



Quantum Wasserstein distances for quantum permutation groups

Anshu ^{a,b,c}, David Jekel ^{d,*}, Therese Basa Landry ^e

^a Fields Institute for Research in Mathematical Sciences, 222 College Street, Toronto, M5T 3J1, Ontario, Canada

^b Department of Mathematics and Statistics, York University, Ross Hall, 4700 Keele St, Toronto, M3J 1P3, Ontario, Canada

^c Department of Mathematics, University of Ulsan, 93 Daehak-ro, Nam-gu, Ulsan Metropolitan City, 44610, Korea

^d Department of Mathematical Sciences, University of Copenhagen, Universitetsparken 5, Copenhagen Ø, 2100, Denmark

^e Department of Mathematics, University of California Santa Barbara, South Hall, Santa Barbara, 93106, CA, United States



ARTICLE INFO

Article history:

Received 17 June 2025

Received in revised form 27 August 2025

Accepted 31 August 2025

Available online 8 September 2025

MSC:

46L67

49Q22

46L89

81R60

Keywords:

Quantum permutation group

Quantum metric space

Quantum optimal transport

Hamming distance

ABSTRACT

We seek an analog for the quantum permutation group S_n^+ of the normalized Hamming distance for permutations. We define three distances on the tracial state space of $C(S_n^+)$ that generalize the L^1 -Wasserstein distance of probability measures on S_n equipped with the normalized Hamming metric, for which we demonstrate basic metric properties, subadditivity under convolution, and density of the Lipschitz elements in the C^* -algebra.

© 2025 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

1. Introduction

1.1. Motivation

Our goal in this paper is to “quantize” the Hamming distance on the permutation group S_n to obtain a “quantum metric” on $C(S_n^+)$. This is motivated on the one hand by non-commutative geometry and Marc A. Rieffel’s notion of quantum metric spaces [1,2], which was adapted to the setting of tracial states by Jacelon [3,4], and on the other hand by recent developments in non-commutative optimal transport theory for quantum states, e.g. [5–7].

The program of non-commutative geometry initiated by Alain Connes adapts classical tools from topology and Riemannian geometry to the operator algebraic setting. For example, Connes showed in [8,9] that the geodesic distance d_g on a compact Riemannian spin manifold M can be recovered from the C^* -algebra $C(M)$, the Hilbert space H of L^2 -spinor fields, and the Dirac operator.¹ He further noted that many of the arguments did not require commutativity of

* Corresponding author.

E-mail addresses: aanshu@yorku.ca (Anshu), daj@math.ku.dk (D. Jekel), tlandry@ucsb.edu (T.B. Landry).

¹ Similarly, for a general oriented compact Riemannian manifold, the geodesic distance can be recovered from the Hodge Laplacian.

the underlying C^* -algebra and hence should generalize to non-commutative operator algebras. Rieffel developed a corresponding notion of metric spaces in the non-commutative setting [1,2,10]. The idea is to capture the underlying metric in a dual way through the associated L^1 -Wasserstein distance on states. Given a Dirac operator D one can define a Lipschitz norm on the C^* -algebra A by $\|a\|_{\text{Lip}} = \|[a, D]\|$, and then the “ L^1 -Wasserstein distance” on states is given by $W_1(\varphi, \psi) = \sup\{|\varphi(a) - \psi(a)| : \|a\|_{\text{Lip}} \leq 1\}$. Rieffel investigated when this metric recovers the weak- $*$ topology on the state space [1]. In some cases, this metric also gives rise to a metric on the density space of a unital C^* -algebra whose induced topology coincides with that of the Bures metric from quantum information theory [11].

For a discrete abelian group G , the dual group \widehat{G} is compact and the (full or reduced) group C^* -algebra $C^*(G)$ is isomorphic to the continuous functions $C(\widehat{G})$. Thus, a natural class of examples to apply quantum metric theory in the sense of Rieffel arises from group C^* -algebras. In particular, Christ and Rieffel studied quantum metric space structures on $C_r^*(G)$ where the metric is defined via a length function [12]. Length functions and the associated quantum metric spaces have also been investigated in the more general discrete *quantum groups*, as well as the quantum permutation group S_n^+ [13,14], and quantum metric structures have also been developed for homogeneous spaces of quantum $SU(2)$ [15–17].

Compact quantum groups, as defined and developed by Woronowicz [18–20], can be viewed as quantum analogues of the continuous functions on a group $C(G)$, where G is a locally compact group. To define multiplication in this setting, one must again dualize. For a group G , there is a natural coassociative operation $\Delta : C(G) \rightarrow C(G \times G)$ given by $\Delta f(s, t) = f(st)$, called the *comultiplication*, which encodes the group multiplication structure. Thus, a quantum group structure on a C^* -algebra A is defined via a coassociative operation $\Delta : A \rightarrow A \otimes A$ and a counit $\varepsilon : A \otimes A \rightarrow A$ satisfying certain axioms.

Our work focuses on the quantum permutation groups S_n^+ . In answer to a question posed by Connes, Wang developed the notion of *quantum symmetries* for an n -point set by defining the quantum group S_n^+ [21]. The underlying C^* -algebra $C(S_n^+)$ is the universal (unital) C^* -algebra generated by elements $(u_{ij})_{i,j=1}^n$ such that u_{ij} is a projection and $\sum_i u_{ij} = 1$ and $\sum_j u_{ij} = 1$. The comultiplication on the generators is given by $\Delta(u_{ij}) = \sum_k u_{ik} \otimes u_{kj}$ which generalizes the matrix multiplication structure in a natural way; u_{ij} corresponds in the classical case to the continuous function on S_n that evaluates the ij entry of the permutation matrix, or $\mathbb{1}_{\sigma(i)=j}$.

Since $C(S_n^+)$ is not a discrete quantum group for $n \geq 4$, it is not clear to us how to adapt the notion of length function per se [14]. Rather, we take inspiration from cost operators used in quantum optimal transport theory. One of the goals of this subject is to find appropriate analogs of Wasserstein distances on the space of quantum states of some system (often the space of states or density operators on a finite-dimensional C^* -algebra). Quantum Wasserstein distances can be used as a measure of the effort it takes to transform one state to another by elementary operations (see e.g. [7]), as well as to define a Riemannian structure analogous to the manifold of probability measures [22].

As in classical optimal transport, one needs both a *cost* to measure the distances between points as well as a notion of a *transport plan* or a *coupling* for joining two states. Let us start with the cost operator. If the underlying operator algebra A is analogous to $C(X)$, then a cost or distance function $X \times X \rightarrow [0, \infty)$ would correspond to a positive element C of $A \otimes A$. The symmetry of distance means that C is symmetric under the tensor flip, and the triangle inequality means that $\iota_{1,3}(C) \leq \iota_{1,2}(C) + \iota_{2,3}(C)$ where $\iota_{i,j}$ is the inclusion of $A \otimes A$ into the i th and j th copies in $A \otimes A \otimes A$.

To construct such a cost operator on the quantum permutation group, recall that the normalized Hamming distance on the permutation group S_n is given by

$$d_H(\sigma, \sigma') = \frac{1}{n} |\{i : \sigma(i) \neq \sigma'(i)\}|.$$

This can be expressed in terms of the entries of the permutation matrices as follows:

$$d_H(\sigma, \sigma') = 1 - \frac{1}{n} |\{i : \sigma(i) = \sigma'(i)\}| = 1 - \frac{1}{n} |\{i : \sigma^{-1}\sigma'(i) = i\}| = 1 - \text{tr}_n(\sigma^* \sigma') = 1 - \frac{1}{n} \sum_{i,j=1}^n \sigma_{ij} \sigma'_{ij}.$$

This expression represents an element of $C(S_n \times S_n)$ since $\sigma \mapsto \sigma_{ij}$ is a continuous function on S_n . To obtain the quantum version, we just have to replace this continuous function on S_n with the element $u_{ij} \in C(S_n^+)$. This leads to the following cost operator in $C(S_n^+) \otimes C(S_n^+)$:

$$C_H = 1 - \frac{1}{n} \sum_{i,j=1}^n u_{ij} \otimes u_{ij}.$$

We show that C_H satisfies the triangle inequality (see §4.2). We also show a subadditivity property under comultiplication (see §4.4). This is a natural property if we want to obtain metrics on the state space that interact with convolution operation on states of $C(S_n^+)$; for instance, in order to study convolution semigroups on S_n^+ [23].

The notion of coupling in the non-commutative setting is much more subtle. A natural approach is to define a coupling of two states φ_1 and φ_2 on a C^* -algebra A as a state φ on $A \otimes A$ with $\varphi(a \otimes 1) = \varphi_1(a)$ and $\varphi(1 \otimes a) = \varphi_2(a)$; see e.g. [5,6]. However, there are major difficulties in proving the triangle inequality because of the inability to amalgamate couplings, and sometimes the self-distance is nonzero. One can also consider unital completely positive maps $A \rightarrow A$ transforming the one state into the other. One thing that helps in the von Neumann algebraic setting is that by imposing appropriate

modular conditions on the states, one can guarantee that couplings can be amalgamated via the relative tensor product [24]. However, it is unclear how to adapt this method to the setting of C^* -algebras, especially with states that are not faithful.

There are other approaches that are not based on a cost operator per se, but rather on other structures of the underlying space. For instance, [25] defined a quantum Hamming distance on the state space for n qdits, which gives a quantum version of quantifying changes one bit at a time. Very recently Rieffel [26] explained how similar quantum Hamming distances give rise to quantum compact metric spaces, and also shows how to encode the metric data in Dirac operator type structures. In fact, another Hamming-like distance has been used to study the ideal structure of AF algebras [27]. However, the Hamming distance in this paper is distinct and it is unclear how they relate, since the quantum permutation group does not relate naturally to an n -fold tensor product structure.

Many of these issues become easier to deal with in the setting of *traces* as opposed to general states. A trace on a C^* -algebra automatically produces a tracial von Neumann algebra on which the trace is faithful (see Lemma 2.2), while in general it is difficult to tell whether a state on a C^* -algebra gives rise to a faithful state on a von Neumann algebra and if so, how its modular group behaves. The work of Biane and Voiculescu [28] gives natural L^p -Wasserstein distances on tracial states on a certain universal C^* -algebra with specified generators; here the couplings are given by embeddings into an arbitrary larger tracial von Neumann algebra, in which one can measure the distance between the two copies of the generators. Moreover, the trace space of a C^* -algebra, unlike the state space, is always a Choquet simplex, meaning that every trace has a unique decomposition in terms of extreme points (see [29]).

Although restricting to tracial states may seem untenable for applications in quantum theory, many C^* -algebras admit an abundance of traces, which have played a key role in the stably finite case of the classification of unital separable simple nuclear \mathcal{Z} -stable C^* -algebras satisfying the UCT; see [30] for a version of the classification theorem and its history. Motivated by the classification program, Jacelon [3,31,4] studied certain metrics on the trace space of classifiable C^* -algebras A , showing that they are in duality with Lipschitz seminorms that vanish on the trace kernel of A , and hence are conceptually not far from quantum metrics defined by Rieffel. In fact, he showed that any compact metric space can be realized as the extreme boundary of the trace space of some classifiable C^* -algebra [31, Theorem 4.4], [4, Theorem A].

While classifiable C^* -algebras with large trace spaces require some effort to construct, it is much easier to see that many universal C^* -algebras given by generators and relations have extremely large trace spaces. For instance, the trace space on the universal C^* -algebra generated by self-adjoints $\|x_j\| \leq R$ is not separable with respect to the Biane-Voiculescu-Wasserstein distance [32, §5.5]. The C^* -algebra of the quantum permutation group studied in this paper is also a universal C^* -algebra that similarly has an abundance of traces, including many coming from finite-dimensional representations (see e.g. [33]).

1.2. Results

In this work, we study three analogs of the Hamming distance on trace space of $C(S_n^+)$, which we call $W_{H,\wedge}$, $W_{H,1}$, and $W_{H,\otimes}$. The first two distances are based on the notion of tracial couplings as in Biane and Voiculescu’s work [28], while the third is based on tracial tensor product couplings.

Definition 1.1 (*Tracial couplings*). Let A be a C^* -algebra and let $\varphi_1, \varphi_2 \in \mathcal{T}(A)$. A *tracial coupling* of φ_1 and φ_2 is a tuple $(M, \tau, \alpha_1, \alpha_2)$ where (M, τ) is a tracial von Neumann algebra and $\alpha_j : A \rightarrow M$ is a $*$ -homomorphism such that $\varphi_j = \tau \circ \alpha_j$ for $j = 1, 2$.

For a tracial coupling $(M, \tau, \alpha_1, \alpha_2)$, the two projections $\alpha_1(u_{ij})$ and $\alpha_2(u_{ij})$ need not commute with each other, and $1 - (1/n) \sum_{i,j} \tau(\alpha_1(u_{ij})\alpha_2(u_{ij}))$ need not define a metric. However, there are at least two ways to modify this expression to obtain a metric. The first is to replace the product of the two projections by the intersection $\alpha_1(u_{ij}) \wedge \alpha_2(u_{ij})$.

Definition 1.2 (*Free Hamming distance via intersections*). For $\varphi_1, \varphi_2 \in \mathcal{T}(C(S_n^+))$, define

$$W_{H,\wedge}(\varphi_1, \varphi_2) = \inf_{(M, \tau, \alpha_1, \alpha_2)} \left[1 - \frac{1}{n} \sum_{i,j=1}^n \tau(\alpha_1(u_{ij}) \wedge \alpha_2(u_{ij})) \right],$$

where the infimum is over all tracial couplings $(M, \tau, \alpha_1, \alpha_2)$.

As noted in §3.1, some tracial coupling always exists (e.g. by taking a tensor product), and although the class of tracial couplings is not a set, we can rephrase the infimum in terms of sets in the situations that we care about, thanks to Lemma 3.2. Hence, the infimum is well-defined and finite.

Another natural choice of distance looks at the non-commutative L^1 -distances $\|\alpha_1(u_{ij}) - \alpha_2(u_{ij})\|_{L^1(M, \tau)}$. To motivate this, recall that for two permutation matrices σ and σ' , we have

$$d_H(\sigma, \sigma') = \frac{1}{n} \sum_{i,j=1}^n \frac{1}{2} |\sigma_{ij} - \sigma'_{ij}|.$$

Indeed, if $\sigma(j) = \sigma'(j)$, then the j th columns of the two matrices agree, but if $\sigma(j) \neq \sigma'(j)$, then the j th columns of the matrices differ in exactly two rows, so $\sum_i |\sigma_{ij} - \sigma'_{ij}| = 2$. Since u_{ij} is analogous to the value of the ij entry of the matrix, we make the following definition.

Definition 1.3 (Free Hamming distance via L^1 -norms). For $\varphi_1, \varphi_2 \in \mathcal{T}(C(S_n^+))$, define

$$W_{H,1}(\varphi_1, \varphi_2) = \inf_{(M, \tau, \alpha_1, \alpha_2)} \frac{1}{n} \sum_{i,j=1}^n \frac{1}{2} \|\alpha_1(u_{ij}) - \alpha_2(u_{ij})\|_{L^1(M, \tau)},$$

where the infimum is over all tracial couplings $(M, \tau, \alpha_1, \alpha_2)$.

We summarize our results about $W_{H,\wedge}$ and $W_{H,1}$ in the following theorem.

Theorem 1.4.

- (1) $W_{H,\wedge}$ defines a metric on $\mathcal{T}(C(S_n^+))$. (See Proposition 3.6.)
- (2) For tracial states $\varphi_1, \varphi_2, \psi_1, \psi_2$, we have

$$W_{H,\wedge}(\varphi_1 * \psi_1, \varphi_2 * \psi_2) \leq W_{H,\wedge}(\varphi_1, \varphi_2) + W_{H,\wedge}(\psi_1, \psi_2).$$

(See Proposition 3.8.)

- (3) The infimum in the definition of $W_{H,\wedge}$ is a minimum (see Lemma 3.5).
- (4) $W_{H,\wedge}$ is lower semi-continuous with respect to the weak-* topology, and convergence in $W_{H,\wedge}$ implies weak-* convergence. (See Lemma 3.7.)

The same properties also hold for $W_{H,1}$. (See Lemma 3.9 and Proposition 3.10.)

We remark that the proof of the triangle inequality uses the amalgamated free product construction, just as for Biane and Voiculescu’s Wasserstein distance in free probability. The traciality of φ and ψ is essential for the triangle inequality because we need to use the fact that for projections p and q in a tracial von Neumann algebra, $\tau(p \wedge q) + \tau(p \vee q) = \tau(p) + \tau(q)$, which fails for general states.

The third notion of couplings that we consider is the naïve one of states on the tensor product with the given marginals, except that these states are restricted to be tracial. The corresponding distance $W_{H,\otimes}$ satisfies the triangle inequality but the self-distance is not necessarily zero. Our results on $W_{H,\otimes}$ are summarized as follows.

Theorem 1.5.

- (1) $W_{H,\otimes}$ is nonnegative and symmetric and satisfies the triangle inequality. (See Lemma 4.13.)
- (2) We have

$$W_{H,\otimes}(\varphi_1, \varphi_2) \geq \frac{1}{2} W_{H,\otimes}(\varphi_1, \varphi_1) + \frac{1}{2} W_{H,\otimes}(\varphi_2, \varphi_2).$$

(See Lemma 4.15.)

- (3) For tracial states $\varphi_1, \varphi_2, \psi_1, \psi_2$, we have

$$W_{H,\otimes}(\varphi_1 * \psi_1, \varphi_2 * \psi_2) \leq W_{H,\otimes}(\varphi_1, \varphi_2) + W_{H,\otimes}(\psi_1, \psi_2).$$

(See Proposition 4.19.)

A key point in proving the triangle inequality for the tensor couplings is the ability to amalgamate tensor couplings of (φ_1, φ_2) and (φ_2, φ_3) . This relies on the fact that the space of tracial states is Choquet simplex, which fails for general states on a non-commutative C^* -algebra. This means that $W_{H,\otimes}$ can be expressed in terms of classical couplings of probability measures on the extreme boundary $\partial_e \mathcal{T}(C(S_n^+))$, and the cost only depends on the values of $\varphi(u_{ij})$ for φ in this extreme boundary.

We relate the three distances to each other and to the classical Hamming distance as follows.

Proposition 1.6.

- (1) For traces $\varphi_1, \varphi_2 \in \mathcal{T}(C(S_n^+))$, we have

$$W_{H,1}(\varphi_1, \varphi_2) \leq W_{H,\wedge}(\varphi_1, \varphi_2) \leq W_{H,\otimes}(\varphi_1, \varphi_2).$$

(See Proposition 5.1.)

- (2) We furthermore have $W_{H,\wedge}(\varphi_1, \varphi_2) \leq \frac{1}{2} \|\varphi_1 - \varphi_2\|$. (See Proposition 5.2.)
- (3) If φ_1 and φ_2 are both induced by classical probability measures μ_1 and μ_2 on S_n , then $W_{H,1}$, $W_{H,\wedge}$, and $W_{H,\otimes}$ all agree with the classical L^1 Wasserstein distance of μ_1 and μ_2 with respect to d_H . (See Proposition 5.3.)
- (4) For $\varphi_1, \varphi_2 \in \mathcal{T}(C(S_n^+))$, each of the three distances is less than or equal to 1 with equality if and only if $\sum_{i,j=1}^n \varphi_1(u_{ij})\varphi_2(u_{ij}) = 0$. (See Proposition 5.4.)

Returning to the theme of quantum compact metric spaces, any distance d on the trace space of a C^* -algebra A induces a Lipschitz seminorm on A given by

$$\|a\|_{\text{Lip}(d)} = \sup_{\varphi_1 \neq \varphi_2} \frac{|\varphi_1(a) - \varphi_2(a)|}{d(\varphi_1, \varphi_2)}.$$

The domain of Lipschitz seminorm is defined as the set of a where $\|a\|_{\text{Lip}(d)} < \infty$.

Proposition 1.7. *If a is in the $*$ -algebra generated by $\{u_{ij} : i, j \in [n]\}$ in $C(S_n^+)$, then $\|a\|_{\text{Lip}(W_{H,\wedge})} \leq \|a\|_{\text{Lip}(W_{H,1})} < \infty$. So in particular, Lipschitz elements are dense for each of the two metrics. (See Proposition 5.5.)*

Our work leaves open many basic questions, which we discuss in §5.3, including how the topologies induced by these metrics relate to each other, whether there is a reasonable way to extend the definitions to non-tracial states, and more. In particular, it highlights the need for a better understanding of the trace space and the representation theory of $C(S_n^+)$. See 5.3 for further discussion.

1.3. Organization

The rest of the paper is organized as follows.

- §2 recalls background on C^* -algebras, von Neumann algebras, and quantum groups.
- §3 constructs $W_{H,\wedge}$ and proves Theorem 1.4.
- §4 constructs $W_{H,\otimes}$ and proves Theorem 1.5.
- §5.1 relates the distances to each other and proves Proposition 1.6.
- §5.2 proves Proposition 1.7 on Lipschitz seminorms.
- §5.3 gives questions for further research.

2. Preliminaries

We assume basic familiarity with C^* -algebras and von Neumann algebras; see e.g. [34–36]. Nonetheless, given that our paper spans several subdisciplines, we recall some basic notation and facts here for later use.

2.1. C^* -algebras

2.1.1. States and GNS construction

All C^* -algebras in this paper are assumed to be unital. The state space of a C^* -algebra will be denoted $\mathcal{S}(A)$ and the space of tracial states will be denoted $\mathcal{T}(A)$. These are both weak- $*$ convex and compact spaces.

For a state φ , we denote by (H_φ, π_φ) the GNS construction (see e.g. [34, §3.1]). Recall the sesquilinear form $\langle a, b \rangle_\varphi = \varphi(a^*b)$ is nonnegative, and hence defines an inner product on A/N_φ , where $N_\varphi = \{a : \varphi(a^*a) = 0\}$. The completion of A/N_φ is denoted by H_φ . For $a \in A$, we denote by \widehat{a} the corresponding element of H_φ , i.e. the equivalence class $a + N_\varphi$ in A/N_φ . Moreover, $\pi_\varphi : A \rightarrow B(H_\varphi)$ denotes the representation given by $\pi_\varphi(x)\widehat{a} = \widehat{xa}$.

In the course of the argument, we use the following facts relating states, $*$ -homomorphisms, and GNS representations.

Lemma 2.1. *Let $\theta : A \rightarrow B$ be a $*$ -homomorphism, and let φ be a state on B .*

- (1) Then $\varphi \circ \theta$ is a state on A .
- (2) If φ is tracial, then $\varphi \circ \theta$ is tracial.
- (3) There is an isometric map $U : H_{\varphi \circ \theta} \rightarrow H_\varphi$ given by $U(\widehat{a}) = \widehat{\theta(a)}$, and for $x \in A$, we have $U\pi_{\varphi \circ \theta}(x) = \pi_\varphi \circ \theta(x)U$.
- (4) If θ is surjective, then U is surjective and there is an isomorphism $\pi_{\varphi \circ \theta}(A) \cong \pi_\varphi(B)$ given by $T \mapsto UTU^*$.
- (5) If φ is faithful, then there is an isomorphism $\pi_{\varphi \circ \theta}(A) \rightarrow \theta(A)$ given by $\pi_{\varphi \circ \theta}(x) \mapsto \theta(x)$.

Proof. (1) and (2) are immediate.

(3) Note that

$$\langle a, a' \rangle_{\varphi \circ \theta} = \varphi \circ \theta(a^*a') = \varphi(\theta(a)^*\theta(a')) = \langle \theta(a), \theta(a') \rangle_\varphi.$$

Hence, the mapping θ is isometric with respect to the Hilbert space norms, and so gives a well-defined and isometric operator $U : H_{\varphi \circ \theta} \rightarrow H_\varphi$. Next, note that for $x, a \in A$, we have

$$U\pi_{\varphi \circ \theta}(x)\widehat{a} = U\widehat{xa} = \widehat{\theta(xa)} = \widehat{\theta(x)\theta(a)} = \pi_\varphi(\theta(x))\widehat{\theta(a)} = \pi_\varphi(\theta(x))U\widehat{a}.$$

Hence, $U\pi_{\varphi \circ \theta}(x) = \pi_\varphi(\theta(x))U$ as desired.

(4) If θ is surjective, the image of U is dense in H_φ , so being an isometry U is also surjective. We then have $U\pi_{\varphi \circ \theta}(x)U^* = \pi_\varphi(\theta(x))$ which yields the asserted isomorphism.

(5) In general, $\text{Ran}(U)$ is the closure of $\widehat{\theta(a)}$ for $a \in A$ (the range again is closed because U is an isometry). Note that this is an invariant subspace with respect to $\pi_\varphi(\theta(A))$. Therefore, we obtain a $*$ -homomorphism $\alpha : \theta(A) \rightarrow B(\text{Ran}(U))$ given by $\theta(x) \mapsto \theta(x)UU^* = UU^*\theta(x)$. The map α is injective because if $\alpha(y) = 0$, then $yUU^*\widehat{1} = y\widehat{1} = \widehat{y} = 0$, which implies that $y = 0$ by faithfulness of φ . Hence, α^{-1} is a well-defined $*$ -homomorphism.

Now consider the map $\beta : \pi_{\varphi \circ \theta}(A) \rightarrow B(\text{Ran } U)$ given by $\beta(T) = UTU^*$. This is a $*$ -homomorphism since $U^*U = 1$ and it is unital since UU^* is the identity on $\text{Ran}(U)$. Moreover, β is a $*$ -isomorphism onto its image since $U^*\beta(T)U = T$. By (1), we obtain

$$\beta(\pi_{\varphi \circ \theta}(x)) = U\pi_{\varphi \circ \theta}(x)U^* = \pi_\varphi(\theta(x))UU^* = \alpha(\theta(x)).$$

Therefore, $\alpha^{-1} \circ \beta$ gives a $*$ -isomorphism $\pi_{\varphi \circ \theta}(A) \rightarrow \theta(A)$. \square

2.1.2. Universal C^* -algebras, free products, and tensor products

Next, we discuss universal C^* -algebras given by generators and relations (see [37]). For a given set of generators $\mathcal{G} = \{x_\alpha\}_\alpha$ and a set of relations \mathcal{R} where the relation r is allowed in \mathcal{R} if r is given as the following

$$\|p(x_{\alpha_1}, \dots, x_{\alpha_n}, x_{\alpha_1}^*, \dots, x_{\alpha_n}^*)\| \leq \eta$$

where $p \in \mathbb{C}\langle y_1, \dots, y_{2n} \rangle$, the variables y_i 's do not commute and $\eta \geq 0$.

A representation ϕ of $(\mathcal{G}, \mathcal{R})$ is a set of operators $\{y_\alpha\}_\alpha$ in $B(H)$ for a Hilbert space H such that $\{y_\alpha = \phi(x_\alpha)\}_\alpha$ satisfy the relations \mathcal{R} . Consider the free $*$ -algebra $\mathcal{F}(\mathcal{G})$ generated by \mathcal{G} and \mathcal{R} . The map ϕ induces a unique $*$ -homomorphism

$$\tilde{\phi} : \mathcal{F}(\mathcal{G}) \rightarrow B(H).$$

For a given set of generators and relations $(\mathcal{G}, \mathcal{R})$, we can construct $C^*(\mathcal{G}, \mathcal{R})$ as long as it is *admissible*, that is, if there is a representation of $(\mathcal{G}, \mathcal{R})$ and whenever $\{\phi_j, H_j\}$ is a family of representations of $(\mathcal{G}, \mathcal{R})$, then $\bigoplus_j \phi_j$ is a representation of $(\mathcal{G}, \mathcal{R})$ on $\bigoplus_j H_j$. For any $f \in \mathcal{F}(\mathcal{G})$, the norm is defined by $\|f\| := \sup_j \{\|\phi_j(f)\| : \phi_j \text{ is a representation of } (\mathcal{G}, \mathcal{R})\}$. The above two conditions imply that $\|f\| < \infty$ and is a C^* -seminorm on $\mathcal{F}(\mathcal{G})$. Now the universal C^* -algebra for $(\mathcal{G}, \mathcal{R})$, denoted by $C^*(\mathcal{G}, \mathcal{R})$, is defined as the completion of $\mathcal{F}(\mathcal{G})/\{f \in \mathcal{F}(\mathcal{G}) : \|f\| = 0\}$.

Given any two unital C^* -algebras A and B , the *universal* or *full free product* $A * B$ is the universal unital C^* -algebra generated by copies of A and B [38]. It is constructed so that for any two representations π_1 and π_2 on a Hilbert space H , there is a unique representation of π on H extending π_1 and π_2 .

The *maximal tensor product* $A \otimes_{\max} B$ of C^* -algebras is the universal C^* -algebra generated by commuting copies of A and B (see [34, Theorem 6.3.5]). It is constructed so that the images $A \otimes 1$ and $1 \otimes B$ commutes, and for any two representations π_1 and π_2 on a Hilbert space H such that $\pi_1(A)$ and $\pi_2(B)$ commute, there is a unique representation of π on H extending π_1 and π_2 .

The *minimal tensor product* $A \otimes_{\min} B$ is the C^* -algebra generated by the tensor product of given faithful representations of A and B respectively (see [34, Theorem 6.3.3]). Due to its common usage with quantum groups, we denote the minimal tensor product of C^* -algebras simply by \otimes .

2.2. Von Neumann algebras

2.2.1. Basic properties

A *von Neumann algebra* is defined as a unital $*$ -subalgebra $M \subseteq B(H)$ closed in the weak operator topology (WOT). A state φ on a von Neumann algebra M is *normal* if φ is WOT-continuous on $(M)_1$. A *tracial von Neumann algebra* is a pair (M, τ) where M is a von Neumann algebra and τ is a faithful, normal, tracial state. We recall that any C^* -algebra with a state gives rise to a von Neumann algebra via the GNS construction. The following statement is basic and well known but we were unable to find a specific reference, so we include the proof here.

Lemma 2.2. *Let A be a C^* -algebra, and let $\varphi \in \mathcal{T}(A)$. Let (H_φ, π_φ) be the GNS construction of (A, φ) , let $M = \pi_\varphi(A)''$ be the von Neumann algebra generated by $\pi_\varphi(A)$, and let $\tau(x) = \langle \widehat{1}, x\widehat{1} \rangle$ for $x \in M$. Then (M, τ) is a tracial von Neumann algebra.*

Proof. First, note that τ is a normal state on M because it is a vector state. Moreover, by virtue of the GNS construction, we have $\tau \circ \pi_\varphi = \varphi$. In particular, τ is tracial on $\pi_\varphi(A)$. Now if $x, y \in M$, fix nets x_i and y_j converging in the strong operator topology to x and y . Then

$$\tau(xy) = \lim_i \tau(x_i y) = \lim_i \lim_j \tau(x_i y_j) = \lim_i \lim_j \tau(y_j x_i) = \lim_i \tau(y x_i) = \tau(yx).$$

For the faithfulness of τ , we use the right representation of A given by $\rho(a)\widehat{b} = \widehat{ba}$. To show that this is well-defined, note that

$$\|\widehat{ba}\|_{H_\varphi}^2 = \tau(a^* b^* ba) = \tau(baa^* b^*) \leq \|aa^*\| \tau(bb^*) = \|a\|^2 \tau(b^* b) = \|a\|^2 \|\widehat{b}\|_{H_\varphi}^2.$$

Therefore, right multiplication by a passes to a well-defined bounded operator on $L^2(H_\varphi)$. It is easy to verify that the left multiplication by $a \in A$ and the right multiplication by $a' \in A$ commute. In particular, $\rho(A) \subseteq \pi_\varphi(A)'$, and so $\pi_\varphi(A)''$ commutes with $\rho(A)$. Therefore, if $x \in \pi_\varphi(A)$ and $\tau(x^* x) = 0$, this means that $x\widehat{1} = 0$. For $a \in A$, we have $0 = \rho(a)x\widehat{1} = x\rho(a)\widehat{1} = x\widehat{a}$. Thus, x is zero on a dense subset of H_φ , and so $x = 0$. Therefore, τ is faithful. Thus, (M, τ) is a tracial von Neumann algebra as desired. \square

The GNS construction for a tracial von Neumann algebra $M \subseteq B(H)$ gives a standard representation which for most purposes can be used instead of the given representation on H . We recall a few standard facts. See e.g. [39, §1.2]. If $M \subseteq B(H)$ is a von Neumann algebra with faithful normal tracial state τ , let $L^2(M, \tau) = H_\tau$ be the associated GNS space. Then the GNS representation $\lambda : M \rightarrow B(L^2(M, \tau))$ is faithful, the image $\lambda(M)$ is a tracial von Neumann algebra, λ defines a WOT homeomorphism from M_1 onto $\lambda(M)_1$. There is also a right multiplication action $\rho(x)$ given by $\rho(x)\widehat{y} = \widehat{yx}$, which defines a $*$ -homomorphism $M^{\text{op}} \rightarrow B(L^2(M))$, where M^{op} is the algebra with reversed multiplication. Moreover, $\lambda(M)$ and $\rho(M)$ commute, and in fact $\lambda(M)' = \rho(M)$ [39, Theorem 1.2.4]. In particular, λ and ρ produce a representation $\lambda \otimes \rho$ of $M \otimes_{\text{max}} M^{\text{op}}$ on $B(L^2(M, \tau))$.

We remark that a tracial von Neumann algebra is determined up to isomorphism by the evaluation of the trace on $*$ -polynomials in the generators. A $*$ -polynomial in variables $(x_i)_{i \in I}$ is a linear combination of products of the terms x_i and x_i^* , for instance, $x_1 x_2^* + 3x_2^2 x_1^* x_3$. The following is another folklore result that has not adequately been written down, so we include the proof here.

Lemma 2.3. *Let (M, τ) and $(\tilde{M}, \tilde{\tau})$ be tracial von Neumann algebras. Let $(x_i)_{i \in I}$ and $(\tilde{x}_i)_{i \in I}$ be elements of M and \tilde{M} respectively. Suppose N and \tilde{N} be the von Neumann subalgebras generated by $(x_i)_{i \in I}$ and $(\tilde{x}_i)_{i \in I}$ respectively. Suppose that for every $*$ -polynomial, we have $\tau(p((x_i)_{i \in I})) = \tilde{\tau}(\tilde{p}((\tilde{x}_i)_{i \in I}))$. Then there is a $*$ -isomorphism $\pi : N \rightarrow \tilde{N}$ such that $\pi(x_i) = \tilde{x}_i$ and π gives a WOT-homeomorphism $(N)_1$ to $(\tilde{N})_1$.*

Proof. Since the standard representation of N on $L^2(N, \tau|_N)$ is faithful and gives the same WOT-topology on the unit ball, it suffices to prove the case where $M = N$ and $\tilde{M} = \tilde{N}$, and we may assume without loss of generality that N and \tilde{N} are subalgebras of $B(L^2(N, \tau))$ and $B(L^2(\tilde{N}, \tilde{\tau}))$ respectively.

By the bicommutant theorem, N is the SOT-closure of $p((x_i)_{i \in I})$ for $*$ -polynomials p . In particular, $\widehat{p((x_i)_{i \in I})}$ is dense in $L^2(N, \tau)$. The analogous statement holds for \tilde{N} . By assumption, we have $\tau[p((x_i)_{i \in I})^* q((x_i)_{i \in I})] = \tilde{\tau}[p((\tilde{x}_i)_{i \in I})^* q((\tilde{x}_i)_{i \in I})]$. Therefore, there is a unitary isomorphism $U : L^2(N, \tau) \rightarrow L^2(\tilde{N}, \tilde{\tau})$ satisfying $U[\widehat{p((x_i)_{i \in I})}] = U[\widehat{p((\tilde{x}_i)_{i \in I})}]$. It is immediate that

$$U\lambda(x_i)\widehat{p((x_i)_{i \in I})} = \lambda(\tilde{x}_i)U\widehat{p((\tilde{x}_i)_{i \in I})},$$

and therefore $\lambda(\tilde{x}_i) = U\lambda(x_i)U^*$. Thus, the map $\text{ad}_U : T \mapsto UTU^*$ sends the $*$ -algebra generated by $(x_i)_{i \in I}$ to the $*$ -algebra generated by $(\tilde{x}_i)_{i \in I}$. The map ad_U is isometric with respect to operator norm and also a WOT-homeomorphism $B(L^2(N, \tau)) \rightarrow B(L^2(\tilde{N}, \tilde{\tau}))$. Hence, it maps the von Neumann algebra generated by $(x_i)_{i \in I}$ onto the von Neumann algebra generated by $(\tilde{x}_i)_{i \in I}$, or it maps N onto \tilde{N} . \square

In order to define $W_{H,1}$, we use the non-commutative L^1 space associated to a tracial von Neumann algebra. We denote by $L^1(M, \tau)$ the completion of M with respect to the norm $\|x\|_{L^1(M, \tau)} = \tau(|x|)$ where $|x| = (x^* x)^{1/2}$. There is a natural inclusion of $L^2(M, \tau)$ into $L^1(M, \tau)$. In fact, one can define non-commutative L^p spaces $L^p(M, \tau)$ for every $p \in [1, \infty]$, where $L^\infty(M, \tau)$ is M itself. These all reside within the algebra of affiliated operators and satisfy the non-commutative Hölder's inequality $\|xy\|_{L^p(M, \tau)} \leq \|x\|_{L^{p_1}(M, \tau)} \|y\|_{L^{p_2}(M, \tau)}$ when $1/p = 1/p_1 + 1/p_2$. However, we only need this for the values 1, 2, and ∞ . For background on non-commutative L^p spaces, see e.g. [40].

2.2.2. Tensor products and amalgamated free products of von Neumann algebras

Given two tracial von Neumann algebras (M, τ) and (N, σ) , one defines their von Neumann algebraic tensor product $(M \overline{\otimes} N, \tau \otimes \sigma)$ as follows. Consider $M \subseteq B(L^2(M, \tau))$ and $N \subseteq B(L^2(N, \sigma))$ using the standard left multiplication representation. Then form the minimal C^* tensor product $M \otimes N$ on $L^2(M, \tau) \otimes L^2(N, \sigma)$, and let $M \overline{\otimes} N$ be the WOT-closure of $M \otimes N$ in $B(L^2(M, \tau) \otimes L^2(N, \sigma))$. The vector $\widehat{1}_M \otimes \widehat{1}_N$ defines a normal tracial state on $M \otimes N$ which sends $a \otimes b$ to $\tau(a)\sigma(b)$. Moreover, since $A \odot B$ is WOT-dense in $A \overline{\otimes} B$, there is only one normal state on $A \overline{\otimes} B$ that gives $\tau \otimes \sigma$ on $A \odot B$. Faithfulness of $\tau \otimes \sigma$ on $M \overline{\otimes} N$ follows using the right multiplication action of $A \otimes B$ as in the proof of Lemma 2.2.

As in [28], in order to show the triangle inequality for free Wasserstein distances, we need to be able to embed any two given tracial von Neumann algebras (M_1, τ_1) and (M_2, τ_2) into a larger one (M, τ) , and to arrange that the embeddings of M_1 and M_2 agree on a common von Neumann subalgebra. One construction that accomplishes this is the amalgamated free product. We recall first the existence of trace-preserving conditional expectations.

Fact 2.4 (See e.g. [39, §3.1]). *Let (M, τ) be a tracial von Neumann algebra and let N be a von Neumann subalgebra. Then for each $x \in M$, there is a unique $E_N(x) \in N$ such that $\tau(xy) = \tau(E_N(x)y)$ for all $y \in N$.*

The map E_N is called the trace-preserving conditional expectation from M to N .

Definition 2.5 (Free independence with amalgamation). *Let (M, τ) be a tracial von Neumann algebra and N a von Neumann subalgebra. For $i \in I$, let M_i be a von Neumann subalgebra with $N \subseteq M_i \subseteq M$. We say that $(M_i)_{i \in I}$ are *freely independent with amalgamation over N* if whenever $i_1 \neq i_2 \neq \dots \neq i_k$ are indices in I where consecutive indices are not equal, and $x_j \in M_{i_j}$ with $E_N[x_j] = 0$, then*

$$E_N[x_{i_1} \dots x_{i_k}] = 0.$$

Fact 2.6 (Amalgamated free product). *Let $(M_i, \tau_i)_{i \in I}$ and (N, σ) be given tracial von Neumann algebras, and let $\iota_i : N \rightarrow M_i$ be a trace-preserving $*$ -homomorphism. Then there exists a tracial von Neumann algebra (M, τ) and trace-preserving $*$ -homomorphisms $\gamma_i : M_i \rightarrow M$ and $\gamma : N \rightarrow M$ such that*

- $\gamma_i \circ \iota_i = \gamma$ for all $i \in I$.
- M is generated by $\gamma_i(M_i)$ for $i \in I$.
- The subalgebras $\gamma_i(M_i)$ for $i \in I$ are freely independent with amalgamation over $\gamma(N)$.

Moreover, M is unique up to isomorphism in the sense that if $(\tilde{M}, \tilde{\gamma}_i, \tilde{\gamma})$ also satisfy this, then there is a unique isomorphism $\Phi : M \rightarrow \tilde{M}$ with $\Phi \circ \gamma_i = \tilde{\gamma}_i$.

Note that the free product in this theorem is not a *universal* free product but rather a *reduced* free product, one constructed as acting on a particular Hilbert space rather than from all possible representations (there is also a universal amalgamated free product for C^* -algebras). The construction of amalgamated (reduced) free product for C^* -algebras was given in [41], and the version for tracial von Neumann algebras was given in [42]. For a detailed proof of this fact, see the construction of the C^* -amalgamated free product in [43, Theorem 4.7.2], apply [44, Proposition 3.8.5] to establish traciality of the state, and then extend the C^* -algebra to a tracial von Neumann algebra by Lemma 2.2. The uniqueness of the amalgamated free product of tracial von Neumann algebras follows from the C^* -case in [43, Theorem 4.7.2] by extending to the von Neumann algebra along similar lines as Lemma 2.3.

2.2.3. Operations with projections in von Neumann algebras

The set of projections in a von Neumann algebra is closed under unions and intersections. We recall the following elementary facts for later use.

Fact 2.7. *Let p and q be two projections in a von Neumann algebra M . Then $(pq)^k \rightarrow p \wedge q$ in SOT as $k \rightarrow \infty$. If τ is a normal trace on M , then $\tau(p \wedge q) = \lim_{k \rightarrow \infty} \tau((pq)^k) = \inf_{k \in \mathbb{N}} \tau((pq)^k)$.*

Proof sketch. Since $(pq)^k p \geq (pq)^{k+1} p$, the sequence $(pq)^k p$ converges in SOT to some limit r , which is in M since M is SOT-closed. By standard manipulations, one can show that $r \leq p$ and $r \leq q$ and $p \wedge q \leq r$. If τ is a normal trace, then by SOT continuity $\tau((pq)^k p) \rightarrow \tau(p \wedge q)$, and by monotonicity, the limit over k is the same as the infimum. \square

Fact 2.8. *Let (M, τ) be a tracial von Neumann algebra. If p and q are projections in M , then*

$$\tau(p \vee q) + \tau(p \wedge q) = \tau(p) + \tau(q).$$

Proof sketch. As noted in [36, III.1.1.3], letting v be the partial isometry in the polar decomposition of $q(1 - p)$, we have $v^*v = (p \vee q) - p$ and $vv^* = q - (p \wedge q)$ and therefore $(p \vee q) - p$ and $q - (p \wedge q)$ have the same trace. \square

The next two facts are well-known and straightforward to check.

Fact 2.9. *Let $(p_i)_{i=1}^n$ and $(q_i)_{i=1}^n$ be projections in some von Neumann algebra such that $p_i p_j = 0 = q_i q_j$ for $i \neq j$. Then*

$$\left(\sum_{i=1}^n p_i \right) \wedge \left(\sum_{i=1}^n q_i \right) \geq \sum_{i=1}^n p_i \wedge q_i.$$

Fact 2.10. Let p_1, p_2 be projections in some von Neumann algebra M and let q_1, q_2 be projections in some von Neumann algebra N . Then in $M \overline{\otimes} N$, we have $(p_1 \otimes q_1) \wedge (p_2 \otimes q_2) = (p_1 \wedge p_2) \otimes (q_1 \wedge q_2)$.

2.3. Compact quantum groups

For a tensor product $A_1 \otimes \dots \otimes A_n$ and for distinct indices $j_1, \dots, j_k \in [n]$, we denote by ι_{j_1, \dots, j_k} , the inclusion $A_{j_1} \otimes \dots \otimes A_{j_k} \rightarrow A_1 \otimes \dots \otimes A_n$ that maps $a_1 \otimes \dots \otimes a_k$ to the simple tensor with a_i in the j_i 's component and 1's in the remaining components. For instance, for a three-fold tensor product,

$$\iota_{1,2}(a \otimes b) = a \otimes b \otimes 1, \quad \iota_{3,2}(a \otimes b) = 1 \otimes b \otimes a.$$

Definition 2.11 (Woronowicz [20]). A compact quantum group is a unital C^* -algebra A together with a $*$ -homomorphism $\Delta : A \rightarrow A \otimes A$ such that:

1. Δ is coassociative i.e. $(\mathbb{1} \otimes \Delta)\Delta = (\Delta \otimes \mathbb{1})\Delta$.
2. $\text{span}(\Delta(A)(A \otimes \mathbb{1})) = \text{span}(\Delta(A)(\mathbb{1} \otimes A)) = A \otimes A$.

Remark 2.12. Any compact topological group G is a compact quantum group as follows. The pair $(C(G), \Delta)$, where $\Delta : C(G) \rightarrow C(G) \otimes C(G)$ defined as $\Delta(f)(s, t) = f(st)$.

Definition 2.13 ([18]). A Woronowicz Hopf C^* -algebra (compact matrix quantum group) is a unital C^* -algebra A together with a dense $*$ -subalgebra \mathcal{A} generated by u_{ij} for some $n \in \mathbb{N}$ and $i, j \in \{1, \dots, n\}$, a unital $*$ -homomorphism $\Delta : A \rightarrow A \otimes A$ such that, letting $U = [u_{ij}]_{i,j=1}^n \in M_n(A)$,

1. U^T is invertible in $M_n(A)$.
2. U is a unitary element of $M_n(A)$.
3. For every i, j , $\Delta(u_{ij}) = \sum_{k=1}^n u_{ik} \otimes u_{kj}$.

The quantum permutation group S_n^+ is defined as follows. The underlying C^* -algebra $C(S_n^+)$ is the universal C^* -algebra generated by $u_{ij} : 1 \leq i, j \leq n$ such that

$$u_{ij}^* = u_{ij} = u_{ij}^2, \quad \sum_i u_{ij} = 1, \quad \sum_j u_{ij} = 1$$

Writing $\mathbb{M}_n = \mathbb{M}_n(\mathbb{C})$ for $n \times n$ complex matrices, the comultiplication for S_n^+ is given by

$$\Delta(u_{ij}) = \sum_{k=1}^n u_{ik} \otimes u_{kj}.$$

U will denote the matrix $U = [u_{ij}]_{i,j=1}^n \in \mathbb{M}_n(C(S_n^+)) = \mathbb{M}_n \otimes C(S_n^+)$. The relations above are equivalent to $U^*U = UU^* = 1$ and each u_{ij} being a projection. The comultiplication is given by

$$(\text{id}_{\mathbb{M}_n} \otimes \Delta)(U) = \iota_{1,2}(U)\iota_{1,3}(U) \in \mathbb{M}_n \otimes C(S_n^+) \otimes C(S_n^+).$$

The quantum permutation group is the universal final object in the category \mathcal{C} of quantum groups that act on an n -point space X_n ; see [21]

For each $n \in \mathbb{N}$, the quantum permutation group S_n^+ contains the classical permutation group S_n as a quantum subgroup, which means that there is a quotient map $C(S_n^+) \rightarrow C(S_n)$.

Theorem 2.14 ([21, Theorem 3.1]). Let $\theta : C(S_n^+) \rightarrow C(S_n)$ be the surjective $*$ -homomorphism given by the indicator function $\theta(u_{ij}) = \mathbb{1}_{\sigma(i)=j}$. Then, θ is an embedding $S_n \subset S_n^+$ of quantum groups, that is, $\Delta_{C(S_n)} \circ \theta = (\theta \otimes \theta) \circ \Delta_{C(S_n^+)}$.

Remark 2.15. For $n \leq 3$, S_n and S_n^+ are isomorphic. Otherwise, S_n^+ is neither classical nor finite-dimensional.

Compact quantum groups have a natural analog of the Haar measure, known as the Haar state. Its existence and uniqueness for CQGs were proved by Woronowicz in the 1980s [19]. As it turns out, for the quantum permutation group S_n^+ , the Haar state is also tracial and faithful, and it can be described combinatorially; see [45]. Moreover, $C(S_n^+)$ has a large number of finite-dimensional representations and traces [33]. Other examples of quantum permutation matrices can be seen through the connection with quantum games [46].

3. Distance via free couplings

3.1. Definition

Recall (Definition 1.1) that for a C^* -algebra A and $\varphi_1, \varphi_2 \in \mathcal{T}(A)$, a *tracial coupling* is a tuple $(M, \tau, \alpha_1, \alpha_2)$ where (M, τ) is a tracial von Neumann algebra and $\alpha_j : A \rightarrow M$ is a $*$ -homomorphism such that $\varphi_j = \tau \circ \alpha_j$ for $j = 1, 2$. Note that the set of tracial couplings is nonempty because we can take M to be the von Neumann algebra generated by the GNS representation of $\varphi_1 \otimes \varphi_2$ on $A \otimes_{\min} A$ (see Lemma 2.2). We next give an alternative description of tracial couplings in terms of traces on the universal free product $A * A$ that restrict to φ_1 and φ_2 on the first and second copies of A respectively. We remark that this is closely related to the bijection between factorizable quantum channels and traces on free products [47].

Definition 3.1. Let $A * A$ be the universal unital C^* free product (see §2.1.2) and let $\iota_1, \iota_2 : A \rightarrow A * A$ be the coordinate inclusions. For $\varphi_1, \varphi_2 \in \mathcal{T}(A)$, let

$$\mathcal{T}_{\varphi_1, \varphi_2}(A * A) = \{\varphi \in \mathcal{T}(A * A) : \varphi \circ \iota_j = \varphi_j \text{ for } j = 1, 2\}.$$

Lemma 3.2. Let A be a unital C^* -algebra. Let $\varphi_1, \varphi_2 \in \mathcal{T}(A)$.

- (1) Given $\varphi \in \mathcal{T}_{\varphi_1, \varphi_2}(A * A)$, there exists a tracial coupling $(M, \tau, \alpha_1, \alpha_2)$ such that $\varphi = \tau \circ (\alpha_1 * \alpha_2)$.
- (2) Conversely, for every coupling $(M, \tau, \alpha_1, \alpha_2)$, we have $\tau \circ (\alpha_1 * \alpha_2) \in \mathcal{T}_{\varphi_1, \varphi_2}(A * A)$.
- (3) Let $(M, \tau, \alpha_1, \alpha_2)$ and $(\tilde{M}, \tilde{\tau}, \tilde{\alpha}_1, \tilde{\alpha}_2)$ be two tracial couplings of φ_1, φ_2 and let φ and $\tilde{\varphi}$ be the associated traces in $\mathcal{T}_{\varphi_1, \varphi_2}(A * A)$. Suppose that $M = W^*(\alpha_1(A), \alpha_2(A))$ and $\tilde{M} = W^*(\tilde{\alpha}_1(A), \tilde{\alpha}_2(A))$. Then $\varphi = \tilde{\varphi}$ if and only if there is a $*$ -isomorphism $f : M \rightarrow \tilde{M}$ such that $\tau = \tilde{\tau} \circ f$ and $f \circ \alpha_j = \tilde{\alpha}_j$ for $j = 1, 2$.

Proof. (1) Let $\varphi \in \mathcal{T}(A * A)$. Let $\pi_\varphi : A * A \rightarrow B(H_\varphi)$ be the GNS construction associated to $A * A$ and φ . Let $M = \pi_\varphi(A * A)''$, which is a tracial von Neumann algebra by Lemma 2.2. Let $\alpha_j = \pi_\varphi \circ \iota_j$, and then $\tau \circ \alpha_j = \varphi \circ \iota_j = \varphi_j$, so $(M, \tau, \alpha_1, \alpha_2)$ is a tracial coupling.

(2) Let $(M, \tau, \alpha_1, \alpha_2)$ be a tracial coupling. Let $\alpha_1 * \alpha_2 : A * A \rightarrow M$ be the $*$ -homomorphism given by the universal property of the full free product, and let $\varphi := \tau \circ (\alpha_1 * \alpha_2)$. Then $\varphi \circ \iota_j = \tau \circ \alpha_j = \varphi_j$ and hence $\varphi \in \mathcal{T}_{\varphi_1, \varphi_2}(A * A)$.

(3) Suppose that $\varphi = \tilde{\varphi}$. Note that $(\alpha_1(a))_{a \in A} \cup (\alpha_2(b))_{b \in A}$ is a generating set for M and analogously $(\tilde{\alpha}_1(a))_{a \in A} \cup (\tilde{\alpha}_2(b))_{b \in A}$ is a generating set for \tilde{M} . Since $\varphi = \tilde{\varphi}$, we have in particular that for $*$ -polynomials

$$\begin{aligned} \tau(p((\alpha_1(a))_{a \in A}, (\alpha_2(b))_{b \in B})) &= \varphi(p((\iota_1(a))_{a \in A}, (\iota_2(b))_{b \in B})) \\ &= \tilde{\varphi}(p((\iota_1(a))_{a \in A}, (\iota_2(b))_{b \in B})) \\ &= \tilde{\tau}(p((\tilde{\alpha}_1(a))_{a \in A}, (\tilde{\alpha}_2(b))_{b \in B})). \end{aligned}$$

Therefore, by Lemma 2.3, we obtain the desired isomorphism $M \rightarrow \tilde{M}$ sending $\alpha_1(a)$ to $\tilde{\alpha}_1(a)$ and $\alpha_2(b)$ to $\tilde{\alpha}_2(b)$ for $a, b \in A$. Conversely, if there is such a $*$ -isomorphism f , then

$$\begin{aligned} \varphi(p((\iota_1(a))_{a \in A}, (\iota_2(b))_{b \in B})) &= \tau(p((\alpha_1(a))_{a \in A}, (\alpha_2(b))_{b \in B})) \\ &= \tilde{\tau}(p((\tilde{\alpha}_1(a))_{a \in A}, (\tilde{\alpha}_2(b))_{b \in B})) \\ &= \tilde{\varphi}(p((\iota_1(a))_{a \in A}, (\iota_2(b))_{b \in B})). \end{aligned}$$

Thus, φ and $\tilde{\varphi}$ agree on the $*$ -subalgebra generated by $\iota_1(A)$ and $\iota_2(A)$, and therefore they agree on all of $A *_{\text{full}} A$. \square

Definition 3.3 (Quantum Hamming distance). For $\varphi_1, \varphi_2 \in \mathcal{T}(C(S_n^+))$, let

$$W_{H, \wedge}(\varphi_1, \varphi_2) = \inf_{\text{tracial couplings } (M, \tau, \alpha_1, \alpha_2)} \left(1 - \frac{1}{n} \sum_{i, j=1}^n \tau(\alpha_1(u_{ij}) \wedge \alpha_2(u_{ij})) \right). \tag{3.1}$$

Using Lemma 3.2 and Fact 2.7, we can rewrite the quantum Hamming cost in terms of $\mathcal{T}_{\varphi_1, \varphi_2}(A * A)$.

Lemma 3.4. For $j \in \{1, 2\}$, let ι_j be as defined in (3.1). Given any $\varphi_1, \varphi_2 \in \mathcal{T}(C(S_n^+))$, the Hamming cost in (3.3) is equivalent to the following:

$$W_{H, \wedge}(\varphi_1, \varphi_2) = \inf_{\varphi \in \mathcal{T}(A * A)} \sup_k \left(1 - \frac{1}{n} \sum_{i, j=1}^n \varphi([\iota_1(u_{ij}) \iota_2(u_{ij})]^k) \right).$$

Proof. If $\varphi_1, \varphi_2 \in \mathcal{T}(C(S_n^+))$, then

$$\begin{aligned} W_{H,\wedge}(\varphi_1, \varphi_2) &= \inf_{\text{tracial couplings } (M, \tau, \alpha_1, \alpha_2)} \left(1 - \frac{1}{n} \sum_{i,j=1}^n \tau(\alpha_1(u_{ij}) \wedge \alpha_2(u_{ij})) \right) \\ &= \inf_{\text{tracial couplings } (M, \tau, \alpha_1, \alpha_2)} \left(1 - \frac{1}{n} \sum_{i,j=1}^n \tau(\lim_{k \rightarrow \infty} [\alpha_1(u_{ij})\alpha_2(u_{ij})]^k) \right) \\ &= \inf_{\text{tracial couplings } (M, \tau, \alpha_1, \alpha_2)} \sup_k \left(1 - \frac{1}{n} \sum_{i,j=1}^n \tau([\alpha_1(u_{ij})\alpha_2(u_{ij})]^k) \right) \\ &\stackrel{3.2}{=} \inf_{\varphi \in \mathcal{T}(A * A)} \sup_k \left(1 - \frac{1}{n} \sum_{i,j=1}^n \varphi[(\iota_1(u_{ij})\iota_2(u_{ij}))^k] \right). \quad \square \end{aligned}$$

Next, we note that the infimum in the definition of the Hamming cost is always achieved; this is analogous to [28, Proposition 1.4] for the Wasserstein distance in free probability.

Lemma 3.5. Let $A = C(S_n^+)$, and let $\varphi_1, \varphi_2 \in \mathcal{T}(A)$. Then there exists a tracial coupling $(M, \tau, \alpha_1, \alpha_2)$ that achieves the infimum in (3.1).

Proof. Define $f_k : \mathcal{T}_{\varphi_1, \varphi_2}(A * A)$ by

$$f_k(\varphi) = 1 - \frac{1}{n} \sum_{i,j=1}^n \varphi[(\iota_1(u_{ij})\iota_2(u_{ij}))^k].$$

Let $f(\varphi) = \sup_{k \in \mathbb{N}} f_k(\varphi)$. Since f_k is weak-* continuous on $\mathcal{T}_{\varphi_1, \varphi_2}(A * A)$, then f is an supremum of lower semi-continuous functions and hence lower semi-continuous. Note that $\mathcal{T}_{\varphi_1, \varphi_2}(A * A)$ is a closed subset of the compact set $\mathcal{T}(A * A)$, hence compact. Therefore, f achieves an infimum on $\mathcal{T}_{\varphi_1, \varphi_2}(A * A)$ at some point φ . Let $(M, \tau, \alpha_1, \alpha_2)$ be the associated tracial coupling by Lemma 3.2. Then, in light of Lemma 3.4, $(M, \tau, \alpha_1, \alpha_2)$ achieves the infimum in (3.1). \square

3.2. Metric properties

Proposition 3.6. $W_{H,\wedge}$ defines a metric on $\mathcal{T}(C(S_n^+))$.

Proof. Nonnegativity: Let φ_1, φ_2 be traces in $\mathcal{T}(C(S_n^+))$, and let $(M, \tau, \alpha_1, \alpha_2)$ be a tracial coupling. Then

$$\begin{aligned} &1 - \frac{1}{n} \sum_{i,j=1}^n \tau(\alpha_1(u_{ij}) \wedge \alpha_2(u_{ij})) \\ &= \sum_{i,j=1}^n \frac{1}{n} (\tau(\alpha_1(u_{ij})) - \tau(\alpha_1(u_{ij}) \wedge \alpha_2(u_{ij}))) \geq 0 \end{aligned}$$

because $\alpha_1(u_{ij}) \geq (\alpha_1(u_{ij}) \wedge \alpha_2(u_{ij}))$ for $i, j = 1, \dots, n$. Hence, taking the infimum over tracial couplings, we have $W_{H,\wedge}(\varphi_1, \varphi_2) \geq 0$.

Nondegeneracy: We show $W_{H,\wedge}(\varphi_1, \varphi_2) = 0$ if and only if $\varphi_1 = \varphi_2$. If $\varphi_1 = \varphi_2$, then we can take (M, τ) to be the von Neumann algebra generated by the GNS construction of $(C(S_n^+), \varphi_1)$ as in Lemma 2.2, and let $\alpha_1 = \alpha_2$ be the GNS representation itself. Then $\alpha_1(u_{ij}) \wedge \alpha_2(u_{ij}) = \alpha_1(u_{ij})$, and so

$$1 - \frac{1}{n} \sum_{i,j=1}^n \tau(\alpha_1(u_{ij}) \wedge \alpha_2(u_{ij})) = 1 - \frac{1}{n} \sum_{i,j=1}^n \tau(\alpha_1(u_{ij})) = 1 - \frac{1}{n} \sum_{i=1}^n 1 = 0,$$

so $W_{H,\wedge}(\varphi_1, \varphi_1) = 0$.

Conversely, suppose that $W_{H,\wedge}(\varphi_1, \varphi_2) = 0$. By Lemma 3.5, there is a tracial coupling $(M, \tau, \alpha_1, \alpha_2)$ such that

$$0 = W_{H,\wedge}(\varphi_1, \varphi_2) = 1 - \frac{1}{n} \sum_{i,j=1}^n \tau(\alpha_1(u_{ij}) \wedge \alpha_2(u_{ij})) = \frac{1}{n} \sum_{i,j=1}^n (\tau(\alpha_1(u_{ij})) - \tau(\alpha_1(u_{ij}) \wedge \alpha_2(u_{ij}))).$$

Since we are adding nonnegative elements, we get $\tau(\alpha_1(u_{ij}) - \alpha_1(u_{ij}) \wedge \alpha_2(u_{ij})) = 0$ for all i, j . Thus, faithfulness of τ implies that $\alpha_1(u_{ij}) = \alpha_1(u_{ij}) \wedge \alpha_2(u_{ij})$ for all i, j . Symmetrically, $\alpha_2(u_{ij}) = \alpha_1(u_{ij}) \wedge \alpha_2(u_{ij})$, and therefore $\alpha_1(u_{ij}) = \alpha_2(u_{ij})$ for all i, j . Since $\{u_{ij}\}_{i,j}$ generate $C(S_n^+)$, we have $\alpha_1 = \alpha_2$, hence $\varphi_1 = \tau \circ \alpha_1 = \tau \circ \alpha_2 = \varphi_2$.

Symmetry: Observe that $(M, \tau, \alpha_1, \alpha_2)$ is a coupling of (φ_1, φ_2) if and only if $(M, \tau, \alpha_2, \alpha_1)$ is a coupling of (φ_2, φ_1) and $\alpha_1(u_{ij}) \wedge \alpha_2(u_{ij}) = \alpha_2(u_{ij}) \wedge \alpha_1(u_{ij})$. Therefore,

$$\begin{aligned} W_{H,\wedge}(\varphi_1, \varphi_2) &= \inf_{\text{tracial couplings } (M, \tau, \alpha_1, \alpha_2)} \left(1 - \frac{1}{n} \sum_{i,j=1}^n \tau(\alpha_1(u_{ij}) \wedge \alpha_2(u_{ij})) \right) \\ &= \inf_{\text{tracial couplings } (M, \tau, \alpha_2, \alpha_1)} \left(1 - \frac{1}{n} \sum_{i,j=1}^n \tau(\alpha_2(u_{ij}) \wedge \alpha_1(u_{ij})) \right) \\ &= W_{H,\wedge}(\varphi_2, \varphi_1). \end{aligned}$$

Triangle inequality: Let $\varphi_1, \varphi_2, \varphi_3 \in \mathcal{T}(C(S_n^+))$ and consider couplings $(M_1, \tau_1, \alpha_1, \alpha_2), (M_2, \tau_2, \beta_1, \beta_2)$. Since $\tau_1 \circ \alpha_2 = \varphi_2 = \tau_2 \circ \beta_1$, the von Neumann subalgebras generated of M_1 and M_2 generated respectively by $\alpha_2(C(S_n^+))$ and $\beta_1(C(S_n^+))$ are $*$ -isomorphic in a trace-preserving way by Lemma 2.3. Therefore, there is a free product (M, τ) of (M_1, τ_1) and (M_2, τ_2) with amalgamation over $\alpha_2(C(S_n^+)) \cong \beta_1(C(S_n^+))$. Let η_j be the inclusion of M_j into M . Let

$$\begin{aligned} \gamma_1 &= \eta_1 \circ \alpha_1 : C(S_n^+) \rightarrow M \\ \gamma_2 &= \eta_1 \circ \alpha_2 = \eta_2 \circ \beta_1 : C(S_n^+) \rightarrow M \\ \gamma_3 &= \eta_2 \circ \beta_2 : C(S_n^+) \rightarrow M \end{aligned}$$

Then $\tau \circ \gamma_1 = \tau \circ \eta_1 \circ \alpha_1 = \tau_1 \circ \alpha_1 = \varphi_1$. Similarly, $\tau \circ \gamma_2 = \tau \circ \eta_2 \circ \beta_2 = \tau_2 \circ \beta_2 = \varphi_3$. Therefore, $(M, \tau, \gamma_1, \gamma_3)$ is a tracial coupling of (φ_1, φ_3) .

For each $i, j = 1, \dots, n$ and $k = 1, 2, 3$, the element $\gamma_k(u_{ij})$ is a projection, and therefore their maximum and minimum are also projections. By Fact 2.8, for each i, j , we have

$$\begin{aligned} \tau(\gamma_1(u_{ij}) \wedge \gamma_2(u_{ij})) + \tau(\gamma_2(u_{ij}) \wedge \gamma_3(u_{ij})) &= \tau(\gamma_1(u_{ij}) \wedge \gamma_2(u_{ij})) \wedge (\gamma_2(u_{ij}) \wedge \gamma_3(u_{ij})) \\ &\quad + \tau((\gamma_1(u_{ij}) \wedge \gamma_2(u_{ij})) \vee (\gamma_2(u_{ij}) \wedge \gamma_3(u_{ij}))) \\ &\leq \tau(\gamma_1(u_{ij}) \wedge \gamma_3(u_{ij})) + \tau(\gamma_2(u_{ij})) \end{aligned} \tag{3.2}$$

since $\gamma_1(u_{ij}) \wedge \gamma_2(u_{ij}) \wedge \gamma_3(u_{ij}) \leq \gamma_1(u_{ij}) \wedge \gamma_3(u_{ij})$ and $(\gamma_1(u_{ij}) \wedge \gamma_2(u_{ij})) \vee (\gamma_2(u_{ij}) \wedge \gamma_3(u_{ij})) \leq \gamma_2(u_{ij})$ and τ is positive. Therefore, by (3.2),

$$-\frac{1}{n} \sum_{i,j=1}^n \tau(\gamma_1(u_{ij}) \wedge \gamma_3(u_{ij})) - \frac{1}{n} \sum_{i,j=1}^n \tau(\gamma_2(u_{ij})) \leq -\frac{1}{n} \sum_{i,j=1}^n \tau(\gamma_1(u_{ij}) \wedge \gamma_2(u_{ij})) - \frac{1}{n} \sum_{i,j=1}^n \tau(\gamma_2(u_{ij}) \wedge \gamma_3(u_{ij})).$$

By adding 1 both sides and using $\sum_{i,j=1}^n \tau(\gamma_2(u_{ij})) = 1$, we get

$$1 - \frac{1}{n} \sum_{i,j=1}^n \tau(\gamma_1(u_{ij}) \wedge \gamma_3(u_{ij})) \leq 1 - \frac{1}{n} \sum_{i,j=1}^n \tau(\gamma_1(u_{ij}) \wedge \gamma_2(u_{ij})) + 1 - \frac{1}{n} \sum_{i,j=1}^n \tau(\gamma_2(u_{ij}) \wedge \gamma_3(u_{ij})) \tag{3.3}$$

$$= 1 - \frac{1}{n} \sum_{i,j=1}^n \tau_1(\alpha_1(u_{ij}) \wedge \alpha_2(u_{ij})) + 1 - \frac{1}{n} \sum_{i,j=1}^n \tau_2(\beta_1(u_{ij}) \wedge \beta_2(u_{ij})), \tag{3.4}$$

where in the last step we used that the embeddings of M_1 and M_2 into M are trace-preserving. Since the tracial couplings of (φ_1, φ_2) and (φ_2, φ_3) were arbitrary, (3.2) and taking the infimum yields

$$W_{H,\wedge}(\varphi_1, \varphi_3) \leq W_{H,\wedge}(\varphi_1, \varphi_2) + W_{H,\wedge}(\varphi_2, \varphi_3). \quad \square$$

We next show the properties of $W_{H,\wedge}$ with respect to the weak- $*$ topology.

Lemma 3.7. $W_{H,\wedge}$ is a weak- $*$ lower semicontinuous function on $\mathcal{T}(C(S_n^+)) \times \mathcal{T}(C(S_n^+))$. Moreover, if $W_{H,\wedge}(\varphi_k, \varphi) \rightarrow 0$, then $\varphi_k \rightarrow \varphi$ in the weak- $*$ topology.

Proof. Since $\mathcal{T}(C(S_n^+))$ is compact and metrizable, it suffices to show that, given sequences $\varphi_{1,m} \rightarrow \varphi_1$ and $\varphi_{2,m} \rightarrow \varphi_2$, we have

$$W_{H,\wedge}(\varphi_1, \varphi_2) \leq \liminf_{m \rightarrow \infty} W_{H,\wedge}(\varphi_{1,m}, \varphi_{2,m}).$$

Let f be as in the proof of Lemma 3.5, and let $\psi_m \in \mathcal{T}_{\varphi_{1,m}, \varphi_{2,m}}$ achieve the infimum of f over $\mathcal{T}_{\varphi_{1,m}, \varphi_{2,m}}(A * A)$. Since $\mathcal{T}(A * A)$ is compact, there exists a subsequence ψ_{m_ℓ} that weak-* converges to some $\psi \in \mathcal{T}(A * A)$. Note that

$$\psi \circ \iota_j = \lim_{\ell \rightarrow \infty} \psi_{m_\ell} \circ \iota_j = \lim_{\ell \rightarrow \infty} \varphi_{j, m_\ell} = \varphi_j \text{ for } j = 1, 2,$$

hence $\psi \in \mathcal{T}_{\varphi_1, \varphi_2}(A * A)$. Then

$$W_{H,\wedge}(\varphi_1, \varphi_2) \leq f(\psi) \leq \liminf_{\ell \rightarrow \infty} f(\psi_{m_\ell}) \leq \liminf_{m \rightarrow \infty} f(\psi_m) = \liminf_{m \rightarrow \infty} W_{H,\wedge}(\varphi_{1,m}, \varphi_{2,m}),$$

which completes the proof of lower semi-continuity.

For the second claim, suppose $\varphi_k \rightarrow \varphi$ with respect to $W_{H,\wedge}$. Since $\mathcal{T}(C(S_n^+))$ is compact and metrizable, it suffices to show that for every subsequence of φ_k has a further subsequence that converges to φ . If φ_{k_ℓ} is a subsequence, pass to a subsequence that converges to some ψ . By lower semi-continuity, $W_{H,\wedge}(\psi, \varphi) \leq \liminf_{\ell \rightarrow \infty} W_{H,\wedge}(\varphi_{k_\ell}, \varphi) = 0$, and hence $\psi = \varphi$ by non-degeneracy of the metric. \square

3.3. Behavior under convolution

Recall that the comultiplication map Δ on $C(S_n^+)$, is given by $\Delta(u_{ij}) := \sum_{k=1}^n u_{ik} \otimes u_{kj}$ and the convolution of two states $\tau, \sigma \in \mathcal{T}(C(S_n^+)) \in \mathcal{T}(M \otimes N)$.

$$\tau * \sigma = (\tau \otimes \sigma) \circ \Delta.$$

Proposition 3.8. *Let $\varphi_1, \varphi_2, \psi_1, \psi_2 \in \mathcal{T}(C(S_n^+))$. Then*

$$W_{H,\wedge}(\varphi_1 * \psi_1, \varphi_2 * \psi_2) \leq W_{H,\wedge}(\varphi_1, \varphi_2) + W_{H,\wedge}(\psi_1, \psi_2).$$

Proof. Consider tracial couplings $(M, \tau, \alpha_1, \alpha_2)$ for (φ_1, φ_2) and $(N, \sigma, \beta_1, \beta_2)$ for (ψ_1, ψ_2) . Consider the tracial von Neumann algebra $(M \overline{\otimes} N, \tau \otimes \sigma)$, and define $\gamma_1, \gamma_2 : C(S_n^+) \rightarrow M \overline{\otimes} N$ by $\gamma_j := (\alpha_j \otimes \beta_j) \circ \Delta$. For $j = 1, 2$, we have

$$\begin{aligned} (\tau \otimes \sigma) \circ \gamma_j &= (\tau \otimes \sigma) \circ (\alpha_j \otimes \beta_j) \circ \Delta \\ &= ((\tau \circ \alpha_j) \otimes (\sigma \circ \beta_j)) \circ \Delta \\ &= (\varphi_j \otimes \psi_j) \circ \Delta \\ &= \varphi_j * \psi_j. \end{aligned}$$

Therefore, $(N \overline{\otimes} M, \tau \otimes \sigma, \gamma_1, \gamma_2)$ is a tracial coupling for $\varphi_1 * \psi_1$ and $\varphi_2 * \psi_2$. For any $i, j \in [n]$, note that $\alpha_1(u_{ik}) \otimes \beta_1(u_{kj})$ and $\alpha_1(u_{ik'}) \otimes \beta_1(u_{k'j})$ are orthogonal for $k \neq k'$ (and similarly with α_2 and β_2 instead of α_1 and β_1), and so we compute using Facts 2.9 and 2.10 that

$$\begin{aligned} \gamma_1(u_{ij}) \wedge \gamma_2(u_{ij}) &= (\alpha_1 \otimes \beta_1) \Delta(u_{ij}) \wedge (\alpha_2 \otimes \beta_2) \Delta(u_{ij}) \\ &= (\alpha_1 \otimes \beta_1) \left(\sum_{k=1}^n u_{ik} \otimes u_{kj} \right) \wedge (\alpha_2 \otimes \beta_2) \left(\sum_{k=1}^n u_{ik} \otimes u_{kj} \right) \\ &= \left(\sum_{k=1}^n (\alpha_1(u_{ik}) \otimes \beta_1(u_{kj})) \right) \wedge \left(\sum_{k=1}^n (\alpha_2(u_{ik}) \otimes \beta_2(u_{kj})) \right) \\ &\geq \sum_{k=1}^n (\alpha_1(u_{ik}) \otimes \beta_1(u_{kj})) \wedge (\alpha_2(u_{ik}) \otimes \beta_2(u_{kj})) \\ &= \sum_{k=1}^n (\alpha_1(u_{ik}) \wedge \alpha_2(u_{ik})) \otimes (\beta_1(u_{kj}) \wedge \beta_2(u_{kj})). \end{aligned}$$

By positivity of $\tau \otimes \sigma$,

$$\begin{aligned}
 (\tau \otimes \sigma)(\gamma_1(u_{ij}) \wedge \gamma_2(u_{ij})) &\geq \sum_{k=1}^n (\tau \otimes \sigma)((\alpha_1(u_{ik}) \wedge \alpha_2(u_{ik})) \otimes (\beta_1(u_{kj}) \wedge \beta_2(u_{kj}))) \\
 &= \sum_{k=1}^n \tau(\alpha_1(u_{ik}) \wedge \alpha_2(u_{ik}))\sigma(\beta_1(u_{kj}) \wedge \beta_2(u_{kj}))
 \end{aligned}$$

Therefore,

$$\begin{aligned}
 &1 - \frac{1}{n} \sum_{i,j=1}^n (\tau \otimes \sigma)(\gamma_1(u_{ij}) \wedge \gamma_2(u_{ij})) \\
 &\leq 1 - \frac{1}{n} \sum_{i,k=1}^n \tau(\alpha_1(u_{ik}) \wedge \alpha_2(u_{ik})) + \frac{1}{n} \sum_{i,k=1}^n \tau(\alpha_1(u_{ik}) \wedge \alpha_2(u_{ik})) \\
 &\quad - \frac{1}{n} \sum_{i,j,k=1}^n \tau(\alpha_1(u_{ik}) \wedge \alpha_2(u_{ik}))\sigma(\beta_1(u_{kj}) \wedge \beta_2(u_{kj})) \\
 &\leq 1 - \frac{1}{n} \sum_{i,k=1}^n \tau(\alpha_1(u_{ik}) \wedge \alpha_2(u_{ik})) + \frac{1}{n} \sum_{i,k=1}^n \tau(\alpha_1(u_{ik}) \wedge \alpha_2(u_{ik}))(1 - \sum_{j=1}^n \sigma(\beta_1(u_{kj}) \wedge \beta_2(u_{kj}))) \\
 &\leq 1 - \frac{1}{n} \sum_{i,k=1}^n \tau(\alpha_1(u_{ik}) \wedge \alpha_2(u_{ik})) + 1 - \frac{1}{n} \sum_{k,j=1}^n \sigma(\beta_1(u_{kj}) \wedge \beta_2(u_{kj})).
 \end{aligned}$$

Since the couplings of (φ_1, φ_2) and (ψ_1, ψ_2) were arbitrary, we get the required inequality. The small thing to note here is the identification for the state as mentioned in the paragraph before (3.8). \square

3.4. Distance via L^1 -norm

Recall the second distance $W_{H,1}$ from Definition 1.3 given by

$$W_{H,1}(\varphi_1, \varphi_2) = \inf_{(M, \tau, \alpha_1, \alpha_2)} \frac{1}{n} \sum_{i,j=1}^n \frac{1}{2} \|\alpha_1(u_{ij}) - \alpha_2(u_{ij})\|_{L^1(M, \tau)},$$

where $L^1(M, \tau)$ is the non-commutative L^1 space (see §2.2.1).

Lemma 3.9. *The infimum in (1.3) is achieved.*

Proof. Let C_1 be the element of the universal free product $A * A$ given by

$$C_1 = \frac{1}{2n} \sum_{i,j=1}^n |\iota_1(u_{ij}) - \iota_2(u_{ij})|,$$

where $|x| = (x^*x)^{1/2}$. If $(M, \tau, \alpha_1, \alpha_2)$ is a tracial coupling and $\varphi \in \mathcal{T}_{\varphi_1, \varphi_2}(A * A)$ is the corresponding state, then

$$\inf_{(M, \tau, \alpha_1, \alpha_2)} \frac{1}{n} \sum_{i,j=1}^n \frac{1}{2} \|\alpha_1(u_{ij}) - \alpha_2(u_{ij})\|_{L^1(M, \tau)} = \varphi(C_1).$$

Since $\varphi \mapsto \varphi(C_1)$ is weak- $*$ continuous and $\mathcal{T}_{\varphi_1, \varphi_2}(A * A)$ is compact, the infimum is achieved. \square

We now record the properties of $W_{H,1}$ analogous to those of $W_{H,\wedge}$.

Proposition 3.10. *Let $\varphi_1, \varphi_2, \psi_1, \psi_2 \in \mathcal{T}(C(S_n^+))$.*

1. $W_{H,1}$ is a metric on $\mathcal{T}(C(S_n^+))$.
2. $W_{H,1}$ is weak- $*$ lower semi-continuous on $\mathcal{T}(C(S_n^+)) \times \mathcal{T}(C(S_n^+))$ and convergence in $W_{H,1}$ implies weak- $*$ convergence.
3. $W_{H,1}(\varphi_1 * \psi_1, \varphi_2 * \psi_2) \leq W_{H,1}(\varphi_1, \varphi_2) + W_{H,1}(\psi_1, \psi_2)$.

Proof. (1) $W_{H,1}(\varphi_1, \varphi_2) \geq 0$ is straightforward. For $\varphi_1 = \varphi_2$, choose the tracial coupling $(M, \tau, \alpha_1, \alpha_2)$ such that $\tau \circ \alpha_1 = \varphi_1 = \varphi_2 = \tau \circ \alpha_2$. So, for each i, j , $\tau(\alpha_1(u_{ij}) - \alpha_2(u_{ij})) = \tau(\alpha_1 - \alpha_2)(u_{ij}) = \tau(\alpha_1(u_{ij}) - \alpha_2(u_{ij}))^*(\alpha_1(u_{ij}) - \alpha_2(u_{ij}))$. Since τ is faithful, we get $\alpha_1(u_{ij}) = \alpha_2(u_{ij})$ for all i, j , so $W_{H,1}(\varphi_1, \varphi_2) = 0$. By Lemma 3.9, the infimum is achieved $W_{H,1}(\varphi_1, \varphi_2) = 0 \implies$ there is a coupling $(M, \tau, \alpha_1, \alpha_2)$ such that $\|\alpha_1(u_{ij}) - \alpha_2(u_{ij})\| = 0 \implies \alpha_1(u_{ij}) = \alpha_2(u_{ij})$ for all i, j and therefore $\alpha_1 = \alpha_2$. Then $\varphi_1 = \tau \circ \alpha_1 = \tau \circ \alpha_2 = \varphi_2$. Symmetry is similar to (3.6). The triangle inequality follows from the triangle inequality for the norm of $L^1(M, \tau)$ and constructing the amalgamated free product as in the case of $W_{H,\wedge}$.

(2) The proof is completely analogous to Lemma 3.7, so we leave the details to the reader.

(3) As in Proposition 3.8, consider tracial couplings $(M, \tau, \alpha_1, \alpha_2)$ of φ_1 and φ_2 and $(N, \sigma, \beta_1, \beta_2)$ of ψ_1 and ψ_2 . Let $\gamma_j = (\alpha_j \otimes \beta_j) \circ \Delta : C(S_n^+) \rightarrow M \otimes N$, so that $(M \otimes N, \tau \otimes \sigma, \gamma_1, \gamma_2)$ is a tracial coupling of $\varphi_1 * \psi_1$ and $\varphi_2 * \psi_2$. Note that $\|x \otimes y\|_{L^1(M \otimes N, \tau \otimes \sigma)} = \|x\|_{L^1(M, \tau)} \|y\|_{L^1(N, \sigma)}$. Thus,

$$\begin{aligned} & \|\gamma_1(u_{ij}) - \gamma_2(u_{ij})\|_{L^1(M \otimes N, \tau \otimes \sigma)} \\ &= \|(\alpha_1 \otimes \beta_1)\Delta(u_{ij}) - (\alpha_2 \otimes \beta_2)\Delta(u_{ij})\|_{L^1(M \otimes N, \tau \otimes \sigma)} \\ &= \left\| \sum_{k=1}^n \alpha_1(u_{ik}) \otimes \beta_1(u_{kj}) - \alpha_2(u_{ik}) \otimes \beta_2(u_{kj}) \right\|_{L^1(M \otimes N, \tau \otimes \sigma)} \\ &= \left\| \sum_{k=1}^n (\alpha_1(u_{ik}) - \alpha_2(u_{ik})) \otimes \beta_1(u_{kj}) + \alpha_2(u_{ik}) \otimes (\beta_1(u_{kj}) - \beta_2(u_{kj})) \right\|_{L^1(M \otimes N, \tau \otimes \sigma)} \\ &\leq \sum_{k=1}^n \|\alpha_1(u_{ik}) - \alpha_2(u_{ik})\|_{L^1(M, \tau)} \|\beta_1(u_{kj})\|_{L^1(N, \sigma)} + \|\alpha_2(u_{ik})\|_{L^1(M, \tau)} \|\beta_1(u_{kj}) - \beta_2(u_{kj})\|_{L^1(N, \sigma)} \\ &= \sum_{k=1}^n \|\alpha_1(u_{ik}) - \alpha_2(u_{ik})\|_{L^1(M, \tau)} \sigma(\beta_1(u_{kj})) + \tau(\alpha_2(u_{ik})) \|\beta_1(u_{kj}) - \beta_2(u_{kj})\|_{L^1(N, \sigma)}. \end{aligned}$$

Therefore,

$$\begin{aligned} & \frac{1}{2n} \sum_{i,j=1}^n \|\gamma_1(u_{ij}) - \gamma_2(u_{ij})\|_{L^1(M \otimes N, \tau \otimes \sigma)} \\ &\leq \frac{1}{2n} \sum_{i,j,k=1}^n \|\alpha_1(u_{ik}) - \alpha_2(u_{ik})\|_{L^1(M, \tau)} \sigma(\beta_1(u_{kj})) + \frac{1}{2n} \sum_{i,j,k=1}^n \tau(\alpha_2(u_{ik})) \|\beta_1(u_{kj}) - \beta_2(u_{kj})\|_{L^1(N, \sigma)} \\ &= \frac{1}{2n} \sum_{i,k=1}^n \|\alpha_1(u_{ik}) - \alpha_2(u_{ik})\|_{L^1(M, \tau)} \left[\sum_{j=1}^n \sigma(\beta_1(u_{kj})) \right] + \frac{1}{2n} \sum_{j,k=1}^n \left[\sum_{i=1}^n \tau(\alpha_2(u_{ik})) \right] \|\beta_1(u_{kj}) - \beta_2(u_{kj})\|_{L^1(N, \sigma)} \\ &= \frac{1}{2n} \sum_{i,k=1}^n \|\alpha_1(u_{ik}) - \alpha_2(u_{ik})\|_{L^1(M, \tau)} + \frac{1}{2n} \sum_{j,k=1}^n \|\beta_1(u_{kj}) - \beta_2(u_{kj})\|_{L^1(N, \sigma)}. \end{aligned}$$

Since the couplings were arbitrary, we obtain $W_{H,1}(\varphi_1 * \psi_1, \varphi_2 * \psi_2) \leq W_{H,1}(\varphi_1, \varphi_2) + W_{H,1}(\psi_1, \psi_2)$ as desired. \square

4. Distance via tensor couplings

4.1. Tracial tensor couplings

Definition 4.1. Let A be a C^* -algebra, and let $\varphi_1, \varphi_2 \in \mathcal{T}(A)$. A *tracial tensor coupling* of (φ_1, φ_2) is a trace $\varphi \in \mathcal{T}(A \otimes A)$ such that $\varphi(a \otimes 1) = \varphi_1(a)$ and $\varphi(1 \otimes a) = \varphi_2(a)$ for $a \in A$. We denote the set of tracial tensor couplings by $\Pi_{\text{tr}, \otimes}(\varphi_1, \varphi_2)$.

Definition 4.2. A *cost operator* is an element $C \in (A \otimes A)_+$. We say that C is *symmetric* if $\iota_{1,2}(C) = \iota_{1,2}(C)$. We say that C satisfies the *triangle inequality* if $\iota_{1,3}(C) \leq \iota_{1,2}(C) + \iota_{2,3}(C)$, where we employ tensor-leg notation.

Definition 4.3. For a cost operator C and $\varphi_1, \varphi_2 \in \mathcal{T}(A)$, define

$$W_C(\varphi_1, \varphi_2) = \inf_{\varphi \in \Pi_{\text{tr}, \otimes}(\varphi_1, \varphi_2)} \varphi(C).$$

Remark 4.4. The infimum is always achieved because $\Pi_{\text{tr}, \otimes}$ is a weak- $*$ compact subset of $\mathcal{T}(A)$ and $\varphi \mapsto \varphi(C)$ is weak- $*$ continuous.

Proposition 4.5. Let $C \in (A \otimes A)_+$ be a cost operator and $\varphi_1, \varphi_2, \varphi_3 \in \mathcal{T}(A)$.

1. $W_C(\varphi_1, \varphi_2) \geq 0$.
2. If C is symmetric, then $W_C(\varphi_1, \varphi_2) = W_C(\varphi_2, \varphi_1)$.
3. If C satisfies the triangle inequality, then $W_C(\varphi_1, \varphi_3) \leq W_C(\varphi_1, \varphi_2) + W_C(\varphi_2, \varphi_3)$.

The main challenge in Proposition 4.5 is to prove the triangle inequality. In order to do this, we need to consider $\tau_1 \in \Pi_{\text{tr}, \otimes}(\varphi_1, \varphi_2)$ and $\tau_2 \in \Pi_{\text{tr}, \otimes}(\varphi_2, \varphi_3)$ and amalgamate them to obtain an element of $\Pi_{\text{tr}, \otimes}(\varphi_1, \varphi_3)$.

The key property that makes this proposition work is that the trace space is a Choquet simplex. Recall that for a weak-* closed convex subset X in the dual of some Banach space, we denote by $\partial_e X$ the set of extreme points. Choquet’s theorem (see e.g. [48, §I.4]) says for each $x \in X$, there exists a probability measure μ on $\partial_e X$ such that $x = \int_{\partial_e X} y d\mu(y)$; in this case, we say that x is the barycenter of μ . If the measure μ is always unique, then X is called a Choquet simplex (see e.g. [48, §II.3]). For a C^* -algebra A , both $\mathcal{S}(A)$ and $\mathcal{T}(A)$ are weak-* closed convex subsets of the dual of A . However, $\mathcal{S}(A)$ is never a Choquet simplex unless A is commutative. On the other hand, $\mathcal{T}(A)$ is always a Choquet simplex. A simple proof of this fact is given in [29].

Lemma 4.6. Let A be a unital C^* -algebra. For every $\varphi \in \mathcal{T}(A)$, there exists a unique $\mu \in \mathcal{P}(\partial_e \mathcal{T}(A))$ such that

$$\varphi(a) = \int_{\partial_e \mathcal{T}(A)} \tau(a) d\mu(\tau).$$

In order to study traces on the tensor product, we also need to use the following fact.

Lemma 4.7. Let A and B be C^* -algebras. Let $\tau \in \mathcal{T}(A \otimes B)$. Then τ is extreme if and only if $\tau = \tau_1 \otimes \tau_2$ where $\tau_1 \in \partial_e \mathcal{T}(A)$ and $\tau_2 \in \partial_e \mathcal{T}(B)$. Moreover, this relationship defines a weak-* homeomorphism $\partial_e \mathcal{T}(A \otimes B) \cong \partial_e \mathcal{T}(A) \times \partial_e \mathcal{T}(B)$.

Lemma 4.8. Let A be a unital C^* -algebra. Let $\varphi_1, \varphi_2, \varphi_3 \in \mathcal{T}(A)$. Fix couplings $\rho_1 \in \Pi_{\text{tr}}(\varphi_1, \varphi_2)$ and $\rho_2 \in \Pi_{\text{tr}}(\varphi_2, \varphi_3)$. Then there exists a unique $\sigma \in \mathcal{T}(A \otimes A \otimes A)$ such that

$$\rho \circ \iota_{1,2} = \rho_1, \quad \rho \circ \iota_{2,3} = \rho_2,$$

where $\iota_{i,j}$ is the natural tensor leg inclusion.

Proof. Recall that $\partial_e \mathcal{T}(A \otimes A) \cong \partial_e \mathcal{T}(A) \otimes \partial_e \mathcal{T}(A)$, and let $\pi_1, \pi_2 : \partial_e \mathcal{T}(A \otimes A) \rightarrow \partial_e \mathcal{T}(A)$ be the two canonical coordinate projections.

For $j = 1, 2, 3$ let $\mu_j \in \mathcal{P}(\partial_e \mathcal{T}(A))$ be the unique measure such that $\varphi_j = \int_{\partial_e \mathcal{T}(A)} \tau d\mu_j(\tau)$. Similarly, for $j = 1, 2$, let $\nu_j \in \mathcal{P}(\partial_e \mathcal{T}(A \otimes A))$ be the unique measure such that ρ_j is the barycenter of ν_j . Note that $(\pi_2)_* \nu_1$ gives a measure on $\partial_e \mathcal{T}(A)$ whose barycenter is φ_2 , and therefore since $\mathcal{T}(A)$ is a Choquet simplex, we have $(\pi_2)_* \nu_1 = \mu_2$. Similarly, $(\pi_1)_* \nu_2 = \mu_2$. Thus, ν_1 and ν_2 yield are two probability measures on $\partial_e \mathcal{T}(A) \times \partial_e \mathcal{T}(A)$ such that the second marginal of ν_1 agrees with the first marginal of ν_2 . By standard results on conditional independence in probability theory, there exists a probability measure λ on $\partial_e \mathcal{T}(A) \times \partial_e \mathcal{T}(A) \times \partial_e \mathcal{T}(A) \cong \partial_e \mathcal{T}(A \otimes A \otimes A)$ whose projection on the first two coordinates is ν_1 and whose projection on the second two coordinates is ν_2 . Then let

$$\rho = \int_{\partial_e \mathcal{T}(A)^{\times 3}} \tau_1 \otimes \tau_2 \otimes \tau_3 d\lambda(\tau_1, \tau_2, \tau_3).$$

Then

$$\rho \circ \iota_{1,2} = \int_{\partial_e \mathcal{T}(A)^{\times 3}} \tau_1 \otimes \tau_2 d\lambda(\tau_1, \tau_2, \tau_3) = \int_{\partial_e \mathcal{T}(A)^{\times 2}} \tau_1 \otimes \tau_2 d\nu_1(\tau_1, \tau_2) = \rho_1,$$

and similarly $\rho \circ \iota_{2,3} = \rho_2$. \square

Proof of Proposition 4.5. (1) follows because for any $\varphi \in \Pi_{\text{tr}, \otimes}(\varphi_1, \varphi_2)$, we have $\varphi(C) \geq 0$ by positivity of C .

For (2), note that if $\varphi \in \Pi_{\text{tr}, \otimes}$, then so is φ composed with the tensor flip.

It remains to prove (3) that W_C satisfies the triangle inequality provided that $C_{1,3} \leq C_{1,2} + C_{2,3}$. Let $\varphi_1, \varphi_2, \varphi_3 \in \mathcal{T}(A)$. Fix $\psi_1 \in \Pi_{\text{tr}, \otimes}(\varphi_1, \varphi_2)$ and $\psi_2 \in \Pi_{\text{tr}, \otimes}(\varphi_2, \varphi_3)$. By Lemma 4.8, there exists some $\rho \in \mathcal{T}(A \otimes A \otimes A)$ such that

$$\rho \circ \iota_{1,2} = \psi_1, \quad \rho \circ \iota_{2,3} = \psi_2.$$

Note that $\psi := \rho \circ \iota_{1,3} \in \Pi_{\text{tr}, \otimes}(\varphi_1, \varphi_3)$. Therefore,

$$\begin{aligned} \rho(\iota_{1,3}(C)) &\leq \rho(\iota_{1,2}(C)) + \rho(\iota_{2,3}(C)) \\ \psi(C) &\leq \psi_1(C) + \psi_2(C). \end{aligned}$$

Hence, taking the infimum over $\psi_1 \in \Pi_{\text{tr}, \otimes}(\varphi_1, \varphi_2)$ and $\psi_2 \in \Pi_{\text{tr}, \otimes}(\varphi_2, \varphi_3)$, we have

$$W_C(\varphi_1, \varphi_3) \leq W_C(\varphi_1, \varphi_2) + W_C(\varphi_2, \varphi_3),$$

as desired. \square

Remark 4.9. In fact, both properties (1) and (2) also work if we use general states rather than traces, and consider the set $\Pi_{\otimes}(\varphi_1, \varphi_2)$ of states φ on $A \otimes A$ with $\varphi(a \otimes 1) = \varphi_1(a)$ and $\varphi(1 \otimes a) = \varphi_2(a)$. Moreover, the triangle inequality (3) will hold whenever the middle state φ_2 admits a unique decomposition in terms of extreme points, e.g. if φ_2 is itself extreme. Compare [49, Theorem 1] where certain quantum Wasserstein distances are shown to satisfy the triangle inequality when the middle state is pure.

We finally remark that by decomposing a trace in terms of extreme traces, the tracial couplings can be reduced in some sense to classical couplings. This lemma follows from the same reasoning used in the proof of Lemma 4.8.

Lemma 4.10. Let C be a cost operator, and for $\tau_1, \tau_2 \in \partial_e \mathcal{T}(A)$, define

$$c(\tau_1, \tau_2) := (\tau_1 \otimes \tau_2)(C).$$

Let $\varphi_1, \varphi_2 \in \mathcal{T}(A)$. Let μ_1 and μ_2 be the unique probability measures on $\partial_e \mathcal{T}(A)$ with barycenters φ_1 and φ_2 . Then any $\varphi \in \Pi_{\text{tr}, \otimes}(\varphi_1, \varphi_2)$ is the barycenter of a unique probability measure μ on $\partial_e \mathcal{T}(A) \times \partial_e \mathcal{T}(A)$ which has marginals μ_1 and μ_2 , and we have

$$\varphi(C) = \int_{\partial_e \mathcal{T}(A) \times \partial_e \mathcal{T}(A)} c(\tau_1, \tau_2) d\mu(\tau_1, \tau_2).$$

4.2. Tensor Hamming distance for $C(S_n^+)$

Definition 4.11. The Hamming cost operator is the element $C_H \in C(S_n^+) \otimes C(S_n^+)$ given by

$$C_H = 1 - \frac{1}{n} \sum_{i,j=1}^n u_{ij} \otimes u_{ij}.$$

Remark 4.12. The Hamming cost for quantum permutations can be expressed in another way as follows. Let U be the canonical $n \times n$ unitary matrix in $M_n(\mathbb{C}) \otimes C(S_n^+)$. Then

$$\frac{1}{n} \sum_{i,j=1}^n \rho(u_{ij} \otimes u_{ij}) = [\text{tr}_n \otimes \rho](\iota_{1,2}(U^*)\iota_{1,3}(U)),$$

where the element on the right-hand side is in $M_n(\mathbb{C}) \otimes C(S_n^+) \otimes C(S_n^+)$. Therefore,

$$C_H = 1 - [\text{tr}_n \otimes \rho](\iota_{1,2}(U^*)\iota_{1,3}(U)).$$

Since U is unitary, it is clear that $[\text{tr}_n \otimes \rho](\iota_{1,2}(U^*)\iota_{1,3}(U))$ has norm at most 1. We also have that C_H is self-adjoint since the u_{ij} 's are projections. Therefore, $C_H \geq 0$, so it is a valid cost operator.

Lemma 4.13. The Hamming cost C_H is symmetric and satisfies the triangle inequality $\iota_{1,3}(C_H) \leq \iota_{1,2}(C_H) + \iota_{2,3}(C_H)$. Consequently, $W_{H, \otimes}$ is nonnegative and symmetric, and it satisfies the triangle inequality.

Proof. We already verified $C_H \geq 0$, and it is immediate that C_H is symmetric under the tensor flip. It remains to show that C_H satisfies the triangle inequality, that is,

$$1 - \frac{1}{n} \sum_{i,j=1}^n u_{ij} \otimes 1 \otimes u_{ij} \leq \left(1 - \frac{1}{n} \sum_{i,j=1}^n u_{ij} \otimes u_{ij} \otimes 1 \right) + \left(1 - \frac{1}{n} \sum_{i,j=1}^n 1 \otimes u_{ij} \otimes u_{ij} \right).$$

It suffices to prove for each i that

$$1 - \sum_{j=1}^n u_{ij} \otimes 1 \otimes u_{ij} \leq \left(1 - \sum_{j=1}^n u_{ij} \otimes u_{ij} \otimes 1\right) + \left(1 - \sum_{j=1}^n 1 \otimes u_{ij} \otimes u_{ij}\right),$$

since of course this operator inequality can be summed over i then divided by n . Therefore, fix i , and write

$$p_{1,2} = \sum_{j=1}^n u_{ij} \otimes u_{ij} \otimes 1$$

$$p_{2,3} = \sum_{j=1}^n 1 \otimes u_{ij} \otimes u_{ij}$$

$$p_{1,3} = \sum_{j=1}^n u_{ij} \otimes 1 \otimes u_{ij}.$$

Then we want to prove that $1 - p_{1,3} \leq (1 - p_{1,2}) + (1 - p_{2,3})$. Recall that $(u_{ij})_{j=1}^n$ is a family of projections that sum up to 1, and in particular, they commute. Similarly, $(u_{ji})_{j=1}^n$ is a family of projections that add up to 1. Therefore, for a fixed value of i , all the operators in the equations above commute with each other. Note that $p_{1,2}$ is a projection since $u_{ij} \otimes u_{ij}$ for $j = 1, \dots, n$ are disjoint projections. Similarly, the other four operators are projections. Since we have commuting projections,

$$(1 - p_{1,2}) + (1 - p_{2,3}) \geq (1 - p_{1,2}) \vee (1 - p_{2,3}) = 1 - p_{1,2}p_{2,3}.$$

Using again the fact $\{u_{ij}\}_{j=1}^n$ are orthogonal

$$\begin{aligned} p_{1,2}p_{2,3} &= \left(\sum_{j=1}^n u_{ij} \otimes u_{ij} \otimes 1\right) \left(\sum_{j'=1}^n 1 \otimes u_{ij'} \otimes u_{ij'}\right) \\ &= \sum_{j=1}^n u_{ij} \otimes u_{ij} \otimes u_{ij} \\ &\leq \sum_{j=1}^n u_{ij} \otimes 1 \otimes u_{ij} \\ &= p_{1,3}. \end{aligned}$$

Thus, $1 - p_{1,3} \leq 1 - p_{1,2}p_{2,3} \leq (1 - p_{1,2}) + (1 - p_{2,3})$, as desired. \square

4.3. Bounds in terms of self-cost and equality cases

Lemma 4.10 allows us to rephrase the tensor Hamming distance for the quantum permutation group in terms of a classical optimization problem.

Lemma 4.14. *Let $\varphi_1, \varphi_2 \in \mathcal{T}(C(S_n^+))$. Let μ_j be the unique probability measure on $\partial_e \mathcal{T}(C(S_n^+))$ with barycenter φ_j . Let*

$$c_H : \partial_e \mathcal{T}(C(S_n^+)) \times \partial_e \mathcal{T}(C(S_n^+)) \rightarrow [0, \infty)$$

be given by

$$c_H(\tau_1, \tau_2) = (\tau_1 \otimes \tau_2)(C_H) = 1 - \frac{1}{n} \sum_{i,j=1}^n \tau_1(u_{ij})\tau_2(u_{ij}).$$

Then

$$W_{H,\otimes}(\varphi_1, \varphi_2) = \inf_{\mu} \int_{\partial_e \mathcal{T}(C(S_n^+)) \times \partial_e \mathcal{T}(C(S_n^+))} c_H(\tau_1, \tau_2) d\mu(\tau_1, \tau_2),$$

where μ ranges over probability measures on $\partial_e \mathcal{T}(C(S_n^+)) \times \partial_e \mathcal{T}(C(S_n^+))$ with marginals μ_1 and μ_2 respectively.

Proof. The evaluation of $(\tau_1 \otimes \tau_2)(C_H) = 1 - \frac{1}{n} \sum_{i,j=1}^n \tau_1(u_{ij})\tau_2(u_{ij})$ follows by direct computation. By Lemma 4.10, the tensor couplings φ of φ_1 and φ_2 are in bijection with probability measures μ on $\partial_e \mathcal{T}(C(S_n^+)) \times \partial_e \mathcal{T}(C(S_n^+))$, and the cost of each μ is given by the expression asserted in the lemma. \square

In particular, this allows us to evaluate the self-distance $W_{H,\otimes}(\varphi, \varphi)$ as follows. Note the parallel with the self-distance estimate for a certain quantum Wasserstein distance [50, Theorem 1].

Lemma 4.15. *Let $\varphi_0 \in \mathcal{T}(C(S_n^+))$ and let μ_0 be the corresponding probability measure on $\partial_e \mathcal{T}(C(S_n^+))$. Then*

$$W_{H,\otimes}(\varphi_0, \varphi_0) = 1 - \frac{1}{n} \sum_{i,j=1}^n \int_{\partial_e \mathcal{T}(C(S_n^+))} \tau(u_{ij})^2 d\mu_0(\tau). \tag{4.1}$$

Now let $\varphi_1, \varphi_2 \in \mathcal{T}(C(S_n^+))$ and let μ_1 and μ_2 be the corresponding probability measures on $\partial_e \mathcal{T}(C(S_n^+))$. Then

$$\begin{aligned} W_{H,\otimes}(\varphi_1, \varphi_2) - \frac{1}{2}W_{H,\otimes}(\varphi_1, \varphi_1) - \frac{1}{2}W_{H,\otimes}(\varphi_2, \varphi_2) \\ = \inf_{\mu} \int_{\partial_e \mathcal{T}(C(S_n^+)) \times \partial_e \mathcal{T}(C(S_n^+))} \frac{1}{2n} \sum_{i,j=1}^n (\tau_1(u_{ij}) - \tau_2(u_{ij}))^2 d\mu(\tau_1, \tau_2). \end{aligned} \tag{4.2}$$

Proof. For ease of notation, write $\Omega = \partial_e \mathcal{T}(C(S_n^+))$. Consider $\varphi_1, \varphi_2 \in \mathcal{T}(C(S_n^+))$ and corresponding measures μ_1 and μ_2 , and let μ be a measure on $\partial_e \mathcal{T}(C(S_n^+)) \times \partial_e \mathcal{T}(C(S_n^+))$ with marginals μ_1 and μ_2 . Then we have by the arithmetic-geometric mean inequality,

$$\begin{aligned} \int_{\Omega \times \Omega} \left[1 - \frac{1}{n} \sum_{i,j=1}^n \tau_1(u_{ij})\tau_2(u_{ij}) \right] d\mu(\tau_1, \tau_2) \\ = 1 - \frac{1}{n} \sum_{i,j=1}^n \int_{\Omega \times \Omega} \tau_1(u_{ij})\tau_2(u_{ij}) d\mu(\tau_1, \tau_2) \\ \geq 1 - \frac{1}{n} \sum_{i,j=1}^n \int_{\Omega \times \Omega} \left[\frac{1}{2}\tau_1(u_{ij})^2 + \frac{1}{2}\tau_2(u_{ij})^2 \right] d\mu(\tau_1, \tau_2) \\ = \frac{1}{2} \int_{\Omega} \left[1 - \frac{1}{n} \sum_{i,j=1}^n \tau_1(u_{ij})^2 \right] d\mu_1(\tau_1) + \frac{1}{2} \int_{\Omega} \left[1 - \frac{1}{n} \sum_{i,j=1}^n \tau_2(u_{ij})^2 \right] d\mu_2(\tau_2). \end{aligned}$$

In particular, in the case when $\mu_1 = \mu_2 = \mu_0$, we obtain

$$\int_{\Omega \times \Omega} \left[1 - \frac{1}{n} \sum_{i,j=1}^n \tau_1(u_{ij})\tau_2(u_{ij}) \right] d\mu(\tau_1, \tau_2) \geq \int_{\Omega} \left[1 - \frac{1}{n} \sum_{i,j=1}^n \tau(u_{ij})^2 \right] d\mu_0(\tau),$$

and thus, since μ was arbitrary,

$$W_{H,\otimes}(\varphi_0, \varphi_0) \geq \int_{\partial_e \mathcal{T}(C(S_n^+))} \left[1 - \frac{1}{n} \sum_{i,j=1}^n \tau(u_{ij})^2 \right] d\mu_0(\tau).$$

On the other hand, equality is achieved when we take μ to be a copy of μ_0 supported on the diagonal of $\Omega \times \Omega$. Therefore, we have proved (4.1).

We then observe that

$$\begin{aligned} \int_{\Omega \times \Omega} \left[1 - \frac{1}{n} \sum_{i,j=1}^n \tau_1(u_{ij})\tau_2(u_{ij}) \right] d\mu(\tau_1, \tau_2) - \frac{1}{2}W_{H,\otimes}(\varphi_1, \varphi_1) - \frac{1}{2}W_{H,\otimes}(\varphi_2, \varphi_2) \\ = \int_{\Omega \times \Omega} \left[1 - \frac{1}{n} \sum_{i,j=1}^n \tau_1(u_{ij})\tau_2(u_{ij}) \right] d\mu(\tau_1, \tau_2) \end{aligned}$$

$$\begin{aligned}
 & -\frac{1}{2} \int_{\Omega} \left[1 - \frac{1}{2} \sum_{i,j=1}^n \tau_1(u_{ij})^2 \right] d\mu_1(\tau_1) - \frac{1}{2} \int_{\Omega} \left[1 - \frac{1}{n} \sum_{i,j=1}^n \tau_2(u_{ij})^2 \right] d\mu_2(\tau_2) \\
 & = \int_{\Omega \times \Omega} \frac{1}{2n} \sum_{i,j=1}^n (\tau_1(u_{ij}) - \tau_2(u_{ij}))^2 d\mu(\tau_1, \tau_2).
 \end{aligned}$$

By taking the infimum over μ with marginals μ_1 and μ_2 on both sides of the equation, we obtain (4.2). \square

As a consequence of this formula, we next observe that the self-distance vanishes if and only if φ arises from a classical probability measure on S_n . Thus, “most” tracial states on $C(S_n^+)$ have nonzero self-distance.

Lemma 4.16. *Let φ be a tracial state on $C(S_n^+)$. Then $W_{H,\otimes}(\varphi, \varphi) = 0$ if and only if φ has the form $\varphi(a) = \int_{S_n} \theta(a) dv$ for some $v \in P(S_n)$, where $\theta : C(S_n^+) \rightarrow C(S_n)$ is the canonical quotient map.*

Proof. Suppose that $W_{H,\otimes}(\varphi, \varphi) = 0$. Let μ be the measure on $\partial_e \mathcal{T}(C(S_n^+))$ corresponding to φ . Then

$$\begin{aligned}
 \int_{\partial_e \mathcal{T}(C(S_n^+))} \frac{1}{n} \sum_{i,j=1}^n [\tau(u_{ij}) - \tau(u_{ij})^2] d\mu(\tau) &= \int_{\partial_e \mathcal{T}(C(S_n^+))} \left[\frac{1}{n} \sum_{i=1}^n \tau(1) - \frac{1}{n} \sum_{i,j=1}^n \tau(u_{i,j})^2 \right] d\mu(\tau) \\
 &= 1 - \int_{\partial_e \mathcal{T}(C(S_n^+))} \left[\frac{1}{n} \sum_{i,j=1}^n \tau(u_{ij})^2 \right] d\mu(\tau) \\
 &= W_{H,\otimes}(\varphi, \varphi).
 \end{aligned}$$

Since $0 \leq \tau(u_{ij}) - \tau(u_{ij})^2 \leq 1$, we conclude that for μ -almost every τ , for all i, j , we have $\tau(u_{ij}) = \tau(u_{ij})^2$, and so $\tau(u_{ij}) \in \{0, 1\}$. For any τ that satisfies this condition, the projection $\pi_{\tau}(u_{ij})$ is either 0 or 1 by faithfulness of the trace on $\pi_{\tau}(A)''$. Since $\pi_{\tau}(A)''$ is generated by $\{\pi_{\tau}(u_{ij})\}_{i,j=1}^n$, this forces $\pi_{\tau}(A)''$ to be one-dimensional, and so $\pi_{\tau}(a) = \tau(a)1$ for all a . Hence, τ is $*$ -homomorphism $C(S_n^+) \rightarrow \mathbb{C}$ and so it vanishes on $\ker(\theta)$. Since this holds for μ -almost every τ , we see that φ vanishes on $\ker(\theta)$. Hence, there exists a tracial state $\bar{\varphi}$ on $C(S_n)$ such that $\varphi = \bar{\varphi} \circ \theta$; indeed, one can verify readily that $\bar{\varphi}$ is well-defined as a linear map since φ vanishes on the kernel and then check that $\bar{\varphi}$ is unital, positive, and tracial. But every tracial state on $C(S_n)$ is given by a probability measure on S_n .

For the converse direction, we defer to Proposition 5.3 below which shows that for tracial states that arise from probability measures on S_n , our $W_{H,\otimes}$ agrees with the classical L^1 Wasserstein distance, so in particular the self-distance is zero. Alternatively, one can argue that the Dirac delta masses on $C(S_n)$ produce extreme states on $C(S_n^+)$, so the decomposition of v in terms of such Dirac masses gives a direct description of boundary measure μ to use in (4.1). \square

The preceding lemma allows us to express $W_{H,\otimes}$ in terms of measures on the Birkhoff polytope

$$B_n = \{A \in M_n(\mathbb{R}) : A_{ij} \in [0, 1], \sum_i A_{ij} = 1, \sum_j A_{ij} = 1\}$$

of bistochastic matrices. We remark that if φ is a state on $C(S_n^+)$, then the matrix

$$\omega(\varphi) := [\varphi(u_{ij})]_{i,j=1}^n \text{ is in } B_n$$

because $\sum_i u_{ij} = 1$ and $\sum_j u_{ij} = 1$.

Lemma 4.17. *Let $\varphi_1, \varphi_2 \in \mathcal{T}(C(S_n^+))$, and let μ_j be the probability measure on $\partial_e \mathcal{T}(C(S_n^+))$ with barycenter φ_j . Let $\omega_* \mu_j \in \mathcal{P}(B_n)$ be the pushforward and let W_2 be the Wasserstein distance on $\mathcal{P}(B_n)$ with respect to the normalized Hilbert-Schmidt metric $\|A - B\|_2 = \text{tr}_n[(A - B)^2]^{1/2}$. Then we have*

$$W_{H,\otimes}(\varphi_1, \varphi_2) - \frac{1}{2} W_{H,\otimes}(\varphi_1, \varphi_1) - \frac{1}{2} W_{H,\otimes}(\varphi_2, \varphi_2) = \frac{1}{2} W_2(\omega_* \mu_1, \omega_* \mu_2)^2.$$

In particular, the left-hand side vanishes if and only if $\omega_* \mu_1 = \omega_* \mu_2$.

Proof. First, let φ be a trace that achieves the optimum in the definition of $W_{H,\otimes}(\varphi_1, \varphi_2)$, and let μ be the corresponding probability measure on $\partial_e \mathcal{T}(C(S_n^+)) \times \partial_e \mathcal{T}(C(S_n^+))$. Then $\nu = (\omega \times \omega)_* \mu$ defines a classical coupling of $\nu_1 = \omega_* \mu_1$ and $\nu_2 = \omega_* \mu_2$. It follows from (4.2) and the definition of ν_1, ν_2 , that

$$\begin{aligned}
 W_{H,\otimes}(\varphi_1, \varphi_2) &- \frac{1}{2}W_{H,\otimes}(\varphi_1, \varphi_1) - \frac{1}{2}W_{H,\otimes}(\varphi_2, \varphi_2) \\
 &= \int_{\partial_e \mathcal{T}(C(S_n^+)) \times \partial_e \mathcal{T}(C(S_n^+))} \frac{1}{2n} \sum_{i,j=1}^n (\tau_1(u_{ij}) - \tau_2(u_{ij}))^2 d\mu(\tau_1, \tau_2) \\
 &= \int_{B_n \times B_n} \frac{1}{2n} \sum_{i,j=1}^n (A_{ij} - A'_{ij})^2 d\nu(A, A') \\
 &= \int_{B_n \times B_n} \frac{1}{2} \|A - A'\|_2^2 d\nu(A, A') \\
 &\geq \frac{1}{2} W_2(\nu_1, \nu_2)^2.
 \end{aligned}$$

To prove the opposite inequality, let $\nu \in \mathcal{P}(B_n \times B_n)$ be an optimal coupling of ν_1 and ν_2 for W_2 . To prove the lemma, it suffices to show that $\nu = (\omega \times \omega)_* \mu$ for some coupling μ of (μ_1, μ_2) on $\partial_e \mathcal{T}(C(S_n^+)) \times \partial_e \mathcal{T}(C(S_n^+))$, because then the computations that we did above can be applied to this μ .

For $j = 1, 2$, define a measure $\rho_j \in \mathcal{P}(\partial_e \mathcal{T}(C(S_n^+)) \times B_n)$ by $\rho_j = (\text{id}, \omega)_* \mu_j$. By construction the marginals of ρ_j are μ_j and ν_j . We also have a measure ν on $B_n \times B_n$ whose marginals are ν_1 and ν_2 . By taking conditionally independent joins or fibered products of the probability spaces, there exists a probability measure ρ on $\partial_e \mathcal{T}(C(S_n^+)) \times B_n \times \partial_e \mathcal{T}(C(S_n^+)) \times B_n$ such that

- The marginal of ρ on the first two coordinates is ρ_1 .
- The marginal of ρ on the second and fourth coordinates (the copies of B_n) is ν .
- The marginal of ρ on the last two coordinates is ρ_2 .

Let $\mu \in \mathcal{P}(\partial_e \mathcal{T}(C(S_n^+)) \times \partial_e \mathcal{T}(C(S_n^+)))$ be the marginal of ρ on the first and third coordinates. Thus, the marginals of μ are the marginals of ρ on the first and third coordinates respectively, which are μ_1 and μ_2 . So it remains to show that $(\omega \times \omega)_* \mu = \nu$. For this, note that ρ_1 and ρ_2 are supported on $\{(\tau, A) \in \partial_e \mathcal{T}(C(S_n^+)) \times B_n : \omega(\tau) = A\}$. Hence, ρ is supported on

$$\{(\tau, A, \tau', A') : \omega(\tau) = A, \omega(\tau') = A'\}.$$

Therefore, for each Borel set $E \subseteq B_n \times B_n$, we have

$$\begin{aligned}
 \nu(E) &= \rho(\{(\tau, A, \tau', A') : (A, A') \in E\}) \\
 &= \rho(\{(\tau, A, \tau', A') : \omega(\tau) = A, \omega(\tau') = A', (A, A') \in E\}) \\
 &= \rho(\{(\tau, A, \tau', A') : \omega(\tau) = A, \omega(\tau') = A', (\tau, \tau') \in (\omega \times \omega)^{-1}(E)\}) \\
 &= \rho(\{(\tau, A, \tau', A') : (\tau, \tau') \in (\omega \times \omega)^{-1}(E)\}) \\
 &= \mu((\omega \times \omega)^{-1}(E)).
 \end{aligned}$$

Hence, $\nu = (\omega \times \omega)_* \mu$ as desired. \square

4.4. Behavior under convolution

In this section, we show the subadditivity property for the distance with respect to convolution. This is the dual version or quantum analog of the multiplication operation being 1-Lipschitz in each argument for a metric group.

We define the following $*$ -homomorphism using the Sweedler notation

$$\begin{aligned}
 \tilde{\Delta} : A \otimes A &\rightarrow A \otimes A \otimes A \otimes A \text{ defined as} \\
 \tilde{\Delta}(a \otimes 1) &= \sum_{(a)} a_{(1)} \otimes 1 \otimes a_{(2)} \otimes 1 = \iota_{1,3}(\Delta(a) \otimes 1) \\
 \tilde{\Delta}(1 \otimes a) &= \sum_{(a)} 1 \otimes a_{(1)} \otimes 1 \otimes a_{(2)} = \iota_{2,4}(1 \otimes \Delta(a))
 \end{aligned}$$

where $\iota_{i,j} : A \otimes A \otimes A \otimes A \rightarrow A \otimes A \otimes A \otimes A$ are embeddings to the (i, j) -th legs.

Lemma 4.18. For the Hamming cost $C_H \in (A \otimes A)_+$, we have

$$\tilde{\Delta}(C_H) \leq \iota_{1,2}(C_H) + \iota_{3,4}(C_H)$$

Proof.

$$\begin{aligned} \tilde{\Delta}(C_H) &= \tilde{\Delta} \left(1 - \frac{1}{n} \sum_{i,j=1}^n u_{ij} \otimes u_{ij} \right) \\ &= 1 \otimes 1 \otimes 1 \otimes 1 - \frac{1}{n} \sum_{i,j,k,l=1}^n u_{ik} \otimes u_{il} \otimes u_{kj} \otimes u_{lj} \\ &\leq 1 \otimes 1 \otimes 1 \otimes 1 - \frac{1}{n} \sum_{i,j,k=1}^n u_{ik} \otimes u_{ik} \otimes u_{kj} \otimes u_{kj} \\ &= 1 \otimes 1 \otimes 1 \otimes 1 - \frac{1}{n} \sum_{i,k=1}^n u_{ik} \otimes u_{ik} \otimes 1 \otimes 1 + \frac{1}{n} \sum_{i,k=1}^n u_{ik} \otimes u_{ik} \otimes 1 \otimes 1 \\ &\quad - \frac{1}{n} \sum_{i,j,k=1}^n u_{ik} \otimes u_{ik} \otimes u_{kj} \otimes u_{kj} \\ &= \iota_{1,2}(C_H) + \frac{1}{n} \sum_{i,k=1}^n u_{ik} \otimes u_{ik} \otimes 1 \otimes 1 \left(1 - \sum_{j=1}^n 1 \otimes 1 \otimes u_{kj} \otimes u_{kj} \right) \\ &\leq \iota_{1,2}(C_H) + \iota_{3,4}(C_H). \quad \square \end{aligned}$$

Proposition 4.19. Let $\phi_1, \phi_2, \psi_1, \psi_2 \in \mathcal{T}(A)$. Then

$$W_{H,\otimes}(\phi_1 * \psi_1, \phi_2 * \psi_2) \leq W_{H,\otimes}(\phi_1, \phi_2) + W_{H,\otimes}(\psi_1, \psi_2).$$

Proof. Let ϕ, ψ be the tracial tensor couplings for ϕ_1, ϕ_2 and ψ_1, ψ_2 respectively. We define $\phi \tilde{*} \psi = (\phi \otimes \psi) \circ \tilde{\Delta} \in \mathcal{T}(A \otimes A)$ and show that it is a tracial tensor coupling for the convolutions $\phi_1 * \psi_1$ and $\phi_2 * \psi_2$. For any $a \in A$, we have

$$\begin{aligned} \phi \tilde{*} \psi(a \otimes 1) &= (\phi \otimes \psi) \tilde{\Delta}(a \otimes 1) \\ &= (\phi \otimes \psi) \left(\sum_{(a)} a_{(1)} \otimes 1 \otimes a_{(2)} \otimes 1 \right) \\ &= \sum_{(a)} \phi(a_{(1)} \otimes 1) \psi(a_{(2)} \otimes 1) \\ &= \sum_{(a)} \phi_1(a_{(1)}) \psi_1(a_{(2)}) \\ &= (\phi_1 \otimes \psi_1)(\Delta(a)) = \phi_1 * \psi_1(a). \end{aligned}$$

Similarly, we get $\phi \tilde{*} \psi(1 \otimes a) = \phi_2 * \psi_2(a)$ for all $a \in A$. For any such coupling of the convolutions $\phi_1 * \psi_1$ and $\phi_2 * \psi_2$, we have the following inequality. Since $\phi \otimes \psi$ is positive, by (4.18), we get

$$\begin{aligned} \phi \tilde{*} \psi(C_H) &= (\phi \otimes \psi)(\tilde{\Delta}(C_H)) \\ &\leq \phi \otimes \psi(\iota_{1,2}(C_H)) + \phi \otimes \psi(\iota_{3,4}(C_H)) \\ &= \phi(C_H) \psi(1 \otimes 1) + \phi(1 \otimes 1) \psi(C_H) = \phi(C_H) + \psi(C_H). \end{aligned}$$

Now, taking the infimum over all the couplings of the convolutions gives us the required inequality \square

5. Connections and future directions

5.1. Comparison of distances and recovery of classical case

The next proposition gives inequalities relating the different versions of the Hamming distance that we have defined.

Proposition 5.1. For $\varphi_1, \varphi_2 \in \mathcal{T}(C(S_n^+))$, we have

$$W_{H,1}(\varphi_1, \varphi_2) \leq W_{H,\wedge}(\varphi_1, \varphi_2) \leq W_{H,\otimes}(\varphi_1, \varphi_2) \leq 1.$$

Proof. To show that $W_{H,1}(\varphi_1, \varphi_2) \leq W_{H,\wedge}(\varphi_1, \varphi_2)$, consider a tracial coupling $(M, \tau, \alpha_1, \alpha_2)$. Observe that

$$\begin{aligned} \|\alpha_1(u_{ij}) - \alpha_2(u_{ij})\|_{L^1(M,\tau)} &\leq \|\alpha_1(u_{ij}) - \alpha_1(u_{ij}) \wedge \alpha_2(u_{ij})\|_{L^1(M,\tau)} + \|\alpha_2(u_{ij}) - \alpha_1(u_{ij}) \wedge \alpha_2(u_{ij})\|_{L^1(M,\tau)} \\ &= \tau[\alpha_1(u_{ij}) - \alpha_1(u_{ij}) \wedge \alpha_2(u_{ij})] + \tau[\alpha_2(u_{ij}) - \alpha_1(u_{ij}) \wedge \alpha_2(u_{ij})] \\ &= \tau(\alpha_1(u_{ij})) + \tau(\alpha_2(u_{ij})) - 2\tau(\alpha_1(u_{ij}) \wedge \alpha_2(u_{ij})) \end{aligned}$$

since $\alpha_1(u_{ij}) \geq \alpha_1(u_{ij}) \wedge \alpha_2(u_{ij})$ and the same for $\alpha_2(u_{ij})$. Since $\frac{1}{n} \sum_{i,j=1}^n u_{ij} = 1$, this implies that

$$\frac{1}{n} \sum_{i,j=1}^n \|\alpha_1(u_{ij}) - \alpha_2(u_{ij})\|_{L^1(M,\tau)} \leq 2 \left(1 - \frac{1}{n} \sum_{i,j=1}^n \tau(\alpha_1(u_{ij}) \wedge \alpha_2(u_{ij})) \right).$$

Since the tracial coupling was arbitrary, we obtain the desired inequality upon dividing by 2.

To show that $W_{H,\wedge}(\varphi_1, \varphi_2) \leq W_{H,\otimes}(\varphi_1, \varphi_2)$, let $\varphi \in \Pi_{\text{tr},\otimes}(\varphi_1, \varphi_2)$. Let (H_φ, π_φ) be the GNS construction associated to $A \otimes A$ and φ . Let $M = \pi_\varphi(A \otimes A)''$ and let τ be the vector state on $B(H_\varphi)$ given by $\widehat{1}$. By Lemma 2.2, (M, τ) is a tracial von Neumann algebra. Let $\alpha_1, \alpha_2 : A \rightarrow M$ be the inclusions of the first and second tensorands into $A \otimes A$ composed with the map $\pi_\varphi : A \otimes A \rightarrow M$. Then $(M, \tau, \alpha_1, \alpha_2)$ is a tracial coupling in the sense of Definition 1.1. Therefore,

$$W_{H,\wedge}(\varphi_1, \varphi_2) \leq 1 - \frac{1}{n} \sum_{i,j=1}^n \tau(\alpha_1(u_{ij}) \wedge \alpha_2(u_{ij})).$$

Since $\alpha_1(A)$ and $\alpha_2(A)$ commute, $\alpha_1(u_{ij}) \wedge \alpha_2(u_{ij}) = \alpha_1(u_{ij})\alpha_2(u_{ij})$. Thus,

$$W_{H,\wedge}(\varphi_1, \varphi_2) \leq 1 - \frac{1}{n} \sum_{i,j=1}^n \tau(\alpha_1(u_{ij})\alpha_2(u_{ij})) = 1 - \frac{1}{n} \sum_{i,j=1}^n \varphi(u_{ij} \otimes u_{ij}).$$

Taking the infimum over φ , we have $W_{H,\wedge}(\varphi_1, \varphi_2) \leq W_{H,\otimes}(\varphi_1, \varphi_2)$ as desired.

The inequality $W_{H,\otimes}(\varphi_1, \varphi_2) \leq 1$ is immediate because the cost operator C_H satisfies $C_H \leq 1$. \square

We can further compare the distances defined in this paper with the norm $\|\varphi_1 - \varphi_2\|_{C(S_n^+)^*}$, where $C(S_n^+)^*$ denotes the dual space. This is motivated by the fact that Wasserstein distance is bounded by a constant time the total variation distance for probability measures on a compact metric space. Recall that probability measures on a compact Hausdorff space X are equivalent to states on $C(X)$. The total variation distance of two probability measures is $d_{TV}(\mu, \nu) = \sup_{A \subseteq X} \{|\mu(A) - \nu(A)|\}$ where A ranges over Borel subsets of X , and this agrees with $(1/2)\|\mu - \nu\|_{C(X)^*}$. Hence, for states φ_1, φ_2 on a C^* -algebra A , the quantity $(1/2)\|\varphi_1 - \varphi_2\|_{A^*}$ serves as an analog of the total variation distance.

Proposition 5.2. For $\varphi_1, \varphi_2 \in \mathcal{T}(C(S_n^+))$, we have

$$W_{H,\wedge}(\varphi_1, \varphi_2) \leq \frac{1}{2} \|\varphi_1 - \varphi_2\|_{C(S_n^+)^*}.$$

Proof. We begin by recalling the analog of Jordan decomposition for tracial functionals. Suppose A is a C^* -algebra and $\varphi \in A^*$ is self-adjoint ($\varphi(a^*) = \overline{\varphi(a)}$) and tracial. By [51, Proposition 2.8], φ has a unique decomposition as $\varphi_+ - \varphi_-$, where φ_+ and φ_- are positive and tracial with $\|\varphi\| = \|\varphi_+\| + \|\varphi_-\|$. Furthermore, the proof gives the following description of φ_+ . For $x, y \in A_+$, write $x \sim y$ if there exist elements $(u_n)_{n \in \mathbb{N}}$ in A such that

$$x = \sum_{n \in \mathbb{N}} u_n^* u_n, \quad y = \sum_{n \in \mathbb{N}} u_n u_n^*$$

where the sums are norm-convergent (see [51, p. 136]). Then for $x \in A_+$,

$$\begin{aligned} \varphi_+(x) &= \sup\{\varphi(y) : y, z \in A_+, y + z \sim x\} \\ \varphi_-(x) &= \sup\{-\varphi(y) : y, z \in A_+, y + z \sim x\}. \end{aligned}$$

Now suppose that $\varphi = \varphi_1 - \varphi_2$ where φ_1, φ_2 are tracial states on A (note that such a φ is automatically tracial and self-adjoint). In this case, if $y + z \sim x$, then

$$\varphi(y) = \varphi_1(y) - \varphi_2(y) \leq \varphi_1(y) \leq \varphi_1(y+z) = \varphi_1(x).$$

Taking the supremum over such y, z , we obtain $\varphi_+(x) \leq \varphi_1(x)$. By symmetrical reasoning $\varphi_- \leq \varphi_2$.

Hence, for tracial states φ_1, φ_2 on A , there exist tracial positive functionals φ_+ and φ_- with

$$\begin{aligned} \varphi_1 - \varphi_2 &= \varphi_+ - \varphi_-, & \varphi_+ &\leq \varphi_1, \\ \|\varphi_1 - \varphi_2\|_{A^*} &= \|\varphi_+\|_{A^*} + \|\varphi_-\|_{A^*}, & \varphi_- &\leq \varphi_2, \end{aligned}$$

and furthermore,

$$\|\varphi_+\|_{A^*} - \|\varphi_-\|_{A^*} = \varphi_+(1) - \varphi_-(1) = \varphi_1(1) - \varphi_2(1) = 0,$$

whence

$$\|\varphi_+\|_{A^*} = \|\varphi_-\|_{A^*} = \frac{1}{2} \|\varphi_1 - \varphi_2\|_{A^*}.$$

Now let us specialize to the case of $A = C(S_n^+)$. Note that if $\varphi_1 = \varphi_2$, then the claimed statement holds with both sides being zero. Moreover, if $\|\varphi_1 - \varphi_2\|_{A^*} = 2$, it is also immediate since $W_{H,\wedge}(\varphi_1, \varphi_2) \leq 1$ by Proposition 5.1. Hence, assume without loss of generality that $0 < \|\varphi_1 - \varphi_2\|_{A^*} < 2$, so that $0 < \|\varphi_+\|_{A^*} < 1$. Let

$$\psi = \varphi_1 - \varphi_+ = \varphi_2 - \varphi_-.$$

Let $\bar{\psi} = \psi/\|\psi\|$ and define $\bar{\varphi}_+$ and $\bar{\varphi}_-$ analogously. Thus, $\bar{\psi}, \bar{\varphi}_+$, and $\bar{\varphi}_-$ are tracial states on A . Let $N = \pi_{\bar{\psi}}(A)''$, $M_+ = \pi_{\bar{\varphi}_+}(A)''$, and $M_- = \pi_{\bar{\varphi}_-}(A)''$ be the corresponding von Neumann algebras. Then let

$$M = N \oplus (M_+ \bar{\otimes} M_-), \quad \tau = \psi(1)\bar{\psi} \oplus \varphi_+(1)(\bar{\varphi}_+ \otimes \bar{\varphi}_-).$$

Let $\alpha_1, \alpha_2 : A \rightarrow M$ be given by

$$\begin{aligned} \alpha_1(a) &= \pi_{\bar{\psi}}(a) \oplus (\pi_{\bar{\varphi}_+}(a) \otimes 1) \\ \alpha_2(a) &= \pi_{\bar{\psi}}(a) \oplus (1 \otimes \pi_{\bar{\varphi}_-}(a)). \end{aligned}$$

Thus, $(M, \tau, \alpha_1, \alpha_2)$ is a tracial coupling of $\varphi_1 = \psi + \varphi_+$ and $\varphi_2 = \psi + \varphi_-$. Note that

$$\alpha_1(u_{ij}) \wedge \alpha_2(u_{ij}) \geq \pi_{\bar{\psi}}(u_{ij}) \oplus 0.$$

Hence,

$$\begin{aligned} 1 - \frac{1}{n} \sum_{i,j=1}^n \tau(\alpha_1(u_{ij}) \wedge \alpha_2(u_{ij})) &\leq 1 - \frac{1}{n} \sum_{i,j=1}^n \psi(1)\bar{\psi}(u_{ij}) = 1 - \frac{1}{n} \sum_{i=1}^n \psi(1) \\ &= 1 - \psi(1) = \varphi_+(1) = \frac{1}{2} \|\varphi_1 - \varphi_2\|_{C(S_n^+)^*}, \end{aligned}$$

which proves the asserted inequality. \square

Every probability measure on S_n gives rise to a state on $C(S_n)$ and hence a state on $C(S_n^+)$ via the quotient map θ from Theorem 2.14. We will now show that the non-commutative Wasserstein distance $W_{H,\wedge}(\varphi_1, \varphi_2)$ agrees with the classical L^1 Wasserstein distance of μ_1 and μ_2 with respect to the normalized Hamming distance. To prove this, we must extract tracial couplings of the states from classical couplings of the measures and vice versa. To obtain a classical coupling from a tracial coupling a von Neumann algebra M that is not necessarily commutative, we use the trick of separating the left and right multiplication as in [28, Theorem 1.5], closely related to Connes' joint distribution trick.

Proposition 5.3. *Let μ_1 and $\mu_2 \in \mathcal{P}(S_n)$, and let φ_1, φ_2 be the corresponding traces on $C(S_n^+)$ given by*

$$\varphi_j(a) = \int_{S_n} \theta(a) d\mu_j.$$

Let W_1 be the classical L^1 -Wasserstein distance on $\mathcal{P}(S_n)$ associated to the normalized Hamming distance d_H on S_n . Then

$$W_{H,1}(\varphi_1, \varphi_2) = W_{H,\wedge}(\varphi_1, \varphi_2) = W_{H,\otimes}(\varphi_1, \varphi_2) = W_1(\mu_1, \mu_2).$$

Proof. By Proposition 5.1, it suffices to show that

$$W_{H,\otimes}(\varphi_1, \varphi_2) \leq W_1(\mu_1, \mu_2)$$

and

$$W_1(\mu_1, \mu_2) \leq W_{H,1}(\varphi_1, \varphi_2).$$

For the first claim, suppose that μ is a classical coupling of μ_1 and μ_2 . Let $\varphi \in \mathcal{T}(C(S_n^+) \otimes C(S_n^+))$ be the state given by $\varphi(a) = \int_{S_n \times S_n} (\theta \otimes \theta)(a) d\mu$. Then $\varphi(a \otimes 1) = \varphi_1(a)$ and symmetrically for φ_2 . We also have that

$$\varphi(C_H) = \int_{S_n \times S_n} \left(1 - \frac{1}{n} \sum_{i,j=1}^n \theta(u_{ij}) \otimes \theta(u_{ij}) \right) d\mu = \int_{S_n \times S_n} d_H d\mu.$$

Since μ was arbitrary, $W_{H,\otimes}(\varphi_1, \varphi_2) \leq W_1(\mu_1, \mu_2)$ which completes the proof.

For the second claim, consider a tracial coupling $(M, \tau, \alpha_1, \alpha_2)$ of φ_1 and φ_2 . In order to extract a classical coupling from this, we will first reinterpret $(M, \tau, \alpha_1, \alpha_2)$ as a tracial coupling of states on the commutative $C(S_n)$. Let $\bar{\varphi}_i$ be the state on $C(S_n)$ given by the measure μ_i , so that $\varphi_i = \bar{\varphi}_i \circ \theta$. Note that the image $\alpha_i(C(S_n^+))$ is $*$ -isomorphic to image of the GNS representation $\pi_{\varphi_i}(C(S_n^+))$ using Lemma 2.1 (3) and the faithfulness of τ , and the state $\tau|_{\alpha_i(C(S_n^+)}}$ corresponds to φ_i . Moreover, by Lemma 2.1 (2), θ passes to an isomorphism from $\pi_{\varphi_i}(C(S_n^+))$ to $\pi_{\bar{\varphi}_i}(C(S_n))$ which also preserves the given states. Hence, $\alpha_i(C(S_n^+))$ is isomorphic to $\pi_{\bar{\varphi}_i}(C(S_n))$, which implies that the map α_i factors through $C(S_n)$, or $\alpha_i = \bar{\alpha}_i \circ \theta$ for some $*$ -homomorphism $\bar{\alpha}_i : C(S_n) \rightarrow M$; moreover, $\bar{\varphi}_i = \tau \circ \bar{\alpha}_i$. Therefore, $(M, \tau, \bar{\alpha}_1, \bar{\alpha}_2)$ is a tracial coupling of $\bar{\varphi}_1$ and $\bar{\varphi}_2$.

Next, we construct a classical coupling out of $(M, \tau, \bar{\alpha}_1, \bar{\alpha}_2)$. Recall from §2.2.1 that the left and right multiplication representations $\lambda : M \rightarrow B(L^2(M, \tau))$ and $\rho : M^{\text{op}} \rightarrow B(L^2(M, \tau))$ commute. Hence, $\lambda \circ \bar{\alpha}_1$ and $\rho \circ \bar{\alpha}_2$ are commuting representations of $C(S_n)$ and $C(S_n^{\text{op}})$ on $L^2(M, \tau)$, and of course $C(S_n)^{\text{op}} = C(S_n)$ since $C(S_n)$ is commutative. Thus, $(\lambda \circ \bar{\alpha}_1) \otimes (\rho \circ \bar{\alpha}_2)$ gives a $*$ -representation of $C(S_n) \otimes C(S_n)$ on $L^2(M, \tau)$. Let $\bar{\varphi}$ be the state on $C(S_n) \otimes C(S_n)$ given by

$$\bar{\varphi}(x) = \langle \widehat{1}_M, (\lambda \circ \bar{\alpha}_1) \otimes (\rho \circ \bar{\alpha}_2)(x) \widehat{1}_M \rangle_{L^2(M, \tau)},$$

and let μ be the measure on $S_n \times S_n$ corresponding to the state $\bar{\varphi}$. Note that given $a, b \in C(S_n)$,

$$\bar{\varphi}(a \otimes b) = \langle \widehat{1}, \lambda(\alpha_1(a))\rho(\alpha_2(b))\widehat{1} \rangle = \langle \widehat{1}, \alpha_1(a)\widehat{1}\alpha_2(b) \rangle = \tau(\bar{\alpha}_1(a)\bar{\alpha}_2(b)).$$

In particular, for $a \in C(S_n)$,

$$\int_{S_n \times S_n} a \otimes 1 d\mu = \bar{\varphi}(a \otimes 1) = \tau(\bar{\alpha}_1(a)) = \bar{\varphi}_1(a) = \int_{S_n} a d\mu_1,$$

which shows that the first marginal of μ is μ_1 , and symmetrically the second marginal is μ_2 . Hence, μ is a classical coupling of (μ_1, μ_2) .

The cost of the classical coupling μ is given by

$$\begin{aligned} \int_{S_n \times S_n} d_H(\sigma, \sigma') d\mu &= 1 - \frac{1}{n} \sum_{i,j} \bar{\varphi}(\theta(u_{ij}) \otimes \theta(u_{ij})) = 1 - \frac{1}{n} \sum_{i,j=1}^n \tau(\bar{\alpha}_1(\theta(u_{ij}))\bar{\alpha}_2(\theta(u_{ij}))) \\ &= 1 - \frac{1}{n} \sum_{i,j=1}^n \tau(\alpha_1(u_{ij})\alpha_2(u_{ij})) \\ &= \frac{1}{n} \sum_{i,j=1}^n \left(\frac{1}{2} \tau(\alpha_1(u_{ij})) + \frac{1}{2} \tau(\alpha_2(u_{ij})) - \tau(\alpha_1(u_{ij})\alpha_2(u_{ij})) \right) \\ &= \frac{1}{n} \sum_{i,j=1}^n \frac{1}{2} \|\alpha_1(u_{ij}) - \alpha_2(u_{ij})\|_{L^2(M, \tau)}^2. \end{aligned}$$

Since $\alpha_1(u_{ij})$ and $\alpha_2(u_{ij})$ are nonnegative operators, the Powers-Størmer inequality shows that

$$\|\alpha_1(u_{ij}) - \alpha_2(u_{ij})\|_{L^2(M, \tau)}^2 \leq \|\alpha_1(u_{ij})^2 - \alpha_2(u_{ij})^2\|_{L^1(M, \tau)} = \|\alpha_1(u_{ij}) - \alpha_2(u_{ij})\|_{L^1(M, \tau)},$$

and therefore

$$W_1(\mu_1, \mu_2) \leq \int_{S_n \times S_n} d_H(\sigma, \sigma') d\mu \leq \frac{1}{n} \sum_{i,j=1}^n \frac{1}{2} \|\alpha_1(u_{ij}) - \alpha_2(u_{ij})\|_{L^1(M,\tau)}.$$

Since the tracial coupling $(M, \tau, \alpha_1, \alpha_2)$ was arbitrary, we obtain the asserted inequality. \square

The next proposition characterizes when each of the distances studied here achieves the maximum value of 1. Condition (4) below means that the two states are supported on disjoint sets of entries of the canonical unitary $U = [u_{ij}]$, namely $\{(i, j) : \varphi_1(u_{ij}) > 0\}$ and $\{(i, j) : \varphi_2(u_{ij}) > 0\}$ are disjoint. For instance, if φ_j was induced by a measure μ_j on the classical permutation group as in Proposition 5.3, this would mean that each $\sigma_1 \in \text{supp}(\mu_1)$ and $\sigma_2 \in \text{supp}(\mu_2)$ are supported in disjoint sets of indices, or $\sigma_1^{-1}\sigma_2$ has no fixed points.

Proposition 5.4. *Let $\varphi_1, \varphi_2 \in \mathcal{T}(C(S_n^+))$. Then the following are equivalent:*

- (1) $W_{H,1}(\varphi_1, \varphi_2) = 1$.
- (2) $W_{H,\wedge}(\varphi_1, \varphi_2) = 1$.
- (3) $W_{H,\otimes}(\varphi_1, \varphi_2) = 1$.
- (4) $\sum_{i,j=1}^n \varphi_1(u_{ij})\varphi_2(u_{ij}) = 0$.

Proof. (1) \implies (2) \implies (3) is immediate from Proposition 5.1.

(3) \implies (4), note that $\varphi_1 \otimes \varphi_2$ is a tracial tensor coupling of φ_1, φ_2 , and hence

$$W_{H,\otimes}(\varphi_1, \varphi_2) \leq 1 - \frac{1}{n} \sum_{i,j=1}^n \varphi_1(u_{ij})\varphi_2(u_{ij}).$$

So if $W_{H,\otimes}(\varphi_1, \varphi_2) = 1$, then $\sum_{i,j=1}^n \varphi_1(u_{ij})\varphi_2(u_{ij}) = 0$.

For (4) \implies (1), assume (4). Let $(M, \tau, \alpha_1, \alpha_2)$ be a tracial coupling of (φ_1, φ_2) . For every i, j , either $\varphi_1(u_{ij}) = 0$ or $\varphi_2(u_{ij}) = 0$. Thus, either $\tau(\alpha_1(u_{ij})) = 0$ or $\tau(\alpha_2(u_{ij})) = 0$, and hence either $\alpha_1(u_{ij}) = 0$ or $\alpha_2(u_{ij}) = 0$. Therefore,

$$\|\alpha_1(u_{ij}) - \alpha_2(u_{ij})\|_{L^1(M,\tau)} = \|\alpha_1(u_{ij})\|_{L^1(M,\tau)} + \|\alpha_2(u_{ij})\|_{L^2(M,\tau)} = \tau(\alpha_1(u_{ij})) + \tau(\alpha_2(u_{ij})).$$

Hence,

$$\frac{1}{n} \sum_{i,j=1}^n \frac{1}{2} \|\alpha_1(u_{ij}) - \alpha_2(u_{ij})\|_{L^1(M,\tau)} = \frac{1}{2n} \sum_{i,j=1}^n [\tau(\alpha_1(u_{ij})) + \tau(\alpha_2(u_{ij}))] = 1,$$

and since $(M, \tau, \alpha_1, \alpha_2)$ was arbitrary, (1) holds. \square

5.2. Lipschitz seminorms induced by quantum Wasserstein distances

Recall that a Wasserstein distance on the trace space always induces a corresponding Lipschitz norm on the C^* -algebra. A minimal requirement for a reasonable definition is that the space of Lipschitz elements should be dense in the C^* -algebra, so we now verify this property for $W_{H,\wedge}$ and $W_{H,1}$.

Proposition 5.5. *For all indices $i_1, j_1, \dots, i_\ell, j_\ell \in [n]$, we have*

$$\|u_{i_1 j_1} \dots u_{i_\ell j_\ell}\|_{\text{Lip}(W_{H,\wedge})} \leq \|u_{i_1 j_1} \dots u_{i_\ell j_\ell}\|_{\text{Lip}(W_{H,1})} \leq 2n\ell.$$

In particular, the $$ -algebra generated by the u_{ij} 's is contained in the domains of the Lipschitz seminorms associated to $W_{H,\wedge}$ and $W_{H,1}$.*

Proof. The first inequality follows because $W_{H,1} \leq W_{H,\wedge}$. For the second inequality, let $\varphi_1, \varphi_2 \in \mathcal{T}(C(S_n^+))$, and let $(M, \tau, \alpha_1, \alpha_2)$ be a tracial coupling. Then using the triangle inequality and non-commutative Hölder's inequality,

$$\begin{aligned} |\varphi_1(u_{i_1 j_1} \dots u_{i_\ell j_\ell}) - \varphi_2(u_{i_1 j_1} \dots u_{i_\ell j_\ell})| &\leq |\tau(\alpha_1(u_{i_1 j_1} \dots u_{i_\ell j_\ell}) - \alpha_2(u_{i_1 j_1} \dots u_{i_\ell j_\ell}))| \\ &\leq \sum_{t=1}^{\ell} \|\alpha_1(u_{i_1 j_1} \dots u_{i_{t-1} j_{t-1}})(\alpha_1(u_{i_t j_t}) - \alpha_2(u_{i_t j_t}))\| \\ &\quad \|\alpha_2(u_{i_{t+1} j_{t+1}} \dots u_{i_\ell j_\ell})\|_{L^1(M,\tau)} \end{aligned}$$

$$\begin{aligned} &\leq \sum_{t=1}^{\ell} \|\alpha_1(u_{i_t j_t}) - \alpha_2(u_{i_t j_t})\|_{L^1(M, \tau)} \\ &\leq \ell \max_{i, j} \|\alpha_1(u_{ij}) - \alpha_2(u_{ij})\|_{L^1(M, \tau)} \\ &\leq 2n\ell \cdot \frac{1}{n} \sum_{i, j=1}^n \frac{1}{2} \|\alpha_1(u_{ij}) - \alpha_2(u_{ij})\|_{L^1(M, \tau)}. \end{aligned}$$

Taking the infimum over the coupling $(M, \tau, \alpha_1, \alpha_2)$ shows that

$$|\varphi_1(u_{i_1 j_1} \dots u_{i_\ell j_\ell}) - \varphi_2(u_{i_1 j_1} \dots u_{i_\ell j_\ell})| \leq 2n\ell W_{H,1}(\varphi_1, \varphi_2),$$

which is the desired bound for the Lipschitz norm. \square

5.3. Further questions

We hope that the different distances defined in this paper and these properties are the beginning of further investigations. In particular, we propose the following questions.

Question 5.6 (Duality for Wasserstein distances and Lipschitz seminorms). Do the distances $W_{H,\wedge}$ and $W_{H,1}$ defined in this paper satisfy the duality $d(\varphi, \psi) = \sup\{|\varphi(x) - \psi(x)| : \|x\|_{\text{Lip}} \leq 1\}$?

Question 5.7 (Topologies on $\mathcal{T}(C(S_n^+))$). We now know of at least four topologies on $\mathcal{T}(C(S_n^+))$:

- (1) the weak- $*$ topology,
- (2) the topology generated by $W_{H,1}$,
- (3) the topology generated by $W_{H,\wedge}$,
- (4) the topology generated by the total variation distance, i.e. the restriction of the norm topology of $C(S_n^+)^*$.

Each topology in this list is at least as strong as the ones above it, thanks to Propositions 3.10, 5.1, and 5.2. We do not yet know if all these topologies are distinct, but this is quite plausible. For instance, based on [32, §5.5], we expect that the $W_{H,1}$ would generate a stronger topology than the weak- $*$ topology. Moreover, since projections p_k can converge to p without $p_k \wedge p$ converging p , we expect that $W_{H,\wedge}$ is much stronger than $W_{H,1}$. The computation of $W_{H,\wedge}$ may also lead to interesting combinatorial problems about how to arrange non-commuting projections in a way that maximizes the pairwise intersections.

Question 5.8 (Behavior under iterated convolution). For any faithful state on $C(S_n^+)$, the repeated convolution φ^{*k} converges in the weak- $*$ topology as $k \rightarrow \infty$ to the Haar state h . Under what conditions does φ^{*k} converge to h with respect to $W_{H,1}$ or $W_{H,\wedge}$. Can we describe the limit of φ^{*k} when φ is not faithful, and determine whether convergence occurs with respect to the metrics in this paper? Another related question is how the metrics behave with respect to the convolution semigroups studied in [23]. We remark that convergence in total variation distance for various convolution semigroups of states has been shown in [52,53], though these states are not tracial.

Question 5.9 (Extreme traces on $C(S_n^+)$). The study of $W_{H,\otimes}$ raises many questions about extreme points in the trace space $\mathcal{T}(C(S_n^+))$. In particular, can we classify the extreme points that arise from finite-dimensional representations of $C(S_n^+)$? How does the trace space relate to corepresentations of S_n^+ ? Is $\mathcal{T}(C(S_n^+))$ a Poulsen simplex?

Question 5.10 (Projection of extreme traces onto Birkhoff polytope). Let $B_n = \{A \in M_n(\mathbb{R}) : A_{ij} \in [0, 1], \sum_i A_{ij} = 1, \sum_j A_{ij} = 1\}$ be the Birkhoff polytope of doubly stochastic matrices. Consider the map $\omega : \partial_e \mathcal{T}(C(S_n^+)) \rightarrow B_n$ by sending φ to $[\varphi(u_{ij})]_{i,j=1}^n$. Under what conditions does $\omega(\varphi) = \omega(\psi)$? Are there natural sets on which ω is injective?

Funding

A was partially supported by the grant ‘‘Logic and C*-algebras’’ from the National Sciences and Engineering Research Council (Canada). DJ was partially supported by the Danish Independent Research Fund, grant 1026-00371B; and an EU Horizon Marie Skłodowska Curie Action, FREEINFOGEOM, grant id: 101209517. Views and opinions expressed are those of the author(s) only and do not necessarily reflect those of the European Union or the Research Executive Agency. Neither the European Union nor the granting authority can be held responsible for them.

Acknowledgements

We thank the Fields Institute for their hospitality during the Thematic Program on Operator Algebras in Fall 2023, where this collaboration began. We thank the Institute for Pure and Applied Mathematics for their hospitality during the long program on Noncommutative Optimal Transport in Spring 2025, including travel funding for A and DJ's visits. IPAM is supported by the National Science Foundation (US), grant DMS-1925919.

We thank Bhishan Jacelon for discussions of metrics on trace spaces and for suggesting to show density of the Lipschitz functions. We thank Milad Marvian for suggesting to characterize when the distance is equal to 1. We thank Peixue Wu for references on quantum Wasserstein distances and Mikael Rørdam for references on trace spaces. We thank Anna Wysocańska-Kula and Uwe Franz for comments on the draft and discussion of recent work on convolution semigroups.

Data availability

No data was used for the research described in the article.

References

- [1] M.A. Rieffel, Metrics on states from actions of compact groups, *Doc. Math.* 3 (1998) 215–229.
- [2] M.A. Rieffel, Metrics on state spaces, *Doc. Math.* 4 (1999) 559–600.
- [3] B. Jacelon, Metrics on trace spaces, *J. Funct. Anal.* 285 (4) (2023) 109977, <https://doi.org/10.1016/j.jfa.2023.109977>, <https://www.sciencedirect.com/science/article/pii/S0022123623001349>.
- [4] B. Jacelon, Quantum metric Choquet simplices, preprint, arXiv:2408.04368, 2024.
- [5] A. Skalski, I.G. Todorov, L. Turowska, Coupling capacity in C^* -algebras, *Proc. R. Soc. Edinb., Sect. A, Math.* 155 (1) (2025) 81–103, <https://doi.org/10.1017/prm.2023.81>.
- [6] E. Caputo, A. Gerolin, N. Monina, L. Portinale, Quantum optimal transport with convex regularization, preprint, arXiv:2409.03698.
- [7] R. Araiza, Y. Chen, M. Junge, P. Wu, Resource-dependent complexity of quantum channels, preprint, arXiv:2303.11304, 2023.
- [8] A. Connes, Compact metric spaces, Fredholm modules, and hyperfiniteness, *Ergod. Theory Dyn. Syst.* 9 (1989) 207–220.
- [9] A. Connes, *Noncommutative Geometry*, Academic Press, Inc., San Diego, CA, 1994.
- [10] M.A. Rieffel, Group C^* -algebras as compact quantum metric spaces, *Doc. Math.* 7 (2002) 605–651, <https://doi.org/10.4171/DM/133>.
- [11] K. Aguilar, K. Behera, T. Omland, N. Wu, On the Bures metric, C^* -norm and quantum metric, *Bull. Aust. Math. Soc.* (2025) 1–11, <https://doi.org/10.1017/S0004972724001217>.
- [12] M. Christ, M.A. Rieffel, Nilpotent group C^* -algebras as compact quantum metric spaces, *Can. Math. Bull.* 60 (2017) 77–94.
- [13] J. Bhowmick, C. Voigt, J. Zacharias, Compact quantum metric spaces from quantum groups of rapid decay, *J. Noncommut. Geom.* 9 (2015) 1175–1200, <https://doi.org/10.4171/JNCG/220>.
- [14] A. Austad, D. Kyed, Quantum metrics from length functions on quantum groups, preprint, arXiv:2503.01501, <https://arxiv.org/abs/2503.01501>.
- [15] K. Aguilar, J. Kaad, The Podleś sphere as a spectral metric space, *J. Geom. Phys.* 133 (03 2018), <https://doi.org/10.1016/j.geomphys.2018.07.015>.
- [16] K. Aguilar, J. Kaad, D. Kyed, The Podleś spheres converge to the sphere, *Commun. Math. Phys.* 392 (06 2022), <https://doi.org/10.1007/s00220-022-04363-4>.
- [17] K. Aguilar, J. Kaad, D. Kyed, Polynomial approximation of quantum Lipschitz functions, *Doc. Math.* 27 (2022) 765–787.
- [18] S.L. Woronowicz, Pseudospaces, pseudogroups and Pontryagin duality, in: *Proceedings of the International Conference on Mathematics and Physics, Lausanne*, in: *Lecture Notes in Physics*, vol. 116, 1979, pp. 407–412.
- [19] S.L. Woronowicz, Compact matrix pseudogroups, *Commun. Math. Phys.* 111 (4) (1987) 613–665, <https://doi.org/10.1007/BF01218298>.
- [20] S.L. Woronowicz, Compact quantum groups, 1998.
- [21] S. Wang, Quantum symmetry groups of finite spaces, *Commun. Math. Phys.* 195 (1) (1998) 195–211.
- [22] E.A. Carlen, J. Maas, Gradient flow and entropy inequalities for quantum Markov semigroups with detailed balance, *J. Funct. Anal.* 273 (5) (2017) 1810–1869, <https://doi.org/10.1016/j.jfa.2017.05.003>, <https://www.sciencedirect.com/science/article/pii/S0022123617301878>.
- [23] U. Franz, A. Kula, A. Skalski, Lévy processes on quantum permutation groups, in: *Linear Operators and Linear Systems*, vol. 252, Birkhäuser, Basel, 2016, pp. 193–259.
- [24] R. Duvenhage, Quadratic Wasserstein metrics for von Neumann algebras via transport plans, *J. Oper. Theory* 88 (2) (2022) 289–308.
- [25] G. De Palma, M. Marvian, D. Trevisan, S. Lloyd, The quantum Wasserstein distance of order 1, *IEEE Trans. Inf. Theory* 67 (10) (2021) 6627–6643, <https://doi.org/10.1109/TIT.2021.3076442>.
- [26] M.A. Rieffel, Quantum Hamming metrics, preprint, arXiv:2507.23046, 2025.
- [27] K. Aguilar, Z. Batterman, The Hamming distance and the Fell topology on af algebras, preprint, arXiv:2310.10752.
- [28] P. Biane, D. Voiculescu, A free probability analogue of the Wasserstein metric on the trace-state space, *Geom. Funct. Anal.* 11 (6) (2001) 1125–1138.
- [29] B. Blackadar, M. Rørdam, The space of tracial states on a C^* -algebra, *Expo. Math.* (2024) 125618, <https://doi.org/10.1016/j.exmath.2024.125618>, <https://www.sciencedirect.com/science/article/pii/S0723086924000859>.
- [30] J.R. Carrión, J. Gabe, C. Schafhauser, A. Tikuisis, S. White, Classifying $*$ -homomorphisms I: unital simple nuclear C^* -algebras, arXiv:2307.06480, 2023.
- [31] B. Jacelon, Chaotic tracial dynamics, *Forum Math. Sigma* 11 (2023) e39, <https://doi.org/10.1017/fms.2023.38>.
- [32] W. Gangbo, D. Jekel, K. Nam, D. Shlyakhtenko, Duality for optimal couplings in free probability, *Commun. Math. Phys.* 396 (3) (2022) 903–981.
- [33] A. Freslon, A. Skalski, S. Wang, Tracial central states on compact quantum groups, *J. Funct. Anal.* 289 (7) (2025) 110988, <https://doi.org/10.1016/j.jfa.2025.110988>.
- [34] G.J. Murphy, *C^* -Algebras and Operator Theory*, Academic Press, Boston, 1990.
- [35] K.R. Davidson, *C^* -Algebras by Example*, Fields Institute Monographs, vol. 6, American Mathematical Society, 1996.
- [36] B. Blackadar, *Operator Algebras: Theory of C^* -Algebras and von Neumann Algebras*, Encyclopaedia of Mathematical Sciences, vol. 122, Springer-Verlag, Berlin, Heidelberg, 2006, <https://wolfweb.unr.edu/homepage/bruceb/Cycr.pdf>.
- [37] B. Blackadar, Shape theory for C^* -algebras, *Math. Scand.* 56 (1985) 249–275, <http://eudml.org/doc/166937>.
- [38] D. Avitzour, Free products of C^* -algebras, *Trans. Am. Math. Soc.* 271 (1982) 423–435.
- [39] V. Jones, V.S. Sunder, *Introduction to Subfactors*, London Mathematical Society Lecture Note Series, Cambridge University Press, 1997.
- [40] G. Pisier, Q. Xu, Non-commutative L^p -spaces, in: W.B. Johnson, J. Lindenstrauss (Eds.), *Handbook of the Geometry of Banach Spaces*, vol. 2, Elsevier, 2003, pp. 1459–1517.

- [41] D.-V. Voiculescu, Symmetries of some reduced free product C^* -algebras, in: H. Araki, C.C. Moore, Ş.-V. Stratila, D.-V. Voiculescu (Eds.), *Operator Algebras and Their Connections with Topology and Ergodic Theory*, Springer Berlin Heidelberg, Berlin, Heidelberg, 1985, pp. 556–588.
- [42] S. Popa, Markov traces on universal Jones algebras and subfactors of finite index, *Invent. Math.* 111 (1993) 375–405, <https://doi.org/10.1007/BF01231293>.
- [43] N.P. Brown, N. Ozawa, *C^* -Algebras and Finite-Dimensional Approximations*, Graduate Studies in Mathematics, vol. 88, American Mathematical Society, Providence, 2008.
- [44] D. Voiculescu, K. Dykema, A. Nica, *Free Random Variables*, 1992.
- [45] T. Banica, J. Bichon, B. Collins, Quantum permutation groups: a survey, *Banach Cent. Publ.* 78 (1) (2008) 13–34, <http://eudml.org/doc/282193>.
- [46] M. Lupini, L. Mančinska, D.E. Roberson, Nonlocal games and quantum permutation groups, *J. Funct. Anal.* 279 (5) (2020) 108592, <https://doi.org/10.1016/j.jfa.2020.108592>, <https://www.sciencedirect.com/science/article/pii/S002212362030135X>.
- [47] M. Musat, M. Rørdam, Factorizable maps and traces on the universal free product of matrix algebras, *Int. Math. Res. Not.* 2021 (23) (2021) 17951–17970, <https://doi.org/10.1093/imrn/rnz312>.
- [48] E.M. Alfsen, *Compact Convex Sets and Boundary Integrals*, *Ergebnisse der Mathematik und ihrer Grenzgebiete. 2. Folge*, Springer-Verlag, Berlin, Heidelberg, 1971.
- [49] G. Bunth, J. Pitrik, T. Titkos, D. Virosztek, Metric property of quantum Wasserstein divergences, *Phys. Rev. A* 110 (2024) 022211, <https://doi.org/10.1103/PhysRevA.110.022211>, <https://link.aps.org/doi/10.1103/PhysRevA.110.022211>.
- [50] G. De Palma, D. Trevisan, Quantum optimal transport with quantum channels, *Ann. Henri Poincaré* 22 (2021) 3199–3234, <https://doi.org/10.1007/s00023-021-01042-3>.
- [51] J. Cuntz, G.K. Pedersen, Equivalence and traces on C^* -algebras, *J. Funct. Anal.* 33 (2) (1979) 135–164, [https://doi.org/10.1016/0022-1236\(79\)90108-3](https://doi.org/10.1016/0022-1236(79)90108-3).
- [52] A. Freslon, Cut-off phenomenon for random walks on free orthogonal quantum groups, *Probab. Theory Relat. Fields* 174 (2019) 731–760, <https://doi.org/10.1007/s00440-018-0863-8>.
- [53] A. Freslon, L. Teyssier, S. Wang, Cutoff profiles for quantum Lévy processes and quantum random transpositions, *Probab. Theory Relat. Fields* 183 (2022) 1285–1327, <https://doi.org/10.1007/s00440-022-01121-4>.