Quantum Wasserstein distance based on an optimization over separable states

Géza Tóth[1,2,3,4,5] and József Pitrik[5,6,7]

[1]Theoretical Physics, University of the Basque Country UPV/EHU, ES-48080 Bilbao, Spain

[2]EHU Quantum Center, University of the Basque Country UPV/EHU, Barrio Sarriena s/n, ES-48940 Leioa, Biscay, Spain

[3]Donostia International Physics Center (DIPC), ES-20080 San Sebastián, Spain

[4]IKERBASQUE, Basque Foundation for Science, ES-48011 Bilbao, Spain

[5]Institute for Solid State Physics and Optics, Wigner Research Centre for Physics, HU-1525 Budapest, Hungary

[6]Alfréd Rényi Institute of Mathematics, Reáltanoda u. 13-15., HU-1053 Budapest, Hungary

[7]Department of Analysis and Operations Research, Institute of Mathematics, Budapest University of Technology and Economics, Műegyetem rkp. 3., HU-1111 Budapest, Hungary

We define the quantum Wasserstein distance such that the optimization of the coupling is carried out over bipartite separable states rather than bipartite quantum states in general, and examine its properties. Surprisingly, we find that the self-distance is related to the quantum Fisher information. We present a transport map corresponding to an optimal bipartite separable state. We discuss how the quantum Wasserstein distance introduced is connected to criteria detecting quantum entanglement. We define variance-like quantities that can be obtained from the quantum Wasserstein distance by replacing the minimization over quantum states by a maximization. We extend our results to a family of generalized quantum Fisher information quantities.

Dedicated to the memory of Dénes Petz on the occasion of his 70th birthday.

1 Introduction

A classical Wasserstein distance is a metric between probability distributions μ and ν, induced by the problem of optimal mass transportation [1, 2]. It reflects the minimal effort that is required in order to morph the mass of μ into the mass distribution of ν. Methods based on the theory of optimal transport and advantageous properties of Wasserstein metrics have achieved great success in several important fields of pure mathematics including probability theory [3, 4], theory of (stochastic) partial differential equations [5, 6], variational problems [7, 8] and geometry of metric spaces [9, 10, 11, 12]. In recent years, there have been a lot of results con-

cerning the description of isometries of Wasserstein spaces, too [13, 14, 15, 16, 17, 18, 19]. Moreover, optimal transport and Wasserstein metric are also used as tools in applied mathematics. In particular, there are applications to image and signal processing [20, 21, 22], medical imaging [23, 24] and machine learning [25, 26, 27, 28, 29, 30].

The non-commutative generalization, the so-called quantum optimal transport has been at the center of attention, as it lead to the definition of several new and very useful notions in quantum physics. The first, semi-classical approach of Życzkowski and Slomczynski, has been motivated by applications in quantum chaos [31, 32, 33]. The method of Biane and Voiculescu is related to free probability [34], while the one of Carlen, Maas, Datta, and Rouzé [35, 36, 37, 38, 39] is based on a dynamical interpretation. Caglioti, Golse, Mouhot, and Paul presented an approach based on a static interpretation [40, 41, 42, 43, 44, 45]. Finally De Palma and Trevisan used quantum channels [46], and De Palma, Marvian, Trevisan, and Lloyd defined the quantum earth mover's distance, i.e., the quantum Wasserstein distance order 1 [47], while Bistron, Cole, Eckstein, Friedland and Życzkowski formulated a quantum Wasserstein distance based on an antisymmetric cost function [48, 49, 50].

One of the key results of quantum optimal transport is the definition of the quantum Wasserstein distance [31, 32, 33, 40, 41, 42, 43, 46, 47, 44, 45, 51, 52]. It has the often desirable feature that it is not necessarily maximal for two quantum states orthogonal to each other, which is beneficial, for instance, when performing learning on quantum data [53]. Some of the properties of the new quantities are puzzling, yet point to profound relations between seemingly unrelated fields of quantum physics. For instance, the quantum Wasserstein distance order 2 of the quantum state from itself can be nonzero, while in the classical case the self-distance is always zero. In par-

Géza Tóth: toth@alumni.nd.edu
József Pitrik: pitrik@math.bme.hu

ticular, as we have mentioned, the quantum Wasserstein distance has been defined based on a quantum channel formalism [46], and it has been shown that the square of the self-distance is equal to the Wigner-Yanase skew information of the quantum state [54]. At this point the questions arises: Can other fields of quantum physics help to interpret these results? For example, is it possible to relate the findings above to entanglement theory [55, 56, 57] such that we obtain new and meaningful relations naturally?

Before presenting our results, let us summarize the definitions of the quantum Wasserstein distance.

Definition 1. The square of the distance between two quantum states described by the density matrices ϱ and σ is given by De Palma and Trevisan as [46]

$$
D_{\mathrm{DPT}}(\varrho,\sigma)^2 = \frac{1}{2} \min_{\varrho_{12}} \sum_{n=1}^{N} \mathrm{Tr}[(H_n^T \otimes \mathbb{1} - \mathbb{1} \otimes H_n)^2 \varrho_{12}],
$$
$$
\text{s. t. } \varrho_{12} \in \mathcal{D},
$$
$$
\mathrm{Tr}_2(\varrho_{12}) = \varrho^T,
$$
$$
\mathrm{Tr}_1(\varrho_{12}) = \sigma, \tag{1}
$$

where A^T denotes the matrix transpose of A, and $H_1, H_2, ..., H_N$ are Hermitian operators [1], while $\mathcal{D}$ is the set of density matrices, i.e., Hermitian matrices fulfilling

$$
\varrho_{12} = \varrho_{12}^\dagger, \mathrm{Tr}(\varrho_{12}) = 1, \varrho_{12} \geq 0. \tag{2}
$$

In this approach, there is a bipartite density matrix ϱ_{12}, called coupling, corresponding to any transport map between ϱ and σ, and vice versa, there is a transport map corresponding to any coupling [46]. Moreover, it has been shown that for the self-distance of a state [46]

$$
D_{\mathrm{DPT}}(\varrho,\varrho)^2 = \sum_{n=1}^{N} I_\varrho(H_n) \tag{3}
$$

holds, where the Wigner-Yanase skew information is defined as [54]

$$
I_\varrho(H) = \mathrm{Tr}(H^2 \varrho) - \mathrm{Tr}(H\sqrt{\varrho}H\sqrt{\varrho}). \tag{4}
$$

This profound result connects seemingly two very different notions of quantum physics, as it has been mentioned in the introduction.

The Wasserstein distance has also been defined in a slightly different way.

Definition 2. Golse, Mouhot, Paul and Caglioti defined the square of the distance as [40, 45, 43, 41,

---

[1] We use the convention that in the expression within the trace in Eq. (1), $H_n \otimes \mathbb{1}$ denotes an operator in which H_n acts on subsystem 1 and $\mathbb{1}$ acts on subsystem 2.

42, 44]

$$
D_{\mathrm{GMPC}}(\varrho,\sigma)^2
$$
$$
= \frac{1}{2} \min_{\varrho_{12}} \sum_{n=1}^{N} \mathrm{Tr}[(H_n \otimes \mathbb{1} - \mathbb{1} \otimes H_n)^2 \varrho_{12}],
$$
$$
\text{s. t. } \varrho_{12} \in \mathcal{D},
$$
$$
\mathrm{Tr}_2(\varrho_{12}) = \varrho,
$$
$$
\mathrm{Tr}_1(\varrho_{12}) = \sigma. \tag{5}
$$

In this paper, we will obtain new quantities by restricting the optimization to separable states in the above definitions. We will show that, in this case, the square of the self-distance equals the quantum Fisher information times a constant, while in Eq. (3) it was related to the Wigner-Yanase skew information. The quantum Fisher information is a central quantity in quantum estimation theory and quantum metrology, a field that is concerned with metrological tasks in which the quantumness of the system plays an essential role [58, 59, 60, 61]. Recent findings show that the quantum Fisher information is the convex roof of the variance, apart from a constant factor [62, 63], which allowed, for instance, to derive novel uncertainty relations [64, 65], and will also be used in this article.

The paper is organized as follows. In Sec. 2, we summarize basic facts connected to quantum metrology. In Sec. 3, we summarize entanglement theory. In Sec. 4, we show how to transform the optimization over decompositions of the density matrix into an optimization of an expectation value over separable states. In Sec. 5, we show some applications of these ideas for the Wasserstein distance. We define a novel type of Wasserstein distance based on an optimization over separable states. In Sec. 6, we define variance-like quantities from the Wasserstein distance. In Sec. 7, we discuss how such a Wasserstein distance and the variance-like quantity mentioned above can be used to construct entanglement criteria. In Sec. 8, we introduce further quantities similar to the formulas giving the Wasserstein distance, but they involve the variance of two-body quantities rather than the second moment. In Sec. 9, we consider an optimization over various subsets of the quantum states. In Sec. 10, we extend our ideas to various generalized quantum Fisher information quanitities.

2 Quantum metrology

Before discussing our main results, we review some of the fundamental relations of quantum metrology. A basic metrological task is estimating the small angle θ in a unitary dynamics

$$
U_\theta = \exp(-iH\theta), \tag{6}
$$

where H is the Hamiltonian. The precision is limited by the Cramér-Rao bound as [66, 67, 68, 69, 70, 58,

60, 61, 71, 72, 59, 73]

$$(\Delta\theta)^2 \geq \frac{1}{\nu \mathcal{F}_Q[\varrho, H]}, \qquad (7)$$

where the factor $1/\nu$ in Eq. (7) is the statistical improvement when performing independent measurements on identical copies of the probe state, and the quantum Fisher information is defined by the formula [66, 67, 68, 69, 70]

$$\mathcal{F}_Q[\varrho, H] = 2 \sum_{k,l} \frac{(\lambda_k - \lambda_l)^2}{\lambda_k + \lambda_l} |\langle k|H|l\rangle|^2, \qquad (8)$$

where the density matrix has the eigendecomposition

$$\varrho = \sum_k \lambda_k |k\rangle\langle k|. \qquad (9)$$

The quantum Fisher information is bounded from above by the variance and from below by the Wigner-Yanase skew information as

$$4I_\varrho(H) \leq \mathcal{F}_Q[\varrho, H] \leq 4(\Delta H)_\varrho^2, \qquad (10)$$

where if ϱ is pure then the equality holds for both inequalities [68]. For clarity, we add that the variance is defined as

$$(\Delta H)_\varrho^2 = \langle H^2 \rangle_\varrho - \langle H \rangle_\varrho^2, \qquad (11)$$

where the expectation value is calculated as

$$\langle H \rangle_\varrho = \mathrm{Tr}(\varrho H). \qquad (12)$$

The quantum Fisher information is the convex roof of the variance times four [62, 63, 71]

$$\mathcal{F}_Q[\varrho, H] = 4 \min_{\{p_k, |\Psi_k\rangle\}} \sum_k p_k (\Delta H)_{\Psi_k}^2, \qquad (13)$$

where the optimization is carried out over pure state decompositions of the type

$$\varrho = \sum_k p_k |\Psi_k\rangle\langle\Psi_k|. \qquad (14)$$

For the probabilities $p_k \geq 0$ and $\sum_k p_k = 1$ hold, and the pure states $|\Psi_k\rangle$ are not assumed to be orthogonal to each other.

Apart from the quantum Fisher information, the variance can also be given as a roof [62, 63]

$$(\Delta H)_\varrho^2 = \max_{\{p_k, |\Psi_k\rangle\}} \sum_k p_k (\Delta H)_{\Psi_k}^2. \qquad (15)$$

Note that convex and concave roofs of more complicated expressions can also be computed. For instance,

$$\min_{\{p_k, |\Psi_k\rangle\}} \sum_k p_k \sum_{n=1}^N (\Delta H_n)_{\Psi_k}^2 \geq \frac{1}{4} \sum_{n=1}^N \mathcal{F}_Q[\varrho, H_n], \qquad (16)$$

and the expression

$$\max_{\{p_k, |\Psi_k\rangle\}} \sum_k p_k \sum_{n=1}^N (\Delta H_n)_{\Psi_k}^2 \leq \sum_{n=1}^N (\Delta H_n)_\varrho^2, \qquad (17)$$

are the convex and concave roofs, respectively, of the sum of several variances over the decompositions of ϱ given in Eq. (14). They play a role in the derivation of entanglement conditions [64], and will also appear in our results about the quantum Wasserstein distance. We add that there is an equality in Eq. (17) for $N = 2$ [74, 75].

Finally, we note that the quantum Fisher information can also be given with a minimization over purifications, which has been used, for instance, to study quantum metrology in noisy systems [76, 77, 78]. The relation of this finding to the expression in Eq. (13) is discussed in Ref. [79].

3 Entanglement theory

Next, we review entanglement theory [55, 56, 57]. A bipartite quantum state is separable if it can be given as [80]

$$\sum_k p_k |\Psi_k\rangle\langle\Psi_k| \otimes |\Phi_k\rangle\langle\Phi_k|, \qquad (18)$$

where p_k are for the probabilities, $|\Psi_k\rangle$ and $|\Phi_k\rangle$ are pure quantum states. We will denote the set of separable states by $\mathcal{S}$. The mixture of two separable states is also separable, thus the set $\mathcal{S}$ is convex. If a quantum state cannot be written as Eq. (18) then it is called entangled.

A relevant subset of separable states are the symmetric separable states, which can be given as [81, 82]

$$\sum_k p_k |\Psi_k\rangle\langle\Psi_k| \otimes |\Psi_k\rangle\langle\Psi_k|. \qquad (19)$$

We will denote the set of symmetric separable states by $\mathcal{S}'$. The mixture of two symmetric separable states is also symmetric and separable, thus the set S' is also convex. Clearly

$$\mathcal{S}' \subset \mathcal{S}. \qquad (20)$$

For such states for the expectation value

$$\langle P_s \rangle = 1 \qquad (21)$$

holds, where P_s is the projector to the symmetric subspace defined by the basis vectors $|nn\rangle$ and $(|nm\rangle + |mn\rangle)/\sqrt{2}$ for $n \neq m$. Equivalently, any state in S' fulfills

$$F_{12}\varrho = \varrho F_{12} = \varrho, \qquad (22)$$

where F_{12} is the flip operator for which

$$F_{12}|m\rangle \otimes |n\rangle = |n\rangle \otimes |m\rangle \qquad (23)$$

for all m, n.

The set of quantum states with a positive partial transpose (PPT) $\mathcal{P}$ consists of the states for which

$$\varrho^{T_k} \geq 0 \qquad (24)$$

holds for $k = 1, 2$, where T_k is a partial transposition according to the kth subsystem. $\mathcal{P}$ is clearly a convex set. Moreover, all separable states given in Eq. (18) fulfill Eq. (24), thus

$$\mathcal{P} \supset \mathcal{S} \qquad (25)$$

holds. For two qubits and for a qubit-qutrit system, i.e., for 2×2 and 2×3 systems the set of PPT states equals the set of separable states [83, 84, 85].

It is generally very difficult to decide whether a state is separable or not, while it is very simple to decide whether the condition given in Eq. (24) is fulfilled. Such conditions can be even part of semidefinite programs used to solve various optimization problems (e.g., Ref. [86]). Thus, often the set of PPT quantum states, $\mathcal{P}$ is used instead of the separable states. Since for small systems $\mathcal{P} = \mathcal{S}$, optimization problems over separable states can be solved exactly for those cases. For larger systems, by optimizing over states in $\mathcal{P}$ instead of states in $\mathcal{S}$, we get lower or upper bounds.

Finally, we will define the set of symmetric PPT states $\mathcal{P}'$. States in this set are symmetric, thus Eqs. (21) and (22) hold. For the various sets mentioned above, we have the relation

$$\mathcal{S}' \subset \mathcal{P}' \subset \mathcal{P}, \qquad (26)$$

while for 2×2 and 2×3 systems, we have $P' = S'$.

Convex roofs and concave roofs play also a central role in entanglement theory. The entanglement of formation is defined as a convex roof over pure components of the von Neumann entropy of the reduced state [87, 88]

$$E_F(\varrho) = \min_{\{p_k, |\Psi_k\rangle\}} \sum_k p_k E_E(\Psi_k), \qquad (27)$$

where the optimization is over the decompositions given in Eq. (14), and the entanglement entropy of the pure components is given as

$$E_E(\Psi_k) = S[\text{Tr}_A(\Psi_k)], \qquad (28)$$

where S is the von Neumann entropy. $E_F(\varrho)$ is the minimum entanglement needed to create the state. On the other hand, the entanglement of assistance is obtained as a concave roof [89, 90]

$$E_A(\varrho) = \max_{\{p_k, |\Psi_k\rangle\}} \sum_k p_k E_E(\Psi_k). \qquad (29)$$

The above quantities correspond to the following scenario. Let us assume that the bipartite quantum state ϱ living on parties A and B is realized as the reduced state of a pure state living on parties A, B

and C. Let us assume that party C makes a von Neumann measurement resulting in a state $|\Psi_k\rangle$ on A and B, and it sends the measurement result k to A and B. After repeating this on many copies of the three-partite state, the average entanglement of A and B will be

$$\sum_k p_k E_E(\Psi_k), \qquad (30)$$

where p_k is the probability of the outcome k and and for the $|\Psi_k\rangle$ states Eq. (14) holds. If party C wants to help the parties A and B to have a large average entanglement, then Eq. (30) can reach the entanglement of assistance given in Eq. (29). On the other hand, the average entanglement is always larger than or equal to the entanglement of formation given in Eq. (27).

Next, we will introduce an entanglement condition based on the sum of several variances [91, 92, 93, 94], which will be used later in the article. Let us consider a full set of traceless observables $\{G_n\}_{n=1}^{d^2-1}$ for d-dimensional systems fulfilling

$$\text{Tr}(G_n G_{n'}) = 2\delta_{nn'}. \qquad (31)$$

Any traceless Hermitian observable can be obtained as a linear combination of G_n. In other words, G_n are the SU(d) generators. It is known that for pure states (see e.g., Ref. [94])

$$\sum_{n=1}^{d^2-1} (\Delta G_n)^2 = 2(d-1) \qquad (32)$$

holds. Due to the concavity of the variance it follows that for mixed states we have [94]

$$\sum_{n=1}^{d^2-1} (\Delta G_n)^2 \geq 2(d-1). \qquad (33)$$

Let us now consider a $d \times d$ system. For a product state, $|\Psi\rangle \otimes |\Phi\rangle$, we obtain [91, 94]

$$\sum_{n=1}^{d^2-1} [\Delta(G_n^T \otimes \mathbb{1} - \mathbb{1} \otimes G_n)]_{\Psi \otimes \Phi}^2$$
$$= \sum_{n=1}^{d^2-1} (\Delta G_n^T)_\Psi^2 + \sum_{n=1}^{d^2-1} (\Delta G_n)_\Phi^2$$
$$= 4(d-1), \qquad (34)$$

where in the last inequality we used that for pure states Eq. (32) holds. We also used the fact that if $\{G_n\}_{n=1}^{d^2-1}$ is a full set of observables with the properties mentioned above, then $\{G_n^T\}_{n=1}^{d^2-1}$ is also a full set of such observables. Then, for bipartite separable states given in Eq. (18) [91, 92, 93, 94]

$$\sum_{n=1}^{d^2-1} [\Delta(G_n^T \otimes \mathbb{1} - \mathbb{1} \otimes G_n)]^2 \geq 4(d-1) \qquad (35)$$

holds due to the concavity of the variance. Any state that violates the inequality in Eq. (35) is entangled. The left-hand side of Eq. (35) is zero for the maximally entangled state

$$|\Psi_{\mathrm{me}}\rangle = \frac{1}{\sqrt{d}} \sum_{k=0}^{d-1} |k\rangle|k\rangle. \qquad (36)$$

Thus, we say that the criterion given in Eq. (35) detects entangled states in the vicinity of the state given in Eq. (36).

Let us consider the $d = 2$ case concretely. Let us choose $\{G_n\}_{n=1}^3 = \{\sigma_x, \sigma_y, \sigma_z\}$. Equation (35) can rewritten as

$$\begin{aligned}[\Delta(\sigma_x \otimes \mathbb{1} - \mathbb{1} \otimes \sigma_x)]^2 \\ + [\Delta(\sigma_y \otimes \mathbb{1} + \mathbb{1} \otimes \sigma_y)]^2 \\ + [\Delta(\sigma_z \otimes \mathbb{1} - \mathbb{1} \otimes \sigma_z)]^2 \geq 4, \qquad (37)\end{aligned}$$

where σ_l are Pauli spin matrices defined as

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix},$$
$$\sigma_z = \begin{pmatrix} +1 & 0 \\ 0 & -1 \end{pmatrix}. \qquad (38)$$

The left-hand side of the inequality given in Eq. (37) is zero for the state

$$\frac{1}{\sqrt{2}}(|00\rangle + |11\rangle). \qquad (39)$$

We can have a similar construction with only three operators for any system size. Let us consider the usual angular momentum operators j_x, j_y and j_z, living in a d-dimensional system, fulfilling

$$j_x^2 + j_y^2 + j_z^2 = j(j+1)\mathbb{1}, \qquad (40)$$

where $d = 2j+1$. In particular, let us define the matrix [95]

$$j_z = \mathrm{diag}(-j, -j+1, ..., j-1, j). \qquad (41)$$

Then, we need the ladder operators

$$(j_+)_{m,n} = \delta_{m,n+1}\sqrt{j(j+1) - (j-n)(j+1-n)} \qquad (42)$$

and $j_- = j_+^\dagger$, where $1 \leq m, n \leq d$ and δ_{kl} is the well-known Kronecker delta. Then, we define the x and y components as

$$j_x = \frac{j_+ + j_-}{2}, \qquad j_y = \frac{j_+ - j_-}{2i}. \qquad (43)$$

After clarifying the operators used, we need the uncertainty relation

$$(\Delta j_x)^2 + (\Delta j_y)^2 + (\Delta j_z)^2 \geq j, \qquad (44)$$

which is true for all quantum states. Based on Eq. (44), it can be proved that for separable states

we have [96, 97, 98, 94]

$$\begin{aligned}[\Delta(j_x^T \otimes \mathbb{1} - \mathbb{1} \otimes j_x)]^2 \\ + [\Delta(j_y^T \otimes \mathbb{1} - \mathbb{1} \otimes j_y)]^2 \\ + [\Delta(j_z^T \otimes \mathbb{1} - \mathbb{1} \otimes j_z)]^2 \geq 2j. \qquad (45)\end{aligned}$$

c. f. Eq. (37). The left-hand side of the inequality given in Eq. (45) is zero for the maximally entangled state given in Eq. (36). It is easy to see that Eq. (45) is a tight inequality for separable states, since the $|+j, +j\rangle$ state saturates it.

We now present a simple expression for which the maximum for general states is larger than the maximum for separable quantum states. We know that for separable states [99, 97]

$$\langle(\sigma_x \otimes \mathbb{1} - \mathbb{1} \otimes \sigma_x)^2 + (\sigma_y \otimes \mathbb{1} - \mathbb{1} \otimes \sigma_y)^2\rangle \leq 6, \qquad (46)$$

while the maximum for quantum states is 8 and it is taken by the singlet state

$$\frac{1}{\sqrt{2}}(|01\rangle - |10\rangle). \qquad (47)$$

We add that the product state $|01\rangle_x$ saturates the inequality Eq. (46), where $|.\rangle_x$ is a state given in the x-basis. That is, for a qubit, the basis states in the x-basis are

$$|0\rangle_x = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle),$$
$$|1\rangle_x = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle). \qquad (48)$$

Let us now determine a complementary relation. We will find the minimum for separable states for the left-hand side of Eq. (46). We need to know that for separable states

$$[\Delta(\sigma_x \otimes \mathbb{1} - \mathbb{1} \otimes \sigma_x)]^2 + [\Delta(\sigma_y \otimes \mathbb{1} - \mathbb{1} \otimes \sigma_y)]^2 \geq 2 \quad (49)$$

holds. This can be seen as follows [91, 92]. We need to know that for two-qubit quantum states

$$(\Delta\sigma_x)^2 + (\Delta\sigma_y)^2 \geq 2 \qquad (50)$$

holds. For product states $|\Psi\rangle \otimes |\Phi\rangle$, the left-hand side of Eq. (49) equals the sum of single system variances, which can be bounded from below as

$$(\Delta\sigma_x)_\Psi^2 + (\Delta\sigma_y)_\Psi^2 + (\Delta\sigma_x)_\Phi^2 + (\Delta\sigma_y)_\Phi^2 \geq 4. \quad (51)$$

The bound for separable states given in Eq. (18) is the same as the bound for product states, since the variance is a concave function of the state. The left-hand side of Eq. (49) is zero for the state

$$\frac{1}{\sqrt{2}}(|01\rangle + |10\rangle). \qquad (52)$$

The product state $|11\rangle_x$ saturates the inequality given in Eq. (49) and for that state

$$\begin{aligned}\langle\sigma_x \otimes \mathbb{1} - \mathbb{1} \otimes \sigma_x\rangle = 0, \\ \langle\sigma_y \otimes \mathbb{1} - \mathbb{1} \otimes \sigma_y\rangle = 0 \qquad (53)\end{aligned}$$

hold. Thus, for separable states

$$\langle(\sigma_x \otimes \mathbb{1} - \mathbb{1} \otimes \sigma_x)^2 + (\sigma_y \otimes \mathbb{1} - \mathbb{1} \otimes \sigma_y)^2\rangle \geq 2 \quad (54)$$

holds and it is even a tight inequality for separable states, c. f. Eq. (46).

4 Optimization over the two-copy space

In this section, we first review the formalism that maps the optimization over the decompositions of the density matrix to an optimization of an operator expectation value over bipartite symmetric separable quantum states with given marginals [100] [2]. Then, we show that the same result is obtained if the optimization is carried out over general separable quantum states rather than symmetric separable states.

We start from writing the variance of a pure state $|\Psi\rangle$ as an operator expectation value acting on two copies as

$$(\Delta H)^2_\Psi = \text{Tr}(\Omega|\Psi\rangle\langle\Psi| \otimes |\Psi\rangle\langle\Psi|), \quad (55)$$

where we define the operator

$$\Omega = H^2 \otimes \mathbb{1} - H \otimes H. \quad (56)$$

Based on Eq. (55), the expression of the quantum Fisher information given in Eq. (13) can be rewritten as [62, 100]

$$\begin{aligned} \mathcal{F}_Q[\varrho, H] = \\ 4 \min_{\{p_k, |\Psi_k\rangle\}} \sum_k p_k \text{Tr}[\Omega(|\Psi_k\rangle\langle\Psi_k|)^{\otimes 2}], \\ \text{s. t. } \varrho = \sum_k p_k |\Psi_k\rangle\langle\Psi_k|. \end{aligned} \quad (57)$$

The sum can be moved into the trace and we obtain

$$\begin{aligned} \mathcal{F}_Q[\varrho, H] = \\ 4 \min_{\{p_k, |\Psi_k\rangle\}} \text{Tr}\left\{\Omega\left[\sum_k p_k(|\Psi_k\rangle\langle\Psi_k|)^{\otimes 2}\right]\right\}, \\ \text{s. t. } \varrho = \sum_k p_k |\Psi_k\rangle\langle\Psi_k|. \end{aligned} \quad (58)$$

On the right-hand side in Eq. (58) in the square bracket, we can recognize a mixed state living in the two-copy space

$$\varrho_{12} = \sum_k p_k |\Psi_k\rangle\langle\Psi_k| \otimes |\Psi_k\rangle\langle\Psi_k|, \quad (59)$$

States given in Eq. (59) are symmetric separable states, thus $\varrho_{12} \in \mathcal{S}'$. We can rewrite the optimization in Eq. (58) as an optmization over symmetric separable states given in Eq. (59) as

$$\begin{aligned} \mathcal{F}_Q[\varrho, H] = \min_{\varrho_{12}} 4\text{Tr}(\Omega\varrho_{12}), \\ \text{s. t. } \varrho_{12} \in \mathcal{S}', \\ \text{Tr}_2(\varrho_{12}) = \varrho. \end{aligned} \quad (60)$$

Due to the optimization over symmetric separable states, $\text{Tr}_1(\varrho_{12}) = \varrho$ is fulfilled without adding it as an explicit constraint.

Equation (60) contains an optimization over symmetric separable states, which cannot be computed directly. However, we can consider an optimization over a larger set, the set of symmetric PPT states, which leads to an expression that can be obtained numerically using semidefinite programming [101]

$$\begin{aligned} \mathcal{F}_Q^{(\text{PPT})}[\varrho, H] = \min_{\varrho_{12}} 4\text{Tr}(\Omega\varrho_{12}), \\ \text{s. t. } \varrho_{12} \in \mathcal{P}', \\ \text{Tr}_2(\varrho_{12}) = \varrho. \end{aligned} \quad (61)$$

In general, the relation

$$\mathcal{F}_Q^{(\text{PPT})}[\varrho, H] \leq \mathcal{F}_Q[\varrho, H] \quad (62)$$

holds, while for two qubits we have an equality, since for that system size $\mathcal{S}' = \mathcal{P}'$, as we discussed in Sec. 3.

Let us now change the operator to be optimized, making it permutationally symmetric. Equation (60) can be rewritten as

$$\begin{aligned} \mathcal{F}_Q[\varrho, H] = \min_{\varrho_{12}} 2\text{Tr}[(H \otimes \mathbb{1} - \mathbb{1} \otimes H)^2\varrho_{12}], \\ \text{s. t. } \varrho_{12} \in \mathcal{S}', \\ \text{Tr}_2(\varrho_{12}) = \varrho. \end{aligned} \quad (63)$$

We will now show that the expression in Eq. (63) remains true if we change the set over which we have to optimize to the set of separable states.

Observation 1. The quantum Fisher information can be obtained as an optimization over separable states as

$$\begin{aligned} \mathcal{F}_Q[\varrho, H] = \min_{\varrho_{12}} 2\text{Tr}[(H \otimes \mathbb{1} - \mathbb{1} \otimes H)^2\varrho_{12}], \\ \text{s. t. } \varrho_{12} \in \mathcal{S}, \\ \text{Tr}_2(\varrho_{12}) = \varrho, \\ \text{Tr}_1(\varrho_{12}) = \varrho. \end{aligned} \quad (64)$$

Proof. Using

$$\begin{aligned} \text{Tr}[(H \otimes \mathbb{1} - \mathbb{1} \otimes H)^2\varrho_{12}] \\ = 2\text{Tr}(H^2\varrho) - 2\text{Tr}[(H \otimes H)\varrho_{12}], \end{aligned} \quad (65)$$

---

[2]For a reference in the literature, see Eq. (67) in Ref. [62] and Eqs. (S23)-(S25) in Ref. [100]. In Eqs. (S23)-(S25) in Ref. [100], there are some additional constraints for the expectation values $\langle O_i \rangle_\varrho$, where O_i are some operators. For our discussion, they are not needed.

the right-hand side of Eq. (64) can be reformulated as

$$4\text{Tr}(H^2\varrho) - 4\max_{\varrho_{12}} \text{Tr}[(H \otimes H)\varrho_{12}],$$
$$\text{s. t. } \varrho_{12} \in \mathcal{S},$$
$$\text{Tr}_2(\varrho_{12}) = \varrho,$$
$$\text{Tr}_1(\varrho_{12}) = \varrho. \quad (66)$$

Due to the permutational invariance of $H \otimes H$ and that both marginals must be equal to ϱ, if a separable state given in Eq. (18) maximizes the correlation $\langle H \otimes H \rangle$, then the separable state

$$\sum_k p_k |\Phi_k\rangle\langle\Phi_k| \otimes |\Psi_k\rangle\langle\Psi_k|. \quad (67)$$

also maximizes the correlation. Then, the mixture of the above two separable states also maximize the correlation with the given marginals, which can always be written as

$$\varrho_{12} = \sum_{k=1}^{M} \tilde{p}_k |\tilde{\Psi}_k\rangle\langle\tilde{\Psi}_k| \otimes |\tilde{\Psi}_{\pi_k}\rangle\langle\tilde{\Psi}_{\pi_k}|, \quad (68)$$

where π_k for $k = 1, 2, .., M$ is a permutation of $1, 2, ..., M$ for some M. Based on our arguments, it is sufficient to look for the separable state that maximizes $\langle H \otimes H \rangle$ in the form Eq. (68), and $\tilde{p}_k \geq 0$, $\sum_k \tilde{p}_k = 1$. Then, the correlation can be given as

$$\text{Tr}[(H \otimes H)\varrho_{12}] = \sum_{k=1}^{M} \tilde{p}_k h_k h_{\pi_k}, \quad (69)$$

where we define the expectation values of H on $|\tilde{\Psi}_k\rangle$ as

$$h_k = \text{Tr}(H|\tilde{\Psi}_k\rangle\langle\tilde{\Psi}_k|). \quad (70)$$

Based on the Cauchy-Schwarz inequality we have

$$\sum_k \tilde{p}_k h_k h_{\pi_k} \leq \sqrt{\sum_k \tilde{p}_k h_k^2 \sum_k \tilde{p}_k h_{\pi_k}^2}$$
$$\leq \max\left(\sum_k \tilde{p}_k h_k^2, \sum_k \tilde{p}_k h_{\pi_k}^2\right). \quad (71)$$

Thus, when we maximize $\text{Tr}[(H \otimes H)\varrho_{12}]$ over separable states with the constraints for the marginals, the maximum is taken by a symmetric separable state given in Eq. (59). In this case both inequalities are saturated in Eq. (71). ∎

Based on Eq. (15), we can obtain the variance also as a result of an optimization over a two-copy space.

Observation 2. The variance can also be obtained as an optimization over symmetric separable states as

$$(\Delta H)^2 = \max_{\varrho_{12}} \frac{1}{2}\text{Tr}[(H \otimes \mathbb{1} - \mathbb{1} \otimes H)^2 \varrho_{12}],$$
$$\text{s. t. } \varrho_{12} \in \mathcal{S}',$$
$$\text{Tr}_2(\varrho_{12}) = \varrho,$$
$$\text{Tr}_1(\varrho_{12}) = \varrho. \quad (72)$$

Example 1. Note that if we replace the set of symmetric separable states by the set of separable states in Eq. (72), then we get a different quantity, which can be larger in some cases than $(\Delta H)^2$. Let us see a concrete example. For instance, if $\varrho = \mathbb{1}/2$ then among symmetric separable states, the maximum is 1 and it is attained by the state

$$\varrho_{12} = \frac{1}{4}\left(|00\rangle\langle00| + |11\rangle\langle11| + 2|\Psi^+\rangle\langle\Psi^+|\right), \quad (73)$$

where the Bell state $|\Psi^+\rangle$ is given as

$$|\Psi^+\rangle = \frac{1}{\sqrt{2}}(|01\rangle + |10\rangle). \quad (74)$$

The state given in Eq. (73) can be decomposed into the mixture of symmetric product states as

$$\frac{1}{4}(|\alpha_{+1}\rangle\langle\alpha_{+1}| + |\alpha_{-1}\rangle\langle\alpha_{-1}| + |\alpha_{+i}\rangle\langle\alpha_{+i}| + |\alpha_{-i}\rangle\langle\alpha_{-i}|), \quad (75)$$

where the symmetric product state is defined as

$$|\alpha_q\rangle = (|0\rangle + q|1\rangle) \otimes (|0\rangle + q|1\rangle). \quad (76)$$

Among separable states, the maximum is 2 and it is attained by the state

$$\varrho_{12} = \frac{1}{2}\left(|01\rangle\langle01| + |10\rangle\langle10|\right). \quad (77)$$

Note that ϱ_{12} given in Eq. (77) is not symmetric.

5 Quantum Wasserstein distance based on a separable coupling

Next, we will connect our result in Observation 1 to results available in the literature mentioned in the introduction. We will consider the quantum Wasserstein distance such that the optimization takes place over separable states rather than over general bipartite quantum states. We will consider such modifications of $D_{\text{GMPC}}(\varrho, \sigma)^2$ and $D_{\text{DPT}}(\varrho, \sigma)^2$, and examine their properties.

Our first finding concerning the GMPC distance is the following.

Definition 3. Modifying the definition in Eq. (5) we can define a new type of GMPC distance such that we restrict the optimization over separable states as

$$D_{\text{GMPC,sep}}(\varrho, \sigma)^2$$
$$= \frac{1}{2}\min_{\varrho_{12}} \sum_{n=1}^{N} \text{Tr}[(H_n \otimes \mathbb{1} - \mathbb{1} \otimes H_n)^2 \varrho_{12}],$$
$$\text{s. t. } \varrho_{12} \in \mathcal{S},$$
$$\text{Tr}_1(\varrho_{12}) = \varrho,$$
$$\text{Tr}_2(\varrho_{12}) = \sigma. \quad (78)$$

We can see immediately two relevant properties of the newly defined distance. Observation 1 showed that for $N = 1$ for a given H_1

$$D_{\text{GMPC,sep}}(\varrho, \varrho)^2 = \frac{1}{4}\mathcal{F}_Q[\varrho, H_1] \qquad (79)$$

holds. Moreover, based on Eqs. (5) and (78), we can immediately see that

$$D_{\text{GMPC}}(\varrho, \sigma)^2 \leq D_{\text{GMPC,sep}}(\varrho, \sigma)^2, \qquad (80)$$

since on the left-hand side of Eq. (80) there is a minimization over a larger set of quantum states than on the right-hand side. Based on Eqs. (79) and (80), it follows that for the self-distance for $N = 1$ for a given H_1 we obtain

$$D_{\text{GMPC}}(\varrho, \varrho)^2 \leq \frac{1}{4}\mathcal{F}_Q[\varrho, H_1]. \qquad (81)$$

Now, our goal is to give an optimal transport map (plan) corresponding to an optimal coupling of $D_{GMPC,\text{sep}}(\varrho, \sigma)^2$. Note that there can be several optimal couplings. We look for a CPTP map Φ corresponding to an optimal separable coupling ϱ_{12}. Let us assume that an optimal separable coupling equals the separable state given in Eq. (18). Based on Definition 3, it has the following marginals. We obtain ϱ as in Eq. (14), and for the state σ

$$\sigma = \sum_k p_k |\Phi_k\rangle\langle\Phi_k|, \qquad (82)$$

holds. Let us consider the map [46]

$$\Phi(X) = \sum_k B_k X B_k^\dagger. \qquad (83)$$

where the Kraus operators are given as

$$B_k = \sqrt{p_k} A_k \varrho^{-1/2}, \qquad (84)$$

and ϱ^{-1} is the inverse of ϱ on its support.

For our particular transport problem, let us choose

$$A_k = |\Phi_k\rangle\langle\Psi_k|. \qquad (85)$$

It is clear that Φ is completely positive and it is trace preserving, since

$$\begin{aligned}
\sum_k B_k^\dagger B_k &= \sum_k p_k \varrho^{-1/2} |\Psi_k\rangle\langle\Phi_k|\Phi_k\rangle\langle\Psi_k|\varrho^{-1/2} \\
&= \varrho^{-1/2} \sum_k p_k |\Psi_k\rangle\langle\Psi_k|\varrho^{-1/2} \\
&= \varrho^{-1/2}\varrho\varrho^{-1/2} = \mathbb{1}.
\end{aligned} \qquad (86)$$

Since the map transforms ϱ to σ

$$\begin{aligned}
\Phi(\varrho) &= \sum_k \sqrt{p_k}|\Phi_k\rangle\langle\Psi_k|\varrho^{-1/2}\varrho\varrho^{-1/2}|\Psi_k\rangle\langle\Phi_k|\sqrt{p_k} \\
&= \sum_k p_k |\Phi_k\rangle\langle\Phi_k| = \sigma,
\end{aligned} \qquad (87)$$

the CPTP map Φ given in Eq. (83) gives the optimal transport map we were looking for.

Let us see some properties of the map we have just found. If $|\Psi_k\rangle$ are pairwise orthogonal to each other then

$$B_k = A_k = |\Phi_k\rangle\langle\Psi_k| \qquad (88)$$

holds. In this case, the map can be realized by a von Neumann measurement in the basis given by $\{|\Psi_k\rangle\}$, with a subsequent unitary that transforms $|\Psi_k\rangle$ to $|\Phi_k\rangle$. It is instructive to look at the action of the map on the state

$$\varrho_0 = \sum_k p_k |\Psi_k\rangle\langle\Psi_k|_1 \otimes |\Psi_k\rangle\langle\Psi_k|_2. \qquad (89)$$

Then, we obtain the optimal coupling ϱ_{12} as

$$(\mathbb{1} \otimes \Phi)(\varrho_0) = \sum_k p_k |\Psi_k\rangle\langle\Psi_k|_1 \otimes |\Phi_k\rangle\langle\Phi_k|_2. \qquad (90)$$

Hence, for every map of given by Eqs. (83), (84) and (85) there is a corresponding coupling.

In summary, if we restrict the optimization for separable states, then when computing $D_{\text{GMPC,sep}}(\varrho, \sigma)^2$, for all ϱ and σ, there is a transport map corresponding to all optimal couplings. Note that this was not the case for $D_{\text{GMPC}}(\varrho, \sigma)^2$.

Let us now define another distance based on an optimization over separable states.

Definition 4. Based on the definition of $D_{\text{DPT}}(\varrho, \sigma)^2$ given in Eq. (1), we can also define

$$\begin{aligned}
D_{\text{DPT,sep}}&(\varrho, \sigma)^2 \\
&= \frac{1}{2}\min_{\varrho_{12}} \sum_{n=1}^N \text{Tr}[(H_n^T \otimes \mathbb{1} - \mathbb{1} \otimes H_n)^2 \varrho_{12}], \\
&\text{s. t. } \varrho_{12} \in \mathcal{S}, \\
&\qquad \text{Tr}_2(\varrho_{12}) = \varrho^T, \\
&\qquad \text{Tr}_1(\varrho_{12}) = \sigma.
\end{aligned} \qquad (91)$$

As an important property of $D_{\text{DPT,sep}}(\varrho, \sigma)^2$, we can see that

$$D_{\text{DPT}}(\varrho, \sigma)^2 \leq D_{\text{DPT,sep}}(\varrho, \sigma)^2 \qquad (92)$$

holds, since on the left-hand side of Eq. (92) there is a minimization over a larger set of quantum states than on the right-hand side.

Let us see now, what kind of map corresponds to an optimal separable coupling ϱ_{12}, when we calculate $D_{\text{DPT,sep}}(\varrho, \sigma)^2$. Let us assume that an optimal separable coupling is of the form

$$\varrho_{12} = \sum_k p_k |\Psi_k^*\rangle\langle\Psi_k^*|_1 \otimes |\Phi_k\rangle\langle\Phi_k|_2. \qquad (93)$$

Based on Definition 4, we obtain ϱ as in Eq. (14), and σ as in Eq. (82). It turns out that the map we need is

just $\Phi(\varrho)$ defined in Eq. (83). It is instructive to look at the action of the map on the state

$$\varrho_0^{T_1} = \sum_k p_k |\Psi_k^*\rangle\langle\Psi_k^*|_1 \otimes |\Psi_k\rangle\langle\Psi_k|_2, \qquad (94)$$

where ϱ_0 is defined in Eq. (89). Then, we obtain the optimal coupling ϱ_{12} as

$$(\mathbb{1} \otimes \Phi)(\varrho_0^{T_1}) = \sum_k p_k |\Psi_k^*\rangle\langle\Psi_k^*|_1 \otimes |\Phi_k\rangle\langle\Phi_k|_2. \quad (95)$$

Hence, for every map of given by Eqs. (83), (84) and (85) there is a corresponding coupling.

It is instructive to relate our results to those of Ref. [46]. In Ref. [46], the role of $\varrho_0^{T_1}$ is played by the purification. Moreover, if we compute the self-distance and thus $\varrho = \sigma$, the map given in Eq. (83) is not the identity map. On the other hand, in Ref. [46] the map was the identity map in that case.

Interestingly, the two different distance measures, defined with the transpose and without it, respectively, are equal to each other.

Observation 3. The two quantum Wasserstein distance measures are equal to each other

$$D_{\text{DPT,sep}}(\varrho,\sigma)^2 = D_{\text{GMPC,sep}}(\varrho,\sigma)^2. \qquad (96)$$

Proof. We need to know that for two matrices X and Y acting on a bipartite system

$$\text{Tr}(XY) = \text{Tr}(X^{T_k} Y^{T_k}) \qquad (97)$$

holds for $k = 1, 2$. Then, from Eq. (78) it follows that

$$\begin{aligned}
&D_{\text{GMPC,sep}}(\varrho,\sigma)^2 \\
&= \frac{1}{2} \min_{\varrho_{12}} \sum_{n=1}^N \text{Tr}[(H_n^T \otimes \mathbb{1} - \mathbb{1} \otimes H_n)^2 \varrho_{12}^{T_1}], \\
&\quad \text{s. t. } \varrho_{12}^{T_1} \in \mathcal{S}, \\
&\qquad \text{Tr}_2(\varrho_{12}) = \varrho^T, \\
&\qquad \text{Tr}_1(\varrho_{12}) = \sigma.
\end{aligned} \qquad (98)$$

Then, we arrive at Eq. (91) by noticing that

$$\varrho_{12}^{T_1} \in \mathcal{S} \qquad (99)$$

holds if and only if

$$\varrho_{12} \in \mathcal{S} \qquad (100)$$

holds. ∎

It is instructive to obtain a quantum state that maximizes $D_{\text{DPT,sep}}(\varrho,\varrho)^2$ and $D_{\text{DPT}}(\varrho,\varrho)^2$ for $N = 1$ for a given H_1 as follows. We know that

$$I_\varrho(H_1) \le \mathcal{F}_Q[\varrho, H_1]/4 \le (\Delta H_1)_\varrho^2 \le \frac{1}{4}(h_{\max} - h_{\min})^2 \qquad (101)$$

holds. The first two inequalities are based on Eq. (10). The third one can be obtained as follows. Simple algebra shows that for the state

$$|\Psi_{\text{opt}}\rangle = \frac{1}{\sqrt{2}}(|h_{\min}\rangle + |h_{\max}\rangle), \qquad (102)$$
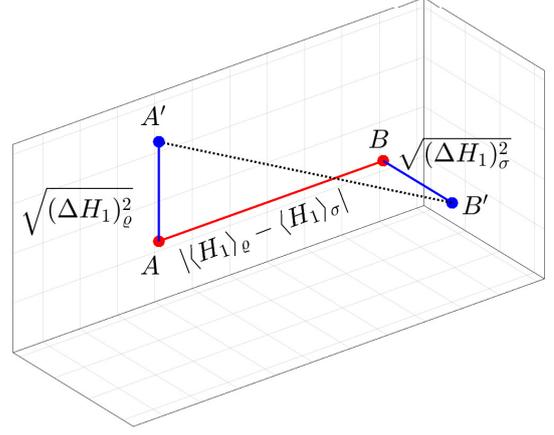

Figure 1: Geometric representation of the quantum Wasserstein distance between a pure state ϱ and a mixed state σ given in Example 2 for $N = 1$ with operator H_1. In Eq. (105), the Wasserstein distance square is $1/2$ times the sum of three terms, corresponding to the two uncertainties and $(\langle H_1\rangle_\varrho - \langle H_1\rangle_\sigma)^2$. Thus, not only the AB distance matters, but the variances computed for the two states. The relevant points are $A(\langle H_1\rangle_\varrho, 0, 0)$, $B(\langle H_1\rangle_\sigma, 0, 0)$, $A'(\langle H_1\rangle_\varrho, \sqrt{(\Delta H_1)_\varrho^2}, 0)$ and $B'(\langle H_1\rangle_\sigma, 0, \sqrt{(\Delta H_1)_\sigma^2})$. The quantum Wasserstein distance equals $1/\sqrt{2}$ times the usual Euclidean distance between A' and B'.

where $|h_{\min}\rangle$ ($|h_{\max}\rangle$) is the eigenstate of H_1 with the minimal eigenvalue $h_{\min}$ (maximal eigenvalue $h_{\max}$) of H_1, the variance $(\Delta H_1)^2$ is maximal, and all inequalities of Eq. (101) are saturated. Hence, the maximal self-distance is achieved by $|\Psi_{\text{opt}}\rangle$

$$\begin{aligned}
&D_{\text{DPT,sep}}(|\Psi_{\text{opt}}\rangle\langle\Psi_{\text{opt}}|, |\Psi_{\text{opt}}\rangle\langle\Psi_{\text{opt}}|)^2 \\
&= D_{\text{DPT}}(|\Psi_{\text{opt}}\rangle\langle\Psi_{\text{opt}}|, |\Psi_{\text{opt}}\rangle\langle\Psi_{\text{opt}}|)^2 \\
&= \frac{1}{4}(h_{\max} - h_{\min})^2.
\end{aligned} \qquad (103)$$

Let us calculate $D_{\text{GMPC,sep}}(\varrho,\sigma)^2$ and the other quantum Wasserstein distance measures for some concrete examples.

Example 2. Let us consider the case when $\varrho = |\Psi\rangle\langle\Psi|$ is a pure state of any dimension and σ is an arbitrary density matrix of the same dimension. Then, when computing the various Wasserstein distance measures between ϱ and σ, the state ϱ_{12} in the optimization is constrained to be the tensor product of the two density matrices. Hence, for the distance from a pure state,

$$\begin{aligned}
&D_{\text{DPT}}(|\Psi\rangle\langle\Psi|, \sigma)^2 \\
&= D_{\text{DPT,sep}}(|\Psi\rangle\langle\Psi|, \sigma)^2 \\
&= D_{\text{GMPC,sep}}(|\Psi\rangle\langle\Psi|, \sigma)^2 \\
&= D_{\text{GMPC}}(|\Psi\rangle\langle\Psi|, \sigma)^2 \\
&= \sum_{n=1}^N \frac{\langle H_n^2\rangle_\varrho + \langle H_n^2\rangle_\sigma}{2} - \langle H_n\rangle_\varrho\langle H_n\rangle_\sigma
\end{aligned} \qquad (104)$$

holds. The last expression in Eq. (104) can be written also as

$$\frac{1}{2}\sum_{n=1}^{N}\left[(\Delta H_n)_\varrho^2 + (\Delta H_n)_\sigma^2 + (\langle H_n\rangle_\varrho - \langle H_n\rangle_\sigma)^2\right],$$
(105)

see Fig. 1.

Example 3. Let us consider the single-qubit states

$$\varrho = \sigma = \frac{1}{2}\mathbb{1}.$$
(106)

Let us take $N = 1$ and $H_1 = \sigma_z$. Then, when computing $D_{\mathrm{GMPC,sep}}(\varrho,\sigma)^2$, the optimum is attained by the separable state

$$\varrho_{12} = \frac{1}{2}|00\rangle\langle 00| + \frac{1}{2}|11\rangle\langle 11|,$$
(107)

and for the self-distance we have

$$D_{\mathrm{GMPC,sep}}\left(\frac{1}{2}\mathbb{1},\frac{1}{2}\mathbb{1}\right)^2 = 0$$
(108)

holds. If H_1 is a different operator then the state ϱ_{12} corresponding to the minimum will be different. It can be obtained from the state given in Eq. (107) with local unitaries. For instance, for $H_1 = \sigma_x$, the optimum is attained by the separable state

$$\varrho_{12} = \frac{1}{2}|00\rangle\langle 00|_x + \frac{1}{2}|11\rangle\langle 11|_x.$$
(109)

Example 4. Let us consider the two-qubit states $\varrho = \sigma = D_p$, where the diagonal state is defined as

$$D_p = \begin{pmatrix} p & 0 \\ 0 & 1-p \end{pmatrix}.$$
(110)

Let us take $N = 1$ and $H_1 = \sigma_z$. When computing $D_{\mathrm{GMPC,sep}}(\varrho,\sigma)^2$, he optimum is reached by the bipartite separable state

$$\varrho_{12} = p|00\rangle\langle 00| + (1-p)|11\rangle\langle 11|.$$
(111)

The self-distance is zero for all p

$$D_{\mathrm{GMPC,sep}}(D_p, D_p)^2 = 0.$$
(112)

Example 5. Let us consider the two single-qubit mixed states

$$\varrho = \frac{1}{2}|1\rangle\langle 1|_x + \frac{1}{2}\cdot\frac{\mathbb{1}}{2},$$
(113)

and

$$\sigma_\phi = e^{-i\frac{\sigma_y}{2}\phi}\varrho e^{+i\frac{\sigma_y}{2}\phi},$$
(114)

ϕ is a real parameter. For $N = 1, H_1 = \sigma_z$ we plotted $D_{\mathrm{DPT,sep}}(\varrho,\sigma_\phi)^2$ and $D_{\mathrm{DPT}}(\varrho,\sigma_\phi)^2$ in Fig. 2. The details of the numerical calculations are in Appendix A. For $\phi = 0$, $\sigma_\phi = \varrho$, hence it this case the two types of distance equal the corresponding types of self-distance of ϱ. That is, based on Eq. (3), we have

$$D_{\mathrm{DPT}}(\varrho,\varrho)^2 = I_\varrho(H_1) = 1 - \frac{\sqrt{3}}{2} \approx 0.1340,$$
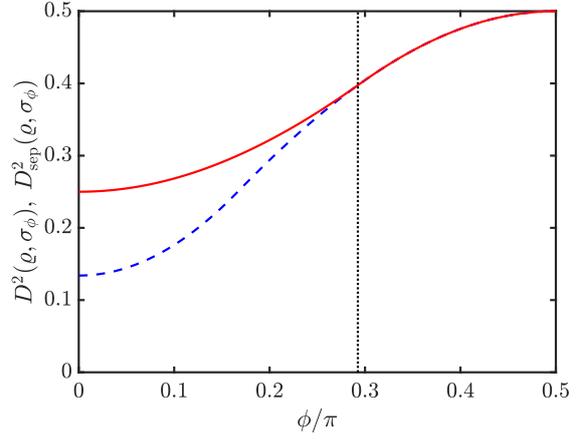(115)

Figure 2: (solid) $D_{\mathrm{DPT,sep}}(\varrho,\sigma_\phi)^2$ and (dashed) $D_{\mathrm{DPT}}(\varrho,\sigma_\phi)^2$ for the states given in Eqs. (113) and (114). The two curves coincide with each other on the right-hand side of the vertical dotted line, where $\phi \geq \phi_0$ and ϕ_0 is given in Eq. (120).

and based on Eq. (79) and taking into account Observation 3, we have

$$D_{\mathrm{DPT,sep}}(\varrho,\varrho)^2 = \frac{\mathcal{F}_Q[\varrho, H_1]}{4} = \frac{1}{4}.$$
(116)

Here, we used the formula giving the quantum Fisher information with the variance for pure states mixed with white noise as [102, 103, 71, 104]

$$\mathcal{F}_Q\left[p|\Psi\rangle\langle\Psi| + (1-p)\frac{\mathbb{1}}{d}, H\right]$$
$$= \frac{p^2}{p + 2(1-p)d^{-1}}4(\Delta H)_\Psi^2,$$
(117)

where d is the dimension of the system. For $\phi = \pi/2$,

$$\sigma_{\pi/2} = \frac{1}{2}|1\rangle\langle 1| + \frac{1}{2}\cdot\frac{\mathbb{1}}{2}.$$
(118)

Numerics show that

$$D_{\mathrm{DPT,sep}}(\varrho,\sigma_\phi)^2 = D_{\mathrm{DPT}}(\varrho,\sigma_\phi)^2$$
(119)

for

$$\phi \geq \phi_0 \approx 0.2946\pi,$$
(120)

while for smaller ϕ an entangled ϱ_{12} is cheaper than a separable one.

We can relate $D_{\mathrm{GMPC,sep}}(\varrho,\sigma)^2$ to the quantum Fisher information.

Observation 4. For the modified GMPC distance defined in Eq. (78) the inequality

$$D_{\mathrm{GMPC,sep}}(\varrho,\sigma)^2 \geq \frac{1}{8}\sum_{n=1}^{N}\mathcal{F}_Q[\varrho, H_n] + \frac{1}{8}\sum_{n=1}^{N}\mathcal{F}_Q[\sigma, H_n]$$
(121)

holds, and for $\varrho = \sigma$ and for $N = 1$ we have equality in Eq. (121).

Proof. We can rewrite the optimization problem in Eq. (78) for separable states as

$$D_{\mathrm{GMPC,sep}}(\varrho, \sigma)^2$$

$$= \frac{1}{2} \min_{\{p_k, |\Psi_k\rangle, |\Phi_k\rangle\}} \sum_{n=1}^{N} \sum_k p_k \Big[(\Delta H_n)^2_{\Psi_k} + (\Delta H_n)^2_{\Phi_k}$$

$$+ (\langle H_n \rangle_{\Psi_k} - \langle H_n \rangle_{\Phi_k})^2 \Big]$$

$$\geq \frac{1}{2} \Big[\min_{\{p_k, |\Psi_k\rangle\}} \sum_{n=1}^{N} \sum_k p_k (\Delta H_n)^2_{\Psi_k}$$

$$+ \min_{\{q_k, |\Phi_k\rangle\}} \sum_{n=1}^{N} \sum_k q_k (\Delta H_n)^2_{\Phi_k} \Big]$$

$$\geq \frac{1}{2} \Big[\sum_{n=1}^{N} \min_{\{p_k, |\Psi_k\rangle\}} \sum_k p_k (\Delta H_n)^2_{\Psi_k}$$

$$+ \sum_{n=1}^{N} \min_{\{q_k, |\Phi_k\rangle\}} \sum_k q_k (\Delta H_n)^2_{\Phi_k} \Big], \qquad (122)$$

where for the first optimization the decomposition of ϱ_{12} is given as Eq. (18), and we have the conditions

$$\mathrm{Tr}_2(\varrho_{12}) = \varrho, \quad \mathrm{Tr}_1(\varrho_{12}) = \sigma. \qquad (123)$$

For the second optimization, the condition with ϱ is given in Eq. (14). For the third optimization, the condition is

$$\sigma = \sum_k q_k |\Phi_k\rangle\langle\Phi_k|, \qquad (124)$$

where for the probabilities $q_k \geq 0$ and $\sum_k q_k = 1$ hold. The fourth and fifth optimization are similar to the second and the third one, however, the order of the sum and the minimization is exchanged. From Eq. (122), the statement follows using the formula giving the quantum Fisher information with the convex roof of the variance in Eq. (13). ∎

In Eq. (122), in the second and third lines we can see that, when computing $D_{\mathrm{GMPC,sep}}(\varrho, \sigma)^2$, the quantity to be minimized is a weighted sum containing the variances of H_n for $|\Psi_k\rangle$ and $|\Phi_k\rangle$, and the expression $(\langle H_n \rangle_{\Psi_k} - \langle H_n \rangle_{\Phi_k})^2$. The two variances, $(\Delta H_n)^2_{\Psi_k}$ and $(\Delta H_n)^2_{\Phi_k}$, are zero if $|\Psi_k\rangle$ and $|\Phi_k\rangle$ are the eigenstates of H_n. The term $(\langle H_n \rangle_{\Psi_k} - \langle H_n \rangle_{\Phi_k})^2$ is zero if

$$\langle H_n \rangle_{\Psi_k} = \langle H_n \rangle_{\Phi_k} \qquad (125)$$

holds.

Based on Example 2 and the proof of Observation 4, we find that the GMPC distance can be obtained with an optimization over separable decompositions as

$$D_{\mathrm{GMPC,sep}}(\varrho, \sigma)^2$$

$$= \min_{\{p_k, |\Psi_k\rangle, |\Phi_k\rangle\}} \sum_k p_k D_{\mathrm{GMPC,sep}}(|\Psi_k\rangle, |\Phi_k\rangle)^2, \qquad (126)$$

where for the optimization the decomposition of ϱ_{12} is given as Eq. (18), and we have the conditions on the marginals given in Eq. (123). This way, we computed the distance for mixed states using the a formula for the distance for pure states and an optimization. An analogous expressions holds also for $D_{\mathrm{DPT,sep}}(\varrho, \sigma)^2$.

Let us see now the consequences of the above observations for the self-distance, which, unlike in the classical case, can be nonzero. It is obtained as

$$D_{\mathrm{GMPC,sep}}(\varrho, \varrho)^2 = \min_{\{p_k, |\Psi_k\rangle\}} \sum_{n=1}^{N} \sum_k p_k (\Delta H_n)^2_{\Psi_k}, \qquad (127)$$

and the condition for the minimization is given in Eq. (14). Then, based on Eq. (80), it follows that the self-distance for the GMPC distance is bounded from above as

$$D_{\mathrm{GMPC}}(\varrho, \varrho)^2 \leq \min_{\{p_k, |\Psi_k\rangle\}} \sum_{n=1}^{N} \sum_k p_k (\Delta H_n)^2_{\Psi_k}. \qquad (128)$$

Let us examine some properties of $D_{\mathrm{GMPC,sep}}(\varrho, \varrho)^2$. From Eq. (127) it follows that for the self-distance for $N = 1$ for a given H_1

$$D_{\mathrm{GMPC,sep}}(\varrho, \varrho)^2 = 0 \qquad (129)$$

holds, if and only if

$$[\varrho, H_1] = 0. \qquad (130)$$

After studying the self-distance, let us look now for similar relations for the distance between two different states. We find that for $N = 1$ for a given H_1

$$D_{\mathrm{GMPC,sep}}(\varrho, \sigma)^2 = 0 \qquad (131)$$

holds, if and only if

$$[\varrho, H_1] = [\sigma, H_1] = 0 \qquad (132)$$

is satisfied and if there is $\{p_k, |\Psi_k\rangle, |\Phi_k\rangle\}$ corresponding to the optimum such that Eq. (125) holds for all k. In this case, if H_1 has a non-degenerate spectrum then $|\Psi_k\rangle$ and $|\Phi_k\rangle$ are the eigenvectors of ϱ, σ and H_1, and thus even $[\varrho, \sigma] = 0$ holds. Finally, based on Observation 3, analogous relations hold for $D_{\mathrm{DPT,sep}}(\varrho, \sigma)^2$.

Let us look for a relation between the distance and the self-distance. From the first inequality in Eq. (122), it follows that for any ϱ and σ

$$D_{\mathrm{GMPC,sep}}(\varrho, \sigma)^2 \geq$$

$$\frac{1}{2} \big[D_{\mathrm{GMPC,sep}}(\varrho, \varrho)^2 + D_{\mathrm{GMPC,sep}}(\sigma, \sigma)^2 \big] \qquad (133)$$

holds.

For a local dimension $d > 2$, the optimization over separable states is difficult to carry out numerically. Thus, it is reasonable to define $D_{\text{DPT,PPT}}(\varrho, \sigma)^2$ and $D_{\text{GMPC,PPT}}(\varrho, \sigma)^2$ that need an optimization over PPT states rather than separable states. It is possible to prove that the two new quantities are equal to each other.

Observation 5. The two quantum Wasserstein distance measures are equal to each other

$$D_{\text{DPT,PPT}}(\varrho, \sigma)^2 = D_{\text{GMPC,PPT}}(\varrho, \sigma)^2. \quad (134)$$

Proof. We have to follow ideas similar to the ones in Observation 3. In particular, we need to use that ϱ_{12} is a PPT quantum state if and only if $\varrho_{12}^{T_1}$ is a PPT quantum state. ∎

In order to define further Wasserstein distance measures based on an optimization over other supersets of separable states, we need to know the separability criterion based on symmetric extensions [105, 106, 107]. A given bipartite state ϱ_{AB} is said to have a $n:m$ symmetric extension if it can be written as the reduced state of a multipartite state $\varrho_{A_1..A_n B_1..B_m}$, which is symmetric under $A_k \leftrightarrow A_l$ for all $k \neq l$ and under $B_{k'} \leftrightarrow B_{l'}$ for all $k' \neq l'$. If we also require that the state is PPT for all bipartitions, then the state has a PPT symmetric extension. The requirement of having a PPT symmetric extension for $n = 1$ and $m = 1$ is equivalent to the PPT condition, while for $n > 1$ or $m > 1$ the condition is stronger. Bipartite separable states have such extensions for arbitrarily large n and m, while the lack of such an extension for some m and n signals the presence of entanglement. In particular, a state is separable if and only if there is an extension for any n and for $m = 1$.

Let us define $D_{\text{DPT,PPT}n}(\varrho, \sigma)^2$ and $D_{\text{GMPC,PPT}n}(\varrho, \sigma)^2$ based on an optimization over quantum states with a PPT symmetric extension for given n and for $m = 1$.

$$D_{\text{DPT,PPT}n}(\varrho, \sigma)^2 \leq D_{\text{DPT,PPT}n'}(\varrho, \sigma)^2 \quad (135)$$

if $n' > n$. Based on Observations 3 and 5, we can also see that

$$D_{\text{DPT,PPT}n}(\varrho, \sigma)^2 = D_{\text{GMPC,PPT}n}(\varrho, \sigma)^2 \quad (136)$$

for every n.

6 Variance-like quantities

When we compare the expression in Eq. (63) defining the quantum Fisher information and the other expression in Eq. (72) defining the variance, we can see that the main difference is that the minimization is replaced by a maximization. We also showed a similar relation between the entanglement of formation given in Eq. (27), and the entanglement of assistance given in Eq. (29). Based on this observation, we can define variance-like quantities from the various forms of

quantum Wasserstein distance by replacing the minimization by maximization.

Such a variance-like quantity can be interpreted in the framework of transport problems as follows. The quantum Wasserstein distance determines the smallest cost possible for the transport problem by a minimization. The variance-like quantities presented in this section determine the largest cost possible for the transport problem. Knowing the largest possible cost is useful when judging how close the cost of a given transport plan is to the optimal cost.

Let us now define the first variance-like quantity.

Definition 5. From the GMPC distance with an optimization restricted over separable states given in Eq. (78), we obtain the following variance like quantity

$$V_{\text{GMPC,sep}}(\varrho, \sigma)$$
$$= \frac{1}{2} \max_{\varrho_{12}} \sum_{n=1}^{N} \text{Tr}[(H_n \otimes \mathbb{1} - \mathbb{1} \otimes H_n)^2 \varrho_{12}],$$
$$\text{s. t.} \quad \varrho_{12} \in \mathcal{S},$$
$$\text{Tr}_2(\varrho_{12}) = \varrho,$$
$$\text{Tr}_1(\varrho_{12}) = \sigma. \quad (137)$$

Clearly, the inequality

$$V_{\text{GMPC,sep}}(\varrho, \sigma) \geq D_{\text{GMPC,sep}}(\varrho, \sigma)^2 \quad (138)$$

holds, since since on the left-hand side of Eq. (138) we maximize over a set of quantum states while on the right-hand side we minimize over the same set.

We can define analogously $V_{\text{DPT,sep}}(\varrho, \sigma)$, $V_{\text{GMPC}}(\varrho, \sigma)$, and $V_{\text{DPT}}(\varrho, \sigma)$, modifying the definition of $D_{\text{DPT,sep}}(\varrho, \sigma)^2$, $D_{\text{GMPC}}(\varrho, \sigma)^2$, and $D_{\text{DPT}}(\varrho, \sigma)^2$, respectively. In all these cases, a relation analogous to the one in Eq. (138) can be obtained.

The value of $V_{\text{GMPC,sep}}(\varrho, \sigma)$ is related to the variance.

Observation 6. The GMPC variance defined in Eq. (137) is bounded from below as

$$V_{\text{GMPC,sep}}(\varrho, \sigma) \geq \frac{1}{2} \sum_{n=1}^{N} (\Delta H_n)^2_\varrho + (\Delta H_n)^2_\sigma. \quad (139)$$

Proof. We can rewrite the expression to be computed for $V_{\text{GMPC,sep}}(\varrho, \sigma)$ as

$$\frac{1}{2} \max_{\varrho_{12} \in \mathcal{S}} \sum_{n=1}^{N} \text{Tr}[(H_n \otimes \mathbb{1} - \mathbb{1} \otimes H_n)^2 \varrho_{12}]$$
$$= \frac{1}{2} \sum_{n=1}^{N} (\Delta H_n)^2_\varrho + (\Delta H_n)^2_\sigma + (\langle H_n \rangle_\varrho - \langle H_n \rangle_\sigma)^2$$
$$- \min_{\varrho_{12} \in \mathcal{S}} C_{\varrho_{12}}, \quad (140)$$

where ϱ_{12} has marginals ϱ and σ, as given in Eq. (123), and we define

$$C_{\varrho_{12}} = \sum_{n=1}^{N} \langle H_n \otimes H_n \rangle_{\varrho_{12}} - \langle H_n \rangle_\varrho \langle H_n \rangle_\sigma. \quad (141)$$

Since the product state $\varrho \otimes \sigma$ is separable and fulfills the conditions on the marginals, it is clear that

$$\min_{\varrho_{12} \in \mathcal{S}} C_{\varrho_{12}} \leq C_{\varrho \otimes \sigma} = 0. \qquad (142)$$

Substituting Eq. (142) into Eq. (140), using the fact $(\langle H_n \rangle_\varrho - \langle H_n \rangle_\sigma)^2 \geq 0$, we can prove the observation. ∎

We can even obtain an upper bound.

Observation 7. The GMPC variance defined in Eq. (137) is bounded from above as

$$V_{\mathrm{GMPC,sep}}(\varrho, \sigma) \leq \sum_{n=1}^{N} (\Delta H_n)^2_\varrho + (\Delta H_n)^2_\sigma$$
$$+ \langle H_n \rangle^2_\varrho + \langle H_n \rangle^2_\sigma. \qquad (143)$$

Proof. We can find an upper bound for the expression to be computed for $V_{\mathrm{GMPC,sep}}(\varrho, \sigma)$ as

$$\frac{1}{2} \max_{\varrho_{12} \in \mathcal{S}} \sum_{n=1}^{N} \mathrm{Tr}[(H_n \otimes \mathbb{1} - \mathbb{1} \otimes H_n)^2 \varrho_{12}]$$
$$\leq \sum_{n=1}^{N} \langle H_n^2 \rangle_\varrho + \langle H_n^2 \rangle_\sigma, \qquad (144)$$

where ϱ_{12} has marginals ϱ and σ, as given in Eq. (123), and we used that

$$H_n^2 \otimes \mathbb{1} + \mathbb{1} \otimes H_n^2 \geq 2 H_n \otimes H_n. \qquad (145)$$

∎

Let us see some concrete examples.

Example 6. Interestingly, the quantity

$$\mathcal{V} := V_{\mathrm{GMPC,sep}}(\varrho, \sigma) \qquad (146)$$

can be larger or smaller than or equal to

$$\overline{\mathcal{V}} := \frac{1}{2} [V_{\mathrm{GMPC,sep}}(\varrho, \varrho) + V_{\mathrm{GMPC,sep}}(\sigma, \sigma)], \quad (147)$$

c. f. Eq. (133). We consider $N = 1$ and $H_1 = \sigma_z$. The three possibilities above are realized by the following states. For the state

$$\varrho = |0\rangle\langle 0|,$$
$$\sigma = \mathrm{diag}(0.25, 0.75), \qquad (148)$$

we have $\mathcal{V} > \overline{\mathcal{V}}$. For the following state

$$\varrho = |0\rangle\langle 0|,$$
$$\sigma = \mathrm{diag}(0.75, 0.25), \qquad (149)$$

we have $\mathcal{V} = \overline{\mathcal{V}}$. Finally, for the following state

$$\varrho = \mathbb{1}/2,$$
$$\sigma = \begin{pmatrix} 0.75 & 0.40 \\ 0.40 & 0.25 \end{pmatrix}, \qquad (150)$$

we have $\mathcal{V} < \overline{\mathcal{V}}$. Let us look now for larger systems. For systems with a local dimension $d = 3$ and for $H_1 = \mathrm{diag}(-1, 0, 1)$, $N = 1$ and for

$$\varrho = \mathbb{1}/3,$$
$$\sigma = \mathrm{diag}(1, 0, 0), \qquad (151)$$

we have $\mathcal{V} < \overline{\mathcal{V}}$. Note that in all examples where ϱ, σ and H_1 were all diagonal, there is an optimal diagonal ϱ_{12}, essentially corresponding to the classical case.

The two different variance-like quantities, defined with the transpose and without it, respectively, are equal to each other.

Observation 8. The two types of quantum Wasserstein variance are equal to each other

$$V_{\mathrm{DPT,sep}}(\varrho, \sigma) = V_{\mathrm{GMPC,sep}}(\varrho, \sigma). \qquad (152)$$

Proof. The proof is analogous to that of Observation 3. ∎

Let us calculate $V_{\mathrm{DPT,sep}}(\varrho, \sigma)$ for some concrete examples.

Example 7. Let us consider the case when $\varrho = |\Psi\rangle\langle\Psi|$ is a pure state of any dimension and σ is an arbitrary density matrix of the same dimension. Then, when computing the various types of quantum Wasserstein distance and quantum Wasserstein variance between ϱ and σ, the state ϱ_{12} in the optimization is constrained to be the tensor product of the two density matrices. Hence, it follows that

$$V_{\mathrm{DPT}}(|\Psi\rangle\langle\Psi|, \sigma) = D_{\mathrm{DPT}}(|\Psi\rangle\langle\Psi|, \sigma)^2 \qquad (153)$$

holds, and analogous equations hold for the quantities $V_{\mathrm{DPT,sep}}(|\Psi\rangle\langle\Psi|, \sigma), V_{\mathrm{GMPC,sep}}(|\Psi\rangle\langle\Psi|, \sigma)$, and $V_{\mathrm{GMPC}}(|\Psi\rangle\langle\Psi|, \sigma)$, where the Wasserstein distance measures for this case are given in Eq. (104).

Example 8. For $\varrho = \sigma = |\Psi\rangle\langle\Psi|$, and for $N = 1$ for a given H_1 we obtain

$$V_{\mathrm{DPT}}(|\Psi\rangle\langle\Psi|, |\Psi\rangle\langle\Psi|) = (\Delta H_1)^2_\Psi. \qquad (154)$$

Example 9. Let us consider the single-qubit states given in Eq. (106). Let us take $N = 1$ and $H_1 = \sigma_z$. Then, when computing $V_{\mathrm{GMPC,sep}}(\varrho, \sigma)$, the optimum is attained by the separable state

$$\varrho_{12} = \frac{1}{2}|01\rangle\langle 01| + \frac{1}{2}|10\rangle\langle 10|, \qquad (155)$$

[c. f. Eq. (107)] and we have

$$V_{\mathrm{GMPC,sep}} \left(\frac{1}{2}\mathbb{1}, \frac{1}{2}\mathbb{1} \right) = 2. \qquad (156)$$

(See also Example 1.) If H_1 is a different operator then the state ϱ_{12} corresponding to the maximum will be different. It can be obtained from the state given in Eq. (155) with local unitaries. For instance, for $H_1 = \sigma_x$, the optimum is reached by the separable state

$$\varrho_{12} = \frac{1}{2}|01\rangle\langle 01|_x + \frac{1}{2}|10\rangle\langle 10|_x, \qquad (157)$$

where $|.\rangle_x$ is a state given in the x-basis.

7 Quantum Wasserstein distance and entanglement criteria

Example 5 highlighted that in certain cases the minimum for separable states is larger than the minimum for general states. In this case, entanglement can help to decrease the quantum Wasserstein distance. In this section, we analyze the relation of the quantum Wasserstein distance and entanglement conditions on the optimal ϱ_{12} couplings.

First, we make the following simple observation.

Observation 9. If

$$D_{\text{DPT,sep}}(\varrho, \sigma)^2 > D_{\text{DPT}}(\varrho, \sigma)^2 \qquad (158)$$

holds then all ϱ_{12} states that minimize the cost for a given ϱ and σ, when computing $D_{\text{DPT}}(\varrho, \sigma)^2$, are entangled. In short, all optimal ϱ_{12} states are entangled. The situation is analogous if

$$D_{\text{GMPC,sep}}(\varrho, \sigma)^2 > D_{\text{GMPC}}(\varrho, \sigma)^2. \qquad (159)$$

Then, all optimal ϱ_{12} states for a given ϱ and σ, when computing $D_{\text{GMPC}}(\varrho, \sigma)^2$, are entangled. Thus, we can even use the quantum Wasserstein distance as an entanglement criterion detecting entanglement in the optimal ϱ_{12} states.

We can even consider conditions with the self-distance. Based on Eqs. (3), (79), and (96), we find that for the $N = 1$ case

$$D_{\text{DPT,sep}}(\varrho, \varrho)^2 > D_{\text{DPT}}(\varrho, \varrho)^2 \qquad (160)$$

is equivalent to a relation between the quantum Fisher information and the Wigner-Yanase skew information

$$\frac{F_Q[\varrho, H_1]}{4} > I_\varrho(H_1). \qquad (161)$$

Thus, the condition in Eq. (161) implies that the optimal ϱ_{12} states, obtained when computing $D_{\text{DPT}}(\varrho, \varrho)^2$, are all entangled.

Let us now use entanglement criteria to construct relations for the quantum Wasserstein distance, that can verify that the coupling ϱ_{12} is entangled. If the inequality given in Eq. (35) holds for separable states, so does the inequality

$$\sum_{n=1}^{d^2-1} \langle (G_n^T \otimes \mathbb{1} - \mathbb{1} \otimes G_n)^2 \rangle \geq 4(d-1), \qquad (162)$$

since the left-hand side of Eq. (162) is never smaller than the left-hand side of Eq. (35). Any state that violates the inequality in Eq. (162) is entangled. It can be shown that Eq. (162) is a tight inequality for separable states as follows. Based on Eq. (34), we see that for pure product states of the form

$$(|\Psi\rangle\langle\Psi|)^T \otimes |\Psi\rangle\langle\Psi|, \qquad (163)$$

for the second moments

$$\sum_{n=1}^{d^2-1} \langle (G_n^T \otimes \mathbb{1} - \mathbb{1} \otimes G_n)^2 \rangle = 4(d-1) \qquad (164)$$

holds since for this state

$$\langle G_n^T \otimes \mathbb{1} - \mathbb{1} \otimes G_n \rangle = 0 \qquad (165)$$

for all n.

Next, we will define a quantum Wasserstein distance related to the entanglement condition in Eq. (162).

Observation 10. Let us consider d-dimensional systems with

$$H_n = G_n \qquad (166)$$

for $n = 1, 2, ..., d^2 - 1$. If

$$D_{\text{DPT}}^{\{G_1, G_2, ..., G_{d^2-1}\}}(\varrho, \sigma)^2 < 2(d-1) \qquad (167)$$

holds, then all optimal ϱ_{12} states are entangled. Here, for clarity, we give explicitly the observables used to define the distance in the superscript.

Clearly, since when calculating $D_{\text{DPT,sep}}(\varrho, \sigma)^2$, we optimize over separable states, we have

$$D_{\text{DPT,sep}}^{\{G_1, G_2, ..., G_{d^2-1}\}}(\varrho, \sigma)^2 \geq 2(d-1). \qquad (168)$$

Thus, independently from what ϱ and σ are, their distance $D_{\text{DPT,sep}}^{\{G_1, G_2, ..., G_{d^2-1}\}}(\varrho, \sigma)^2$ cannot be smaller than a bound. This is true even if $\varrho = \sigma$.

Let us now use another entanglement condition to construct relations for the quantum Wasserstein distance that can verify that the coupling is entangled. We know that the inequality given in Eq. (45) holds for separable states and it is tight. Based on these, we can obtain the following bounds on the quantum Wasserstein distance.

Observation 11. Let us choose the set of operators as

$$\{H_n\} = \{j_x, j_y, j_z\}. \qquad (169)$$

Then, if the inequality

$$D_{\text{DPT}}^{\{j_x, j_y, j_z\}}(\varrho, \sigma)^2 < j \qquad (170)$$

holds, then all optimal ϱ_{12} states are entangled.

Clearly, since when calculating $D_{\text{DPT,sep}}(\varrho, \sigma)^2$, we optimize over separable states, we have

$$D_{\text{DPT,sep}}^{\{j_x, j_y, j_z\}}(\varrho, \sigma)^2 \geq j. \qquad (171)$$

Thus, again, independently from what ϱ and σ are, their distance $D_{\text{DPT,sep}}^{\{j_x, j_y, j_z\}}(\varrho, \sigma)^2$ cannot be smaller than a bound. This is true even if $\varrho = \sigma$.

So far we studied the relation of the quantum Wasserstein distance to entanglement. Next, let us consider the relation of $V_{\text{DPT}}(\varrho, \sigma)$ and $V_{\text{GMPC}}(\varrho, \sigma)$ to entanglement.

Observation 12. If

$$V_{\text{DPT,sep}}(\varrho, \sigma) < V_{\text{DPT}}(\varrho, \sigma) \qquad (172)$$

holds then all optimal ϱ_{12} states for a given ϱ and σ, when computing $V_{\text{DPT}}(\varrho, \sigma)$, are entangled. The situation is analogous if

$$V_{\text{GMPC,sep}}(\varrho, \sigma) < V_{\text{GMPC}}(\varrho, \sigma). \qquad (173)$$

Next, we will determine a set of H_n operators that can be used efficiently to detect entanglement with the quantum Wasserstein variance. For that, we need to know that for separable states the inequality given in Eq. (46) holds.

Observation 13. Let us consider an example with $d = 2$ and $\{H_n\} = \{\sigma_x, \sigma_y\}$. If the condition

$$V_{\text{GMPC}}^{\{\sigma_x, \sigma_y\}}(\varrho, \sigma) > 3 \qquad (174)$$

holds, then all optimal ϱ_{12} states are entangled. Clearly, since when calculating $V_{\text{DPT,sep}}(\varrho, \sigma)$, we optimize over separable states, we have

$$V_{\text{GMPC,sep}}^{\{\sigma_x, \sigma_y\}}(\varrho, \sigma) \le 3. \qquad (175)$$

Let us see now a complementary relation for $D_{\text{GMPC}}(\varrho, \sigma)^2$ and $D_{\text{GMPC,sep}}(\varrho, \sigma)^2$. They will use the same H_n operators that appear in Observation 13. We need to know that for separable states the inequality given in Eq. (54) holds.

Observation 14. Let us consider $d = 2$ and $\{H_n\} = \{\sigma_x, \sigma_y\}$. If

$$D_{\text{GMPC}}^{\{\sigma_x, \sigma_y\}}(\varrho, \sigma)^2 < 1 \qquad (176)$$

then all optimal ϱ_{12} states are entangled. Clearly, since when calculating $D_{\text{DPT,sep}}(\varrho, \sigma)$, we optimize over separable states, we have

$$D_{\text{GMPC,sep}}^{\{\sigma_x, \sigma_y\}}(\varrho, \sigma)^2 \ge 1. \qquad (177)$$

Statements analogous to those of Observations 13 and 14 can be formulated, with identical bounds, for $V_{\text{DPT}}(\varrho, \sigma)^2$, $V_{\text{DPT,sep}}(\varrho, \sigma)^2$, $D_{\text{DPT}}(\varrho, \sigma)^2$, and $D_{\text{DPT,sep}}(\varrho, \sigma)^2$.

8 Optimization of the variance over the two-copy space

In this section, we examine the quantity that we obtain after replacing the second moment by a variance in the optimization in the definition of $D_{\text{GMPC,sep}}(\varrho, \sigma)^2$ given in Eq. (78) and in the definition of $V_{\text{GMPC,sep}}(\varrho, \sigma)$ in Eq. (137). Analogous ideas work also for the other types of quantum Wasserstein distance and quantum Wasserstein variance defined before. We will show that such quantities have interesting properties.

Definition 6. After replacing the second moment by a variance in the optimization in the definition of $D_{\text{GMPC,sep}}(\varrho, \sigma)^2$ given in Eq. (78), we define

$$\tilde{D}_{\text{GMPC,sep}}(\varrho, \sigma)^2$$
$$= \frac{1}{2} \min_{\varrho_{12}} \sum_{n=1}^{N} [\Delta(H_n \otimes \mathbb{1} - \mathbb{1} \otimes H_n)]_{\varrho_{12}}^2,$$
$$\text{s. t. } \varrho_{12} \in \mathcal{S},$$
$$\text{Tr}_2(\varrho_{12}) = \varrho,$$
$$\text{Tr}_1(\varrho_{12}) = \sigma. \qquad (178)$$

Let us see some properties of the quantity we have just introduced. In general, for mixed states,

$$\tilde{D}_{\text{GMPC,sep}}(\varrho, \sigma)^2 = D_{\text{GMPC,sep}}(\varrho, \sigma)^2$$
$$- \frac{1}{2} \sum_n (\langle H_n \rangle_\varrho - \langle H_n \rangle_\sigma)^2 \qquad (179)$$

holds, hence clearly

$$\tilde{D}_{\text{GMPC,sep}}(\varrho, \sigma)^2 \le D_{\text{GMPC,sep}}(\varrho, \sigma)^2. \qquad (180)$$

Due to the relation in Eq. (179), for the self-distance we have

$$\tilde{D}_{\text{GMPC,sep}}(\varrho, \varrho)^2 = D_{\text{GMPC,sep}}(\varrho, \varrho)^2, \qquad (181)$$

where $D_{\text{GMPC,sep}}(\varrho, \sigma)^2$ is given in Eq. (78).

We can write the expression to be optimized as

$$\sum_{n=1}^{N} [\Delta(H_n \otimes \mathbb{1} - \mathbb{1} \otimes H_n)]_{\varrho_{12}}^2$$
$$= \sum_{n=1}^{N} [(\Delta H_n)_\varrho^2 + (\Delta H_n)_\sigma^2] - 2C_{\varrho_{12}}, \quad (182)$$

where ϱ_{12} has marginals ϱ and σ, as given in Eq. (123), and we define $C_{\varrho_{12}}$ as in Eq. (141). For pure states, $\varrho = |\Psi\rangle\langle\Psi|$ and $\sigma = |\Phi\rangle\langle\Phi|$, we have $C = 0$, and hence

$$\tilde{D}_{\text{GMPC,sep}}(|\Psi\rangle\langle\Psi|, |\Phi\rangle\langle\Phi|)^2$$
$$= \frac{1}{2}\Big[\tilde{D}_{\text{GMPC,sep}}(|\Psi\rangle\langle\Psi|, |\Psi\rangle\langle\Psi|)^2$$
$$+ \tilde{D}_{\text{GMPC,sep}}(|\Phi\rangle\langle\Phi|, |\Phi\rangle\langle\Phi|)^2\Big]. \quad (183)$$

We can present lower bounds on the distance.

Observation 15. The modified GMPC distance defined in Eq. (178) is bounded from below as

$$\tilde{D}_{\text{GMPC,sep}}(\varrho, \sigma)^2 \ge \frac{1}{8}\sum_{n=1}^{N} \mathcal{F}_Q[\varrho, H_n] + \frac{1}{8}\sum_{n=1}^{N} \mathcal{F}_Q[\sigma, H_n],$$
$$(184)$$

while for $\varrho = \sigma$ we have equality for $N = 1$ in Eq. (184).

Proof. We can rewrite the optimization problem in Eq. (178) based on Eqs. (122) and (179) as

$$\tilde{D}_{\text{GMPC,sep}}(\varrho,\sigma)^2$$

$$= \frac{1}{2} \min_{\{p_k,|\Psi_k\rangle,|\Phi_k\rangle\}} \sum_{n=1}^{N} \sum_k p_k \Big[(\Delta H_n)^2_{\Psi_k} + (\Delta H_n)^2_{\Phi_k}$$

$$+ (\langle H_n\rangle_{\Psi_k} - \langle H_n\rangle_{\Phi_k})^2 - (\langle H_n\rangle_\varrho - \langle H_n\rangle_\sigma)^2 \Big]$$

$$\geq \frac{1}{2} \Big[\min_{\{p_k,|\Psi_k\rangle\}} \sum_{n=1}^{N} \sum_k p_k (\Delta H_n)^2_{\Psi_k}$$

$$+ \min_{\{q_k,|\Phi_k\rangle\}} \sum_{n=1}^{N} \sum_k q_k (\Delta H_n)^2_{\Phi_k} \Big]$$

$$\geq \frac{1}{2} \Big[\sum_{n=1}^{N} \min_{\{p_k,|\Psi_k\rangle\}} \sum_k p_k (\Delta H_n)^2_{\Psi_k}$$

$$+ \sum_{n=1}^{N} \min_{\{q_k,|\Phi_k\rangle\}} \sum_k q_k (\Delta H_n)^2_{\Phi_k} \Big], \quad (185)$$

where we used in the first inequality that for real x_k

$$\sum_k p_k x_k^2 \geq \left(\sum_k p_k x_k \right)^2 \quad (186)$$

holds due to the fact that $f(x) = x^2$ is convex. In Eq. (185), in the first optimization the decomposition of ϱ_{12} is given as Eq. (18), and we have the conditions given in Eq. (123). For the second optimization, the condition is Eq. (14). For the third optimization, the condition is given in Eq. (124), where for the probabilities $q_k \geq 0$ and $\sum_k q_k = 1$ hold. The fourth and fifth optimization are similar to second and the third one, however, the order of the sum and the minimization is exchanged. From Eq. (185), the statement follows using the formula that obtains the quantum Fisher information with a convex roof of the variance given in Eq. (13). ∎

We can define another variance-like quantity.

Definition 7. Analogously, we can define the quantity that we obtain after replacing the second moment by a variance in the optimization in the definition of $V_{\text{GMPC,sep}}(\varrho,\sigma)^2$ in Eq. (137), as

$$\tilde{V}_{\text{GMPC,sep}}(\varrho,\sigma)$$

$$= \frac{1}{2} \max_{\varrho_{12}} \sum_{n=1}^{N} [\Delta(H_n \otimes \mathbb{1} - \mathbb{1} \otimes H_n)]^2_{\varrho_{12}},$$

$$\text{s. t. } \varrho_{12} \in \mathcal{S},$$
$$\text{Tr}_2(\varrho_{12}) = \varrho,$$
$$\text{Tr}_1(\varrho_{12}) = \sigma. \quad (187)$$

Let us see now some properties of $\tilde{V}_{\text{GMPC,sep}}(\varrho,\sigma)$. In general,

$$\tilde{V}_{\text{GMPC,sep}}(\varrho,\sigma) = V_{\text{GMPC,sep}}(\varrho,\sigma)$$

$$- \frac{1}{2} \sum_n (\langle H_n\rangle_\varrho - \langle H_n\rangle_\sigma)^2 \quad (188)$$

holds, hence clearly

$$\tilde{V}_{\text{GMPC,sep}}(\varrho,\sigma) \leq V_{\text{GMPC,sep}}(\varrho,\sigma). \quad (189)$$

Analogously to Eq. (183), for pure states

$$\tilde{V}_{\text{GMPC,sep}}(|\Psi\rangle\langle\Psi|,|\Phi\rangle\langle\Phi|)$$

$$= \frac{1}{2} \Big[\tilde{V}_{\text{GMPC,sep}}(|\Psi\rangle\langle\Psi|,|\Psi\rangle\langle\Psi|)$$

$$+ \tilde{V}_{\text{GMPC,sep}}(|\Phi\rangle\langle\Phi|,|\Phi\rangle\langle\Phi|) \Big] \quad (190)$$

holds.

Example 10. Let us consider the case when $\varrho = |\Psi\rangle\langle\Psi|$ is a pure state of any dimension and σ is an arbitrary density matrix of the same dimension. Then, when computing the various types of quantum Wasserstein distance and quantum Wasserstein variance between ϱ and σ, the state ϱ_{12} in the optimization is constrained to be the tensor product of the two density matrices. Hence, based on Eqs. (179) and (188), it follows that

$$\tilde{D}_{\text{GMPC,sep}}(|\Psi\rangle\langle\Psi|,\sigma) =$$

$$\tilde{V}_{\text{GMPC,sep}}(|\Psi\rangle\langle\Psi|,\sigma) = \frac{1}{2} \sum_{n=1}^{N} (\Delta H_n)^2_\Psi + (\Delta H_n)^2_\sigma$$

$$\quad (191)$$

holds, c. f. Observation 6.

Finally, one can define similarly $\tilde{D}_{\text{DPT,sep}}(\varrho,\sigma)^2$, $\tilde{D}_{\text{GMPC}}(\varrho,\sigma)^2$, $\tilde{D}_{\text{DPT}}(\varrho,\sigma)^2$, $\tilde{V}_{\text{DPT,sep}}(\varrho,\sigma)$, $\tilde{V}_{\text{GMPC}}(\varrho,\sigma)$, and $\tilde{V}_{\text{DPT}}(\varrho,\sigma)$.

9 Optimization over other subsets of physical states

So far we considered Wasserstein distance based on an optimization over all bipartite physical states, separable states, PPT states, and states with a PPT symmetric extension considered in Sec. 5. In this section we examine other convex sets of quantum states. We will also discuss some relevant couplings. By optimizing over a convex set different from the ones we have considered, we will obtain a Wasserstein distance with a different self-distance.

Let us consider the set of couplings for which the quantum discord is zero [108, 109, 110]. If we assume that

$$\text{Tr}_2(\varrho_{12}) = \varrho,$$
$$\text{Tr}_1(\varrho_{12}) = \sigma \quad (192)$$

hold, then an element of the set is of the form

$$\varrho_{12} = \sum_k \lambda_k |k\rangle\langle k| \otimes \sigma_k, \quad (193)$$

where the eigendecomposition of ϱ is given in Eq. (9) and

$$\sigma = \sum_k \lambda_k \sigma_k. \tag{194}$$

Clearly, such states given in Eq. (193) form a convex set. Such states are called classical-quantum since in subsystem 1 we have a mixture of states that are pairwise orthogonal to each other [110]. We will denote the set of such states as $\mathcal{C}_1$. For such couplings, the map given in Eq. (83) needs only a von Neumann measurement.

Another possibility is the quantum-classical states, which we will denote by $\mathcal{C}_2$. Such states are of the form

$$\sum_k \lambda'_k \varrho_k \otimes |k'\rangle\langle k'|, \tag{195}$$

where the eigendecomposition of σ is

$$\sigma = \sum_k \lambda'_k |k'\rangle\langle k'|, \tag{196}$$

and for the density matrices ϱ_k

$$\varrho = \sum_k \lambda'_k \varrho_k \tag{197}$$

holds. A minimization over $\mathcal{C}_1$ or $\mathcal{C}_2$ will lead to a larger value than a minimization over separable couplings. An optimization over $\mathcal{C}_1$ or $\mathcal{C}_2$ can efficiently be carried out using semidefinite programming for any system size. We can also consider the set of states that are classical-classical, which are the members of both $\mathcal{C}_1$ and $\mathcal{C}_2$.

Let us consider now the case of the GMPC self-distance, when $\varrho = \sigma$. Then, a relevant coupling which is the elements of $\mathcal{C}_1$ and $\mathcal{C}_2$ is

$$\varrho_{cc} = \sum_k \lambda_k |k\rangle\langle k| \otimes |k\rangle\langle k|, \tag{198}$$

where the eigendecomosition of ϱ is given in Eq. (9). For the coupling in Eq. (198), for $N = 1$ the equals

$$\sum_k \lambda_k (\Delta H_1)^2_{|k\rangle} \leq (\Delta H_1)^2_\varrho. \tag{199}$$

So far, we have been obtaining results for the GMPC distance. Analogous statements hold for the DPT distance.

Another relevant case is the product state coupling

$$\varrho_{12} = \varrho \otimes \sigma. \tag{200}$$

Then, we can define the distances given in Eqs. (5) and (1) for product states as

$$D_{\mathrm{GMPC,prod}}(\varrho, \sigma)^2 = D_{\mathrm{DPT,prod}}(\varrho, \sigma)^2$$
$$= \frac{1}{2} \sum_{n=1}^N \mathrm{Tr}[(H_n \otimes \mathbb{1} - \mathbb{1} \otimes H_n)^2 \varrho \otimes \sigma]$$
$$= \frac{1}{2} \sum_{n=1}^N \left[(\Delta H_n)^2_\varrho + (\Delta H_n)^2_\sigma + (\langle H_n\rangle_\varrho - \langle H_n\rangle_\sigma)^2 \right], \tag{201}$$

Set of quantum states	GMPC self-distance	DPT self-distance
General quantum states	$\leq \mathcal{F}_Q[\varrho, H_1]/4$	$I_\varrho(H_1)$
PPT states, PPTn states	$\leq \mathcal{F}_Q[\varrho, H_1]/4$	$\leq \mathcal{F}_Q[\varrho, H_1]/4$
Separable states	$\mathcal{F}_Q[\varrho, H_1]/4$	$\mathcal{F}_Q[\varrho, H_1]/4$
ϱ_{cc} given in Eq. (198)	$\sum_k \lambda_k (\Delta H_1)^2_{|k\rangle}$	$\sum_k \lambda_k (\Delta H_1)^2_{|k\rangle}$
Product state	$(\Delta H_1)^2_\varrho$	$(\Delta H_1)^2_\varrho$

Table 1: Self-distance obtained for the Wasserstein distance considering an optimization over various subsets of the bipartite quantum states for $N = 1$.

c. f. Eq. (105). For the self-distance, the relation

$$D_{\mathrm{GMPC,prod}}(\varrho, \varrho) = D_{\mathrm{DPT,prod}}(\varrho, \varrho) = \sum_{n=1}^N (\Delta H_n)^2_\varrho \tag{202}$$

holds.

In Table 1, we summarized the self-distances obtained for the Wasserstein distance considering an optimization over various subsets of the bipartite quantum states and $N = 1$. For the quantities in the Table, the inequality given in Eq. (10) holds. As expected, a minimization over a larger set will not lead to a larger value and often will lead to a smaller value.

We can consider an optimization over other convex sets of states. For instance, a convex set can be characterzied by constraints like

$$(\Delta A_k)^2 \geq c_k \tag{203}$$

or by the linear constraints

$$\langle B_k \rangle \geq d_k, \tag{204}$$

where A_k and B_k are operators, c_k and d_k are constants. Other possibilities are the convex set of states with negativity not larger than a given bound [86], and the convex set of states not violating certain entanglement conditions. These conditions can be incorporated into the numerical optimization. We can also consider the convex set of states with a local hidden variable model [55, 56, 57].

10 Alternative definition of the Wasserstein distance such that the self-distance equals various generalized quantum Fisher information quantities

In this section, we will modify the definition of the Wasserstein distance such that the self-distance

Accepted in 〈 〉uantum 2023-09-19, click title to verify. Published under CC-BY 4.0.

17

equals a quantity different from the ones we considered so far. In particular, we would like that it equals various generalized quantum Fisher information quantities. The Wigner-Yanase skew information and the quantum Fisher information are two members of this family.

The basic idea of Refs. [111, 112] is that for each standard matrix monotone function $f : \mathbb{R}^+ \to \mathbb{R}^+$, a generalized variance and a corresponding quantum Fisher information are defined. The notion standard means that f must satisfy

$$f(1) = 1, \tag{205a}$$

$$f(t) = tf(t^{-1}). \tag{205b}$$

For a review on generalized variances, generalized quantum Fisher information quantities and covariances see Ref. [113]. Moreover, it is also useful to define the mean based on f as

$$m_f(a, b) = af\left(\frac{b}{a}\right) \tag{206}$$

and use it instead of f. The normalization condition given in Eq. (205a) corresponds to the condition

$$m_f(a, a) = a \tag{207}$$

for the means. The requirement given in Eq. (205b) corresponds to

$$m_f(a, b) = m_f(b, a). \tag{208}$$

A list of generalized quantum Fisher information quantities generated by various well-known means $m_f(a, b)$ can be found in Refs. [111, 112].

Using the normalization suggested by Ref. [62], we arrive at a family of generalized quantum Fisher information quantities $F_Q^f[\varrho, H]$ such that (i) for pure states, we have

$$\mathcal{F}_Q^f[|\Psi\rangle, H] = 4(\Delta H)_\Psi^2. \tag{209}$$

(ii) For mixed states, $F_Q^f[\varrho, H]$ is convex in the state.

The generalized variance $\text{var}_\varrho^f(H)$ fulfills the following two requirements. (i) For pure states, the generalized variance equals the usual variance

$$\text{var}_\Psi^f(H) = (\Delta H)_\Psi^2. \tag{210}$$

(ii) For mixed states, $\text{var}_\varrho^f(H)$ is concave in the state.

A family of generalized quantum Fisher information and generalized variance fulfilling the above requirements are [62]

$$\mathcal{F}_Q^f[\varrho, H] = 2 \sum_{k,l} \frac{m_f(1, 0)}{m_f(\lambda_k, \lambda_l)} (\lambda_k - \lambda_l)^2 |H_{kl}|^2, \tag{211a}$$

$$\text{var}_\varrho^f(H) = \frac{1}{2} \sum_{k,l} \frac{m_f(\lambda_k, \lambda_l)}{m_f(1, 0)} |H_{kl}|^2$$
$$- \frac{1}{2m_f(1, 0)} \left| \sum \lambda_k H_{kk} \right|^2, \tag{211b}$$

where the matrix elements of H in the eigenbasis of ϱ are denoted as

$$H_{kl} = \langle k|H|l\rangle. \tag{212}$$

$\mathcal{F}_Q^f[\varrho, H]/4$ has been called metric-adjusted skew information [114].

Note that the definition of the variance in Eq. (211b) requires that $m_f(1, 0) \equiv f(0)$ is nonzero. In such cases f is called regular [115].

The usual quantum Fisher information and the usual variance corresponds to

$$f_{\max}(x) = \frac{1 + x}{2}, \tag{213}$$

and the arithmetic mean

$$m_{f_{\max}}(a, b) = \frac{a + b}{2}. \tag{214}$$

Note that $f_{\max}(x)$ is the largest among standard matrix monotone functions. Due to this, $\mathcal{F}_Q^{f_{\max}}[\varrho, H]$ is the largest among $\mathcal{F}_Q^f[\varrho, H]$ and $\text{var}_\varrho^{f_{\max}}(H)$ is the smallest among $\text{var}_\varrho^f(H)$.

The Wigner-Yanase skew information corresponds to

$$f_{\text{WY}}(x) = \frac{(\sqrt{x} + 1)^2}{4}, \tag{215}$$

and the mean

$$m_{f_{\text{WY}}}(a, b) = \frac{(\sqrt{a} + \sqrt{b})^2}{4}. \tag{216}$$

Note that we get 4 times the usual Wigner-Yanase skew information due to the chosen normalization

$$\mathcal{F}_Q^{f_{\text{WY}}}[\varrho, H] = 4I_\varrho(H). \tag{217}$$

We now show a method to express the various generalized quantum Fisher information quantities with each other.

Observation 16. Let us define for given f the following matrix in the eigenbasis of ϱ

$$(X_f)_{kl} = \langle k|X|l\rangle = \sqrt{\frac{m_f(1, 0)}{m_f(\lambda_k, \lambda_l)}}. \tag{218}$$

Then, any generalized quantum Fisher information can be expressed as

$$\mathcal{F}_Q^{f_1}[\varrho, H] = \mathcal{F}_Q^{f_2}[\varrho, Q_{f_1, f_2} \circ H], \tag{219}$$

where "$\circ$" denotes element-wise or Hadamard product defined as

$$A \circ B = \sum_{k,l} \langle k|A|l\rangle\langle k|B|l\rangle|k\rangle\langle l|, \tag{220}$$

where $|k\rangle$ and $|l\rangle$ are the eigenvectors of the density matrix, and the coefficient for converting one type of quantum Fisher information into another ons is given as

$$(Q_{f_1, f_2})_{kl} = \begin{cases} \frac{(X_{f_1})_{kl}}{(X_{f_2})_{kl}}, & \text{if } \lambda_k \neq \lambda_l, \\ 0, & \text{if } \lambda_k = \lambda_l. \end{cases} \tag{221}$$

Proof. The statement can be verified by direct comparison of the definition of the quantum Fisher information given in Eqs. (211a) and (219). ∎

It is interesting to compute the matrix needed for $f_1 = f_{\max}$ and $f_2 = f_{WY}$. We obtain, in the basis of the eigenvectors of ϱ,

$$(Q_{f_{\max}, f_{WY}})_{kl} = \begin{cases} \frac{(X_{f_{\max}})_{kl}}{(X_{f_{WY}})_{kl}} = \frac{\sqrt{\lambda_k} + \sqrt{\lambda_l}}{\sqrt{\lambda_k + \lambda_l}}, & \text{if } \lambda_k \neq \lambda_l, \\ 0, & \text{if } \lambda_k = \lambda_l. \end{cases} \tag{222}$$

Here note that in Eq. (222), the denominator of the fraction with λ_k and λ_l is positive if $\lambda_k \neq \lambda_l$. With the matrix in Eq. (222), we can use the Wigner-Yanase skew information to obtain a the quantum Fisher information

$$\mathcal{F}_Q[\varrho, H] = 4 I_\varrho(Q_{f_{\max}, f_{WY}} \circ H). \tag{223}$$

Next, we compute the matrix needed for $f_1 = f_{WY}$ and $f_2 = f_{\max}$. We obtain, in the basis of the eigenvectors of ϱ,

$$(Q_{f_{WY}, f_{\max}})_{kl} = \begin{cases} \frac{(X_{f_{WY}})_{kl}}{(X_{f_{\max}})_{kl}} = \frac{\sqrt{\lambda_k + \lambda_l}}{\sqrt{\lambda_k} + \sqrt{\lambda_l}}, & \text{if } \lambda_k \neq \lambda_l, \\ 0, & \text{if } \lambda_k = \lambda_l. \end{cases} \tag{224}$$

Then, we can also obtain the Wigner-Yanase skew information with the quantum Fisher information as

$$I_\varrho(H) = \frac{1}{4} \mathcal{F}_Q[\varrho, Q_{f_{WY}, f_{\max}} \circ H]. \tag{225}$$

Next, we show how the various generalized quantum Fisher information quantities can be expressed as a convex roof over the decompositions of the density matrix.

Observation 17. The various generalized quantum Fisher information quantities can be expressed as a convex roof as

$$\mathcal{F}_Q^f[\varrho, H] = 4 \min_{\{p_k, |\Psi_k\rangle\}} \sum_k p_k [\Delta(Y_f \circ H)]_{\Psi_k}^2, \tag{226}$$

where the optimization is over pure state decompositions given in Eq. (14), and in the basis of the eigenvectors of ϱ we define

$$(Y_f)_{kl}$$
$$= \begin{cases} \frac{(X_f)_{kl}}{(X_{f_{\max}})_{kl}} = \sqrt{\frac{m_f(1,0)}{m_f(\lambda_k, \lambda_l)}}(\lambda_k + \lambda_l), & \text{if } \lambda_k \neq \lambda_l, \\ 0, & \text{if } \lambda_k = \lambda_l. \end{cases} \tag{227}$$

Proof. It follows from Observation 16 defining $(X_f)_{kl}$ and the definition of the quantum Fisher information with the convex roof of the variance given in Eq. (13). ∎

Next, we show how the various generalized quantum Fisher information quantities can be expressed as an optimization in the two-copy space.

Observation 18. The generalized quantum Fisher information can be obtained as an optimization over separable states

$$\mathcal{F}_Q^f[\varrho, H] = \min_{\varrho_{12}} 2\text{Tr}[(Y_f \circ H \otimes \mathbb{1} - \mathbb{1} \otimes Y_f \circ H)^2 \varrho_{12}],$$
$$\text{s. t. } \varrho_{12} \in \mathcal{S},$$
$$\text{Tr}_2(\varrho_{12}) = \varrho,$$
$$\text{Tr}_1(\varrho_{12}) = \varrho. \tag{228}$$

Proof. It follows from Observation 1 defining the quantum Fisher information with an optimization over bipartite separable quantum states and Observation 16 defining $(X_f)_{kl}$. ∎

Let us calculate a concrete example. For instance, for obtaining the Wigner-Yanase skew information times four as an optimization over separable states, we should use

$$(Y_{f_{WY}})_{kl} = (Q_{f_{WY}, f_{\max}})_{kl}, \tag{229}$$

where $Q_{f_{WY}, f_{\max}}$ is defined in Eq. (224).

Observation 19. We can obtain the various quantum Fisher information quantities as an optimization over general quantum states, rather than over separable states, as

$$\mathcal{F}_Q^f[\varrho, H](\varrho)$$
$$= \min_{\varrho_{12}} 2\text{Tr}\{[(Z_f \circ H)^T \otimes \mathbb{1} - \mathbb{1} \otimes Z_f \circ H]^2 \varrho_{12}\},$$
$$\text{s. t. } \varrho_{12} \in \mathcal{D},$$
$$\text{Tr}_2(\varrho_{12}) = \varrho^T,$$
$$\text{Tr}_1(\varrho_{12}) = \varrho, \tag{230}$$

where in the basis of the eigenvectors of ϱ we define

$$(Z_f)_{kl} = \begin{cases} \frac{(X_f)_{kl}}{(X_{f_{WY}})_{kl}}, & \text{if } \lambda_k \neq \lambda_l, \\ 0, & \text{if } \lambda_k = \lambda_l. \end{cases} \tag{231}$$

Proof. We use the definition of $D_{\text{DPT}}(\varrho, \sigma)^2$ given in Eq. (1) and the equation relating it to the Wigner-Yanase skew information given in Eq. (3), together with Eq. (219). ∎

Let us again calculate a concrete example. For instance, for obtaining the quantum Fisher information as an optimization over general quantum states, we should use

$$(Z_{f_{\max}})_{kl} = Q_{f_{\max}, f_{WY}}, \tag{232}$$

where $Q_{f_{\max}, f_{WY}}$ is defined in Eq. (222).

Based on these, we can define various Wasserstein distance measures, for which the self-distance equals various quantum Fisher information quantities.

Definition 8. A family of Wasserstein distance measures can be defined based on an optimization

over separable states as

$$D_{\mathrm{GMPC,sep}}^{f}(\varrho,\sigma)^2$$
$$= \frac{1}{2}\min_{\varrho_{12}}\mathrm{Tr}[(Y_f\circ H\otimes\mathbb{1}-\mathbb{1}\otimes Y_f\circ H)^2\varrho_{12}],$$
$$\text{s. t. } \varrho_{12}\in\mathcal{S},$$
$$\mathrm{Tr}_2(\varrho_{12})=\varrho,$$
$$\mathrm{Tr}_1(\varrho_{12})=\sigma. \qquad (233)$$

Based on Observation 18, for the self-distance

$$D_{\mathrm{GMPC,sep}}^{f}(\varrho,\varrho)^2=\frac{1}{4}\mathcal{F}_Q^f[\varrho,H] \qquad (234)$$

holds.

Definition 9. A family of Wasserstein distance measures can be defined based on an optimization over general quantum states as

$$D_{\mathrm{DPT}}^{f}(\varrho,\sigma)^2$$
$$= \frac{1}{2}\min_{\varrho_{12}}\mathrm{Tr}\{[(Z_f\circ H)^T\otimes\mathbb{1}-\mathbb{1}\otimes Z_f\circ H]^2\varrho_{12}\},$$
$$\text{s. t. } \varrho_{12}\in\mathcal{D},$$
$$\mathrm{Tr}_2(\varrho_{12})=\varrho^T,$$
$$\mathrm{Tr}_1(\varrho_{12})=\sigma. \qquad (235)$$

Based on Observation 19, for the self-distance

$$D_{\mathrm{DPT}}^{f}(\varrho,\varrho)^2=\frac{1}{4}\mathcal{F}_Q^f[\varrho,H] \qquad (236)$$

holds. It would be interesting to examine the properties of the quantities defined in Definitions 8 and 9.

11 Conclusions

We discussed how to define the quantum Wasserstein distance as an optimization over bipartite separable states rather than an optimization over general quantum states. With such a definition, the self-distance becomes related to the quantum Fisher information. We introduced also variance-like quantities in which we replaced the minimization used in the definition of the quantum Wasserstein distance by a maximization, and examined their properties. We discussed the relation of our findings to entanglement criteria. We examined also the quantity obtained after we considered optimizing the variance rather than the second moment in the usual expression of the quantum Wasserstein distance. Finally, we extended our results to the various generalized quantum Fisher information quantities. The details of the numerical calculations are discussed in Appendix A.

Acknowledgments

A Details of the numerical calculations

We used MATLAB [116] for numerical calculations. We used the semidefinite solver MOSEK [117] and the front-end YALMIP [118]. We also used the QUBIT4MATLAB package [119, 120]. $D_{\mathrm{DPT}}(\varrho,\sigma)^2$ given in Definition 1 and $D_{\mathrm{GMPC}}(\varrho,\sigma)^2$ given in Definition 2 can be obtained using semidefinite programming. $D_{\mathrm{GMPC,sep}}(\varrho,\sigma)^2$ in Definition 3 and $D_{\mathrm{DPT,sep}}(\varrho,\sigma)^2$ in Definition 4 need an optimization over separable states. The optimization over separable states can be carried out numerically using semidefinite programming for two qubits, since in this case the set of PPT states equals the set of separable states.

We included the routines computing the various quantum Wasserstein distance measures as `wdistsquare_GMPC_ppt.m`, `wdistsquare_GMPC.m`, `wdistsquare_DPT_ppt.m`, and `wdistsquare_DPT.m`. We included the various types of the variance-like quantities as `wvar_GMPC_ppt.m`, `wvar_GMPC.m`, `wvar_DPT_ppt.m`, and `wvar_DPT.m`. The usage of these routines is demonstrated in `example_wdistsquare.m`.

References

[1] G. Monge. "Mémoire sur la théory des déblais et des remblais". Mémoires de l'Académie Royale de Sciences de Paris (1781).

[2] L. Kantorovitch. "On the translocation of masses". Management Science **5**, 1–4 (1958). url: http://www.jstor.org/stable/2626967.

[3] Emmanuel Boissard, Thibaut Le Gouic, and Jean-Michel Loubes. "Distribution's template estimate with wasserstein metrics". Bernoulli **21**, 740–759 (2015).

[4] Oleg Butkovsky. "Subgeometric rates of convergence of Markov processes in the Wasserstein metric". Ann. Appl. Probab. **24**, 526–552 (2014).

[5] M. Hairer, J.-C. Mattingly and M. Scheutzow. "Asymptotic coupling and a general form of Harris' theorem with applications to stochastic delay equations". Probab. Theory Relat. Fields **149**, 223–259 (2011).

[6] M. Hairer and J.C. Mattingly. "Spectral Gaps in Wasserstein Distances and the 2D Stochastic Navier-Stokes Equations". Ann. Probab. **36**, 2050–2091 (2008).

[7] A. Figalli, F. Maggi and A. Pratelli. "A mass transportation approach to quantitative isoperimetric inequalities". Invent. Math. **182**, 167–211. (2010).

[8] A. Figalli and F. Maggi. "On the shape of liquid drops and crystals in the small mass regime". Arch. Ration. Mech. Anal. **201**, 143–207 (2011).

[9] J. Lott and C. Villani. "Ricci curvature for metric-measure spaces via optimal transport". Ann. of Math. **169 (3)**, 903–991 (2009).

[10] Max-K. von Renesse and Karl-Theodor Sturm. "Transport inequalities, gradient estimates, entropy, and Ricci curvature". Comm. Pure Appl. Math. **58**, 923–940 (2005).

[11] Karl-Theodor Sturm. "On the geometry of metric measure spaces I". Acta Math. **196**, 65–131 (2006).

[12] Karl-Theodor Sturm. "On the geometry of metric measure spaces II". Acta Math. **196**, 133–177 (2006).

[13] Benoît Kloeckner. "A geometric study of Wasserstein spaces: Euclidean spaces". Annali della Scuola Normale Superiore di Pisa - Classe di Scienze, Scuola Normale Superiore 2010 **IX (2)**, 297–323 (2010).

[14] György Pál Gehér, Tamás Titkos, and Dániel Virosztek. "On isometric embeddings of wasserstein spaces – the discrete case". J. Math. Anal. Appl. **480**, 123435 (2019).

[15] György Pál Gehér, T. Titkos, Dániel Virosztek. "Isometric study of Wasserstein spaces – the real line". Trans. Amer. Math. Soc. **373**, 5855–5883 (2020).

[16] György Pál Gehér, Tamás Titkos, and Dániel Virosztek. "The isometry group of Wasserstein spaces: the Hilbertian case". J. Lond. Math. Soc. **106**, 3865–3894 (2022).

[17] György Pál Gehér, Tamás Titkos, and Dániel Virosztek. "Isometric rigidity of wasserstein tori and spheres". Mathematika **69**, 20–32 (2023).

[18] Gergely Kiss and Tamás Titkos. "Isometric rigidity of wasserstein spaces: The graph metric case". Proc. Am. Math. Soc. **150**, 4083–4097 (2022).

[19] György Pál Gehér, Tamás Titkos, and Dániel Virosztek. "On the exotic isometry flow of the quadratic wasserstein space over the real line". Linear Algebra Appl. (2023).

[20] S. Kolouri, S. R. Park and G. K. Rohde. "The Radon cumulative distribution transform and its application to image classification". IEEE Trans. Image Process. **25**, 920–934 (2016).

[21] W. Wang, D. Slepčev, S. Basu, J. A. Ozolek and G. K. Rohde. "A linear optimal transportation framework for quantifying and visualizing variations in sets of images". Int. J. Comput. Vis. **101**, 254–269 (2013).

[22] S. Kolouri, S. Park, M. Thorpe, D. Slepčev, G. K. Rohde. "Optimal Mass Transport: Signal processing and machine-learning applications". IEEE Signal Processing Magazine **34**, 43–59 (2017).

[23] A. Gramfort, G. Peyré and M. Cuturi. "Fast Optimal Transport Averaging of Neuroimaging Data". Information Processing in Medical Imaging. IPMI 2015. Lecture Notes in Computer Science **9123**, 261–272 (2015).

[24] Z. Su, W. Zeng, Y. Wang, Z. L. Lu and X. Gu. "Shape classification using Wasserstein distance for brain morphometry analysis". Information Processing in Medical Imaging. IPMI 2015. Lecture Notes in Computer Science **24**, 411–423 (2015).

[25] Martin Arjovsky, Soumith Chintala, and Léon Bottou. "Wasserstein generative adversarial networks". In Doina Precup and Yee Whye Teh, editors, Proceedings of the 34th International Conference on Machine Learning. Volume 70 of Proceedings of Machine Learning Research, pages 214–223. PMLR (2017). arXiv:1701.07875.

[26] T. A. El Moselhy and Y. M. Marzouk. "Bayesian inference with optimal maps". J. Comput. Phys. **231**, 7815–7850 (2012).

[27] Gabriel Peyré and Marco Cuturi. "Computational Optimal Transport: With Applications to Data Science". Found. Trends Machine Learn. **11**, 355–602 (2019).

[28] Charlie Frogner, Chiyuan Zhang, Hossein Mobahi, Mauricio Araya, and Tomaso A Poggio. "Learning with a wasserstein loss". In C. Cortes, N. Lawrence, D. Lee, M. Sugiyama, and R. Garnett, editors, Advances in Neural Information Processing Systems. Volume 28. Curran Associates, Inc. (2015). arXiv:1506.05439.

[29] A. Ramdas, N. G. Trillos and M. Cuturi. "On Wasserstein Two-Sample Testing and Related Families of Nonparametric Tests". Entropy **19**, 47. (2017).

[30] S. Srivastava, C. Li and D. B. Dunson. "Scalable Bayes via Barycenter in Wasserstein Space". J. Mach. Learn. Res. **19**, 1–35 (2018). arXiv:1508.05880.

[31] Karol Życzkowski and Wojeciech Slomczynski. "The Monge distance between quantum states". J. Phys. A: Math. Gen. **31**, 9095–9104 (1998).

[32] Karol Życzkowski and Wojciech Slomczynski. "The Monge metric on the sphere and geometry of quantum states". J. Phys. A: Math. Gen. **34**, 6689–6722 (2001).

[33] Ingemar Bengtsson and Karol Życzkowski. "Geometry of quantum states: An introduction to quantum entanglement". Cambridge University Press. (2006).

[34] P. Biane and D. Voiculescu. "A free probability analogue of the Wasserstein metric on the trace-state space". GAFA, Geom. Funct. Anal. **11**, 1125–1138 (2001).

[35] Eric A. Carlen and Jan Maas. "An Analog of the 2-Wasserstein Metric in Non-Commutative Probability Under Which the Fermionic Fokker-Planck Equation is Gradient Flow for the Entropy". Commun. Math. Phys. **331**, 887–926 (2014).

[36] Eric A. Carlen and Jan Maas. "Gradient flow and entropy inequalities for quantum Markov semigroups with detailed balance". J. Funct. Anal. **273**, 1810–1869 (2017).

[37] Eric A. Carlen and Jan Maas. "Non-commutative calculus, optimal transport and functional inequalities in dissipative quantum systems". J. Stat. Phys. **178**, 319–378 (2020).

[38] Nilanjana Datta and Cambyse Rouzé. "Concentration of quantum states from quantum functional and transportation cost inequalities". J. Math. Phys. **60**, 012202 (2019).

[39] Nilanjana Datta and Cambyse Rouzé. "Relating relative entropy, optimal transport and Fisher information: A quantum HWI inequality". Ann. Henri Poincaré **21**, 2115–2150 (2020).

[40] François Golse, Clément Mouhot, and Thierry Paul. "On the mean field and classical limits of quantum mechanics". Commun. Math. Phys. **343**, 165–205 (2016).

[41] François Golse and Thierry Paul. "The Schrödinger equation in the mean-field and semiclassical regime". Arch. Ration. Mech. Anal. **223**, 57–94 (2017).

[42] François Golse and Thierry Paul. "Wave packets and the quadratic Monge-Kantorovich distance in quantum mechanics". Comptes Rendus Math. **356**, 177–197 (2018).

[43] François Golse. "The quantum N-body problem in the mean-field and semiclassical regime". Phil. Trans. R. Soc. A **376**, 20170229 (2018).

[44] E. Caglioti, F. Golse, and T. Paul. "Quantum optimal transport is cheaper". J. Stat. Phys. **181**, 149–162 (2020).

[45] Emanuele Caglioti, François Golse, and Thierry Paul. "Towards optimal transport for quantum densities". arXiv:2101.03256 (2021).

[46] Giacomo De Palma and Dario Trevisan. "Quantum optimal transport with quantum channels". Ann. Henri Poincaré **22**, 3199–3234 (2021).

[47] Giacomo De Palma, Milad Marvian, Dario Trevisan, and Seth Lloyd. "The quantum Wasserstein distance of order 1". IEEE Trans. Inf. Theory **67**, 6627–6643 (2021).

[48] Shmuel Friedland, Michał Eckstein, Sam Cole, and Karol Życzkowski. "Quantum Monge–Kantorovich problem and transport distance between density matrices". Phys. Rev. Lett. **129**, 110402 (2022).

[49] Sam Cole, Michał Eckstein, Shmuel Friedland, and Karol Życzkowski. "Quantum optimal transport". arXiv:2105.06922 (2021).

[50] R. Bistroń, M. Eckstein, and K. Życzkowski. "Monotonicity of a quantum 2-Wasserstein distance". J. Phys. A: Math. Theor. **56**, 095301 (2023).

[51] György Pál Gehér, József Pitrik, Tamás Titkos, and Dániel Virosztek. "Quantum Wasserstein isometries on the qubit state space". J. Math. Anal. Appl. **522**, 126955 (2023).

[52] Lu Li, Kaifeng Bu, Dax Enshan Koh, Arthur Jaffe, and Seth Lloyd. "Wasserstein complexity of quantum circuits". arXiv: 2208.06306 (2022).

[53] Bobak Toussi Kiani, Giacomo De Palma, Milad Marvian, Zi-Wen Liu, and Seth Lloyd. "Learning quantum data with the quantum earth mover's distance". Quantum Sci. Technol. **7**, 045002 (2022).

[54] E. P. Wigner and Mutsuo M. Yanase. "Information contents of distributions". Proc. Natl. Acad. Sci. U.S.A. **49**, 910–918 (1963).

[55] Ryszard Horodecki, Paweł Horodecki, Michał Horodecki, and Karol Horodecki. "Quantum entanglement". Rev. Mod. Phys. **81**, 865–942 (2009).

[56] Otfried Gühne and Géza Tóth. "Entanglement detection". Phys. Rep. **474**, 1–75 (2009).

[57] Nicolai Friis, Giuseppe Vitagliano, Mehul Malik, and Marcus Huber. "Entanglement certification from theory to experiment". Nat. Rev. Phys. **1**, 72–87 (2019).

[58] Vittorio Giovannetti, Seth Lloyd, and Lorenzo Maccone. "Quantum-enhanced measurements: Beating the standard quantum limit". Science **306**, 1330–1336 (2004).

[59] Matteo G. A. Paris. "Quantum estimation for quantum technology". Int. J. Quant. Inf. **07**, 125–137 (2009).

[60] Rafal Demkowicz-Dobrzanski, Marcin Jarzyna, and Jan Kolodynski. "Chapter four - Quantum limits in optical interferometry". Prog. Optics **60**, 345 – 435 (2015). arXiv:1405.7703.

[61] Luca Pezze and Augusto Smerzi. "Quantum theory of phase estimation". In G.M. Tino and M.A. Kasevich, editors, Atom Interferometry (Proc. Int. School of Physics 'Enrico Fermi', Course 188, Varenna). Pages 691–741. IOS Press, Amsterdam (2014). arXiv:1411.5164.

[62] Géza Tóth and Dénes Petz. "Extremal properties of the variance and the quantum Fisher information". Phys. Rev. A **87**, 032324 (2013).

[63] Sixia Yu. "Quantum Fisher Information as the Convex Roof of Variance". arXiv:1302.5311 (2013).

[64] Géza Tóth and Florian Fröwis. "Uncertainty relations with the variance and the quantum Fisher information based on convex decompositions of density matrices". Phys. Rev. Research **4**, 013075 (2022).

[65] Shao-Hen Chiew and Manuel Gessner. "Improving sum uncertainty relations with the quantum Fisher information". Phys. Rev. Research **4**, 013076 (2022).

[66] C. W. Helstrom. "Quantum detection and estimation theory". Academic Press, New York. (1976). url: www.elsevier.com/books/quantum-detection-and-estimation-theory/helstrom/978-0-12-340050-5.

[67] A. S. Holevo. "Probabilistic and statistical aspects of quantum theory". North-Holland, Amsterdam. (1982).

[68] Samuel L. Braunstein and Carlton M. Caves. "Statistical distance and the geometry of quantum states". Phys. Rev. Lett. **72**, 3439–3443 (1994).

[69] Samuel L Braunstein, Carlton M Caves, and Gerard J Milburn. "Generalized uncertainty relations: Theory, examples, and Lorentz invariance". Ann. Phys. **247**, 135–173 (1996).

[70] Dénes Petz. "Quantum information theory and quantum statistics". Springer, Berlin, Heilderberg. (2008).

[71] Géza Tóth and Iagoba Apellaniz. "Quantum metrology from a quantum information science perspective". J. Phys. A: Math. Theor. **47**, 424006 (2014).

[72] Luca Pezzè, Augusto Smerzi, Markus K. Oberthaler, Roman Schmied, and Philipp Treutlein. "Quantum metrology with nonclassical states of atomic ensembles". Rev. Mod. Phys. **90**, 035005 (2018).

[73] Marco Barbieri. "Optical quantum metrology". PRX Quantum **3**, 010202 (2022).

[74] Zoltán Léka and Dénes Petz. "Some decompositions of matrix variances". Probab. Math. Statist. **33**, 191–199 (2013). arXiv:1408.2707.

[75] Dénes Petz and Dániel Virosztek. "A characterization theorem for matrix variances". Acta Sci. Math. (Szeged) **80**, 681–687 (2014).

[76] Akio Fujiwara and Hiroshi Imai. "A fibre bundle over manifolds of quantum channels and its application to quantum statistics". J. Phys. A: Math. Theor. **41**, 255304 (2008).

[77] B. M. Escher, R. L. de Matos Filho, and L. Davidovich. "General framework for estimating the ultimate precision limit in noisy quantum-enhanced metrology". Nat. Phys. **7**, 406–411 (2011).

[78] Rafał Demkowicz-Dobrzański, Jan Kołodyński, and Mădălin Guţă. "The elusive Heisenberg limit in quantum-enhanced metrology". Nat. Commun. **3**, 1063 (2012).

[79] Iman Marvian. "Operational interpretation of quantum fisher information in quantum thermodynamics". Phys. Rev. Lett. **129**, 190502 (2022).

[80] Reinhard F. Werner. "Quantum states with Einstein-Podolsky-Rosen correlations admitting a hidden-variable model". Phys. Rev. A **40**, 4277–4281 (1989).

[81] K. Eckert, J. Schliemann, D. Bruss, and M. Lewenstein. "Quantum correlations in systems of indistinguishable particles". Ann. Phys. **299**, 88–127 (2002).

[82] Tsubasa Ichikawa, Toshihiko Sasaki, Izumi Tsutsui, and Nobuhiro Yonezawa. "Exchange symmetry and multipartite entanglement". Phys. Rev. A **78**, 052105 (2008).

[83] Pawel Horodecki. "Separability criterion and inseparable mixed states with positive partial transposition". Phys. Lett. A **232**, 333–339 (1997).

[84] Asher Peres. "Separability criterion for density matrices". Phys. Rev. Lett. **77**, 1413–1415 (1996).

[85] Paweł Horodecki, Michał Horodecki, and Ryszard Horodecki. "Bound entanglement can be activated". Phys. Rev. Lett. **82**, 1056–1059 (1999).

[86] Géza Tóth and Tamás Vértesi. "Quantum states with a positive partial transpose are useful for metrology". Phys. Rev. Lett. **120**, 020506 (2018).

[87] Scott Hill and William K. Wootters. "Entanglement of a pair of quantum bits". Phys. Rev. Lett. **78**, 5022–5025 (1997).

[88] William K. Wootters. "Entanglement of formation of an arbitrary state of two qubits". Phys. Rev. Lett. **80**, 2245–2248 (1998).

[89] David P. DiVincenzo, Christopher A. Fuchs, Hideo Mabuchi, John A. Smolin, Ashish Thapliyal, and Armin Uhlmann. "Entanglement of assistance". quant-ph/9803033 (1998).

[90] John A. Smolin, Frank Verstraete, and Andreas Winter. "Entanglement of assistance and multipartite state distillation". Phys. Rev. A **72**, 052317 (2005).

[91] Holger F. Hofmann and Shigeki Takeuchi. "Violation of local uncertainty relations as a signature of entanglement". Phys. Rev. A **68**, 032103 (2003).

[92] Otfried Gühne. "Characterizing entanglement via uncertainty relations". Phys. Rev. Lett. **92**, 117903 (2004).

[93] Otfried Gühne, Mátyás Mechler, Géza Tóth, and Peter Adam. "Entanglement criteria based on local uncertainty relations are strictly stronger than the computable cross norm criterion". Phys. Rev. A **74**, 010301 (2006).

[94] Giuseppe Vitagliano, Philipp Hyllus, Iñigo L. Egusquiza, and Géza Tóth. "Spin squeezing inequalities for arbitrary spin". Phys. Rev. Lett. **107**, 240502 (2011).

[95] A. R. Edmonds. "Angular momentum in quantum mechanics". Princeton University Press. (1957).

[96] Géza Tóth. "Entanglement detection in optical lattices of bosonic atoms with collective measurements". Phys. Rev. A **69**, 052327 (2004).

[97] Géza Tóth, Christian Knapp, Otfried Gühne, and Hans J. Briegel. "Optimal spin squeezing inequalities detect bound entanglement in spin models". Phys. Rev. Lett. **99**, 250405 (2007).

[98] Géza Tóth and Morgan W Mitchell. "Generation of macroscopic singlet states in atomic ensembles". New J. Phys. **12**, 053007 (2010).

[99] Géza Tóth. "Detection of multipartite entanglement in the vicinity of symmetric Dicke states". J. Opt. Soc. Am. B **24**, 275–282 (2007).

[100] Géza Tóth, Tobias Moroder, and Otfried Gühne. "Evaluating convex roof entanglement measures". Phys. Rev. Lett. **114**, 160501 (2015).

[101] Lieven Vandenberghe and Stephen Boyd. "Semidefinite programming". SIAM Review **38**, 49–95 (1996).

[102] Géza Tóth. "Multipartite entanglement and high-precision metrology". Phys. Rev. A **85**, 022322 (2012).

[103] Philipp Hyllus, Wiesław Laskowski, Roland Krischek, Christian Schwemmer, Witlef Wieczorek, Harald Weinfurter, Luca Pezzé, and Augusto Smerzi. "Fisher information and multiparticle entanglement". Phys. Rev. A **85**, 022321 (2012).

[104] Géza Tóth, Tamás Vértesi, Paweł Horodecki, and Ryszard Horodecki. "Activating hidden metrological usefulness". Phys. Rev. Lett. **125**, 020402 (2020).

[105] A. C. Doherty, Pablo A. Parrilo, and Federico M. Spedalieri. "Distinguishing separable and entangled states". Phys. Rev. Lett. **88**, 187904 (2002).

[106] Andrew C. Doherty, Pablo A. Parrilo, and Federico M. Spedalieri. "Complete family of separability criteria". Phys. Rev. A **69**, 022308 (2004).

[107] Andrew C. Doherty, Pablo A. Parrilo, and Federico M. Spedalieri. "Detecting multipartite entanglement". Phys. Rev. A **71**, 032333 (2005).

[108] Harold Ollivier and Wojciech H. Zurek. "Quantum discord: A measure of the quantumness of correlations". Phys. Rev. Lett. **88**, 017901 (2001).

[109] L. Henderson and V. Vedral. "Classical, quantum and total correlations". J. Phys. A: Math. Gen. **34**, 6899 (2001).

[110] Anindita Bera, Tamoghna Das, Debasis Sadhukhan, Sudipto Singha Roy, Aditi Sen(De), and Ujjwal Sen. "Quantum discord and its allies: a review of recent progress". Rep. Prog. Phys. **81**, 024001 (2017).

[111] Dénes Petz. "Covariance and Fisher information in quantum mechanics". J. Phys. A: Math. Gen. **35**, 929 (2002).

[112] Paolo Gibilisco, Fumio Hiai, and Dénes Petz. "Quantum covariance, quantum Fisher information, and the uncertainty relations". IEEE Trans. Inf. Theory **55**, 439–443 (2009).

[113] D. Petz and C. Ghinea. "Introduction to quantum Fisher information". Volume 27, pages 261–281. World Scientific. (2011).

[114] Frank Hansen. "Metric adjusted skew information". Proc. Natl. Acad. Sci. U.S.A. **105**, 9909–9916 (2008).

[115] Paolo Gibilisco, Davide Girolami, and Frank Hansen. "A unified approach to local quantum uncertainty and interferometric power by metric adjusted skew information". Entropy **23**, 263 (2021).

[116] MATLAB. "9.9.0.1524771(r2020b)". The MathWorks Inc. Natick, Massachusetts (2020).

[117] MOSEK ApS. "The MOSEK optimization toolbox for MATLAB manual. Version 9.0". (2019). url: docs.mosek.com/9.0/toolbox/index.html.

[118] J. Löfberg. "YALMIP : A Toolbox for Modeling and Optimization in MATLAB". In Proceedings of the CACSD Conference. Taipei, Taiwan (2004).

[119] Géza Tóth. "QUBIT4MATLAB V3.0: A program package for quantum information science and quantum optics for MATLAB". Comput. Phys. Commun. **179**, 430–437 (2008).

[120] The package QUBIT4MATLAB is available at https://www.mathworks.com/matlabcentral/fileexchange/8433, and at the personal home page https://gtoth.eu/qubit4matlab.html.