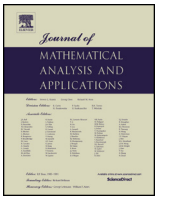
J. Math. Anal. Appl. 527 (2023) 127353

Contents lists available at ScienceDirect

Journal of Mathematical Analysis and Applications

journal homepage: www.elsevier.com/locate/jmaa



Regular Articles

Wasserstein distance between noncommutative dynamical systems

Rocco Duvenhage

Department of Physics, University of Pretoria, Pretoria 0002, South Africa

ARTICLE INFO

Article history:
Received 13 January 2022
Available online 28 April 2023
Submitted by D. Blecher

Keywords:
Optimal transport
Wasserstein distance
von Neumann algebras
States
Dynamical systems
Open systems

ABSTRACT

We introduce and study a class of quadratic Wasserstein distances on spaces consisting of generalized dynamical systems on a von Neumann algebra. We emphasize how symmetry of such a Wasserstein distance arises, but also study the asymmetric case. This setup is illustrated in the context of reduced dynamics, and a number of simple examples are also presented.

1. Introduction

This paper studies a transport plan based approach to quadratic Wasserstein distances on spaces consisting of dynamical systems on a fixed von Neumann algebra, where each system is equipped with a state invariant under the dynamics. Such distances are defined on the states, in line with the usual measure theoretic definition, but with the sets of allowed transport plans "regulated" by the dynamics. This is done by certain balance conditions imposed on the transport plans using the dynamics of two systems, which could also be viewed as covariance conditions. This indeed leads to distances on the space of dynamical systems, which can be metrics, pseudometrics or asymmetric (pseudo)metrics, depending on the assumptions.

The approach taken here builds on [25] and rests on bimodule ideas flowing from the Tomita-Takesaki theory of von Neumann algebras. We can therefore refer to it as a bimodule approach to Wasserstein distances. It is analogous to an approach taken in [8] for traces on C*-algebras in the context of free probability.

One of our motivations for studying Wasserstein distances in this setting, is to apply it to quantum detailed balance, specifically standard quantum detailed balance conditions [29,28]. In particular, to have a natural measure of how far a system is from another satisfying detailed balance. The notion of balance,

E-mail address: rocco.duvenhage@up.ac.za.

https://doi.org/10.1016/j.jmaa.2023.127353

mentioned above, was originally introduced in [26] with this goal in mind. Our eventual aim with Wasserstein distances is to make that paper's approach more precise. This line of investigation will, however, be pursued elsewhere.

The second reason to study Wasserstein distances between noncommutative dynamical systems, rather than just states, is that the mentioned balance conditions can significantly reduce the set of transport plans that need to be considered in determining the distance. In this sense it is a simplification that can ease the study of Wasserstein distances in concrete examples, which in turn can give some insight into the nature and general behaviour of Wasserstein distances in the bimodule approach. This paper indeed investigates a handful of simple examples.

Another motivation is from ideas appearing in classical ergodic theory, in particular in relation to the theory of joinings, the latter which now also has a noncommutative counterpart [22–24,4–6]. In the classical theory, these ideas originate with Ornstein's $\bar{d}$-metric [46,51], developed further in [41,48]. It seems plausible that the Wasserstein distances developed in this paper could have applications to the theory of noncommutative joinings, though this will be left for future work.

In our formalism, we allow multiple dynamics in each system, with the dynamics having minimal structure. For example, we do not assume the semigroup property, the motivation being to allow for non-Markovian dynamics. We argue in favour of this setup by illustrating the formalism in the context of reduced dynamics, namely the dynamics a system inherits from a larger system of which it is part. For reduced dynamics, semigroup properties may not hold, and it will also be seen that further dynamics, other than the dynamics of main interest, come up naturally, with balance conditions having a key role to play.

In order to obtain symmetry of the Wasserstein distances, the modular group (or modular dynamics) of the state of a system has to be included in the set of dynamics of the system, and corresponding modular balance conditions between systems imposed, also involving the KMS-duals of all the dynamics. This essentially generalizes the approach taken in [25]. One of the goals of this paper is to gain a better understanding of the role of modular dynamics and these modular balance conditions by comparing the cases with and without inclusion of the modular balance conditions. Consequently we investigate both (pseudo)metrics and asymmetric (pseudo)metrics. The inclusion of modular dynamics in addition to the dynamics of main interest, in order to attain symmetry, is also another motivating reason not to restrict the formalism to one dynamics per system.

Simple examples will be presented in Section 5 to gain further insight into the general formalism. In particular, it will be seen by example that without the above mentioned modular balance conditions, symmetry can indeed fail. The examples will also cast some light on the behaviour of Wasserstein distances in relation to the dynamics of systems, including modular dynamics. Regarding the latter, certain jumps in the value of the distance in the examples, suggest that it may be more natural to drop the modular balance conditions when we are interested in distances between the states themselves. In other words, it seems possible that asymmetric Wasserstein (pseudo)metrics on states could be more natural in the bimodule approach. This possibility will not be pursued further in this paper though. On the other hand, for essentially the same reasons, the examples also indicate that Wasserstein distances between systems should be of value in discerning qualitative differences between dynamics, while typically being insensitive to perturbations of dynamics. This could be of use in classifying systems according to certain qualitative properties, as is relevant in ideas related to detailed balance and ergodic theory.

Although the theory is developed in quite a general von Neumann algebraic setup, the examples will mostly be for low dimensional matrix cases. The reason for this is that in general it is difficult to determine the relevant sets of transport plans, but in low dimensions one can make progress on this problem by elementary means. We nevertheless study one example with an infinite dimensional algebra as well, namely the quantum (or noncommutative) torus. As mentioned earlier, the prospects to successfully calculate Wasserstein distances improve significantly as the set of allowed transport plans become smaller. Here the balance conditions come into play, making the Wasserstein distances accessible in our examples.

We pause for a moment to give a very brief bird's eye view of noncommutative optimal transport between states, in particular Wasserstein distances, which is currently growing rapidly, to provide further context and motivation for the present paper. The early work by Biane and Voiculescu [8] in the framework of free probability was already mentioned, and lead to further developments such as [39]. Other transport plan (or coupling) based approaches to Wasserstein distances, can be found in work by Ning, Georgiou and Tannenbaum [44], a series of papers by Golse, Mouhot, Paul and co-workers, starting with [37], as well as [21] by de Palma and Trevisan, though these papers in most cases did not manage to obtain metrics. Agredo and Fagnola [2] indeed pointed out pitfalls in this respect. The dynamical approach [7] to transport, from one distribution to another in the classical case, formed the basis for noncommutative versions appearing in series of papers by Carlen and Maas, as well as Chen, Georgiou and Tannenbaum, starting with [12] and [13] respectively, with infinite dimensional extensions by Hornshaw [40] and Wirth [53]. There have been other approaches as well, for example by Ikeda [42] and De Palma, Marvian, Trevisan and Lloyd [20]. Progress continues unabated, with recent work such as [16,31–33], to name but a few. There are many other papers treating further developments, but hopefully the mentioned papers and remarks give the reader an idea of the importance, scope and possibilities of noncommutative (or quantum) optimal transport. A standard introduction to the classical theory, on the other hand, is [52].

As for the structure of the paper: After fixing a few conventions and collecting some preliminary concepts in the next section, we proceed in Section 3 to define our dynamical systems and to obtain Wasserstein metrics and pseudometrics on spaces of such systems. Their asymmetric counterparts, which arise when we drop the modular balance conditions outlined above, are also discussed. It is followed by a study of reduced dynamics in this setting in Section 4, to illustrate the formalism. The paper concludes with examples in Section 5, the goal being to make a number of ideas in the paper concrete in simple cases.

2. Conventions and preliminaries

We consider a σ-finite von Neumann algebra M which is fixed throughout this section and the next. Denote the set of faithful normal states on M by $\mathfrak{F}(M)$. We denote the modular group associated with $\mu \in \mathfrak{F}(M)$ by σ^μ. By the theory of standard forms [3,17,38] (also see [9, Theorem 2.5.31]) we can assume that M is in standard form, meaning that M is a von Neumann algebra on a Hilbert space G, with every faithful normal state $\mu \in \mathfrak{F}(M)$ given by a cyclic and separating vector $\Lambda_\mu \in G$ for M, namely

$$\mu(a) = \langle \Lambda_\mu, a\Lambda_\mu \rangle$$

for all $a \in M$, which allows us to define a state $\mu' \in \mathfrak{F}(M')$ on the commutant of M by

$$\mu'(a') = \langle \Lambda_\mu, a'\Lambda_\mu \rangle$$

for all $a' \in M'$.

Given that we are going to follow a bimodule approach to Wasserstein distances, it is worth pointing out that a standard form can be viewed as a special case of the theory of correspondences (the identity in this case) in the sense of Connes [18, Appendix V.B], which are expressed as bimodules. We'll return to the latter in Section 3.

The unit of a von Neumann will be written as 1, but in cases where multiple von Neumann algebras are involved, we often indicate the algebra as a subscript, i.e., 1_M, for clarification. Define

$$j_\mu := J_\mu(\cdot)^* J_\mu$$

on $\mathcal{B}(G)$ in terms of the modular conjugation J_μ for M associated to $\Lambda_\mu \in G$. Note that $\mu' = \mu \circ j_\mu$.

We use analogous notation and conventions for another von Neumann algebra N, also in standard form on another Hilbert space H, since in Section 4 we occasionally need the duals described below for maps between two different algebras. In addition, the added generality will clarify certain points in Section 4.

We are going to make use of duals and KMS-duals of unital completely positive (u.c.p.) maps. Duals were introduced in [1] (see [26, Section 2] for a summary). KMS-duals and KMS-symmetry were studied and applied in [47,45,35,36,15,28,26].

Definition 2.1. Given a u.c.p. map $E : M \to N$ such that $\nu \circ E = \mu$ for some $\mu \in \mathfrak{F}(M)$ and $\nu \in \mathfrak{F}(N)$, we define its *KMS-dual* (w.r.t. μ and ν) as

$$E^\sigma := j_\mu \circ E' \circ j_\nu : N \to M$$

in terms of the *dual* (w.r.t. μ and ν)

$$E' : N' \to M'$$

of E defined by

$$\langle \Lambda_\mu, aE'(b')\Lambda_\mu \rangle = \langle \Lambda_\nu, E(a)b'\Lambda_\nu \rangle \tag{1}$$

for all $a \in M$ and $b' = N'$. (Consult [1] for the theory behind such duals and [26, Section 2] for a summary.)

Note that according to [1, Proposition 3.1], E' is a u.c.p. map satisfying $\mu' \circ E' = \nu'$. Correspondingly E^σ is u.c.p. and $\mu \circ E^\sigma = \nu$, while $(E^\sigma)^\sigma = E$ follows from $(E')' = E$. We also point out that under the assumptions in the definition above, the maps E, E' and E^σ are normal, i.e., σ-weakly continuous.

A special case of particular interest to us, is $M = N$ and $\mu \circ E = \mu$, where E will be viewed as dynamics leaving the state μ invariant. In such cases the dual and KMS-dual will always be with respect to μ, namely $E^\sigma := j_\mu \circ E' \circ j_\mu$, with E' defined in terms of Λ_μ on both sides of (1).

As in [25], we use the following basic notion as a starting point for optimal transport.

Definition 2.2. A *transport plan from* $\mu \in \mathfrak{F}(M)$ *to* $\nu \in \mathfrak{F}(N)$, is a state ω on the algebraic tensor product $M \odot N'$ such that

$$\omega(a \otimes 1) = \mu(a) \quad \text{and} \quad \omega(1 \otimes b') = \nu'(b')$$

for all $a \in M$ and $b' \in N'$. Denote the set of all transport plans from μ to ν by $T(\mu, \nu)$. Transport plans are also known as *couplings*.

This is a direct extension of the corresponding classical notion, discussed and motivated in [52]. But the commutant is introduced to fit into the bimodule structure to come, and plays a central role in our setup.

As a notational convention, a transport plan from μ to ν will usually be denoted by ω as in the definition above, but from ν to ξ by ψ, and from μ to ξ by φ, for any $\mu, \nu, \xi \in \mathfrak{F}(M)$.

Transport plans from μ to ν are in a one-to-one correspondence with u.c.p. maps $E : M \to N$ such that $\nu \circ E = \mu$: Let

$$\varpi_N : N \odot N' \to \mathcal{B}(H)$$

be the unital $*$-homomorphism defined by extending $\varpi_N(b \otimes b') = bb'$. Note that

$$\delta_\nu := \langle \Lambda_\nu, \varpi_N(\cdot)\Lambda_\nu \rangle$$

is a transport plan from ν to itself. Then there is a unique map

$$E_\omega : M \to N$$

such that

$$\omega(a \otimes b') = \delta_\nu(E_\omega(a) \otimes b') \tag{2}$$

for all $a \in M$ and $b' \in N'$. This map E_ω is linear, normal, u.c.p., and satisfies

$$\nu \circ E_\omega = \mu.$$

Conversely, given a u.c.p. map $E : M \to N$ such that $\nu \circ E = \mu$, it defines a transport plan ω_E from μ to ν by

$$\omega_E(a \otimes b') = \delta_\nu(E(a) \otimes b'),$$

which satisfies $E = E_{\omega_E}$. Technical details can be found in [26, Section 3]. This correspondence appears, for example, in finite dimensions in quantum information theory, where it is known as the Choi-Jamiołkowski duality (with Choi's version [14] corresponding to our setup), and also in the theory of noncommutative joinings [4]. In classical probability theory this correspondence was studied for couplings of a measure with itself in [11], and more generally in [43]. It seems not to be widely used in classical optimal transport, but the basic idea is certainly used in classical ergodic theory (see for example [34, Section 6.2]).

3. Wasserstein distances

Given these conventions and preliminaries, we proceed with only the one von Neumann algebra M in this section, to define two types of Wasserstein distances W and W_σ. The latter is a metric under suitable the conditions, the former only an asymmetric metric.

We introduce the following series of definitions, forming the foundation of our development.

Definition 3.1. A *generalized system* on M is given by $\mathbf{A} = (\alpha, \mu)$, where $\mu \in \mathfrak{F}(M)$, while α consists of the following: Let Υ be any set. To each $\upsilon \in \Upsilon$ corresponds a set Z_υ and *generalized dynamics* α_υ on M, which is given by a u.c.p. map

$$\alpha_{\upsilon,z} : M \to M$$

for every $z \in Z_\upsilon$, such that

$$\mu \circ \alpha_{\upsilon,z} = \mu$$

for all $z \in Z_\upsilon$ and $\upsilon \in \Upsilon$. We then write

$$\alpha = (\alpha_\upsilon)_{\upsilon \in \Upsilon}.$$

Such a generalized system will be referred to simply as a "system" in the sequel. We can view each Z_υ as a set of "points in time" in an abstract sense. Each α_υ is viewed as dynamics, so in effect we have a set of dynamical systems on M, indexed by υ. We do not assume any structure, for example semigroup structure, on Z_υ. In part this is because we do not need any structure, but also since in applications, we want to allow for dynamics that may not have semigroup structure, as will be seen in Section 4.

As mentioned, we allow multiple dynamics, i.e., more than one dynamical system, on M, all collected in $\mathbf{A}$ with each element of Υ corresponding to one such a dynamical system. We have to allow multiple dynamics, since in order to obtain symmetry of a Wasserstein distance, we need to include the *modular dynamics* (i.e., the modular group) in any case, while there are other natural dynamics that can also play a role, an example of which will be seen in Section 4.

We can always include the modular dynamics among the α^{υ}'s, however, to emphasize the role it plays, it will be handled separately in this section.

In the next section, systems on more than one von Neumann algebra will be involved, in which case we add the algebra to the notation for the system, i.e., $\mathbf{A} = (M, \alpha, \mu)$.

For the remainder of this section, we fix Υ and the Z_{υ}'s. The following notational convention will be used: $\mathbf{A}$ will denote (α, μ), as in the definition above, and similarly we write

$$\mathbf{B} = (\beta, \nu) \quad \text{and} \quad \mathbf{C} = (\gamma, \xi)$$

for systems on M, for the same sets Υ and the Z_{υ} used in $\mathbf{A}$. Let

$$X$$

denote the set of all such systems $\mathbf{A}$ on M. Our Wasserstein distances will be defined on X.

Definition 3.2. The *KMS-dual* of a system $\mathbf{A}$, is the system on M given by

$$\mathbf{A}^{\sigma} = (\alpha^{\sigma}, \mu)$$

where $\alpha^{\sigma} = (\alpha^{\sigma}_{\upsilon})_{\upsilon \in \Upsilon}$, and $\alpha^{\sigma}_{\upsilon}$ is given by $\alpha^{\sigma}_{\upsilon,z}$, i.e., we take the KMS-dual of each $\alpha_{\upsilon,z}$ w.r.t. μ.

To obtain our Wasserstein distances on X, we are going to use restricted sets of transport plans. To define them, we use a property which was called balance in [26]:

Definition 3.3. We say that $\mathbf{A}$ and $\mathbf{B}$ (in this order) are in *balance* with respect to $\omega \in T(\mu, \nu)$, written as

$$\mathbf{A} \omega \mathbf{B},$$

if

$$(\alpha_{\upsilon}, \mu)\, \omega\, (\beta_{\upsilon}, \nu)$$

for all $\upsilon \in \Upsilon$, by which we mean that

$$\omega(\alpha_{\upsilon,z}(a) \otimes b') = \omega(a \otimes \beta'_{\upsilon,z}(b')) \tag{3}$$

for all $a \in M$, $b' \in M'$ and $z \in Z_{\upsilon}$, in terms of the dual $\beta'_{\upsilon,z}$ defined in Definition 2.1.

Remark 3.4. Balance has an obvious extension (see [26]) to the case where $\mathbf{A}$ and $\mathbf{B}$ do not necessarily have the same von Neumann algebra, say M and N respectively. One would then write $(M, \alpha_{\upsilon}, \mu)\, \omega\, (N, \beta_{\upsilon}, \nu)$ to mean (3) for all $a \in M$, $b' \in N'$. In the present paper this extension is not technically needed, though in Section 4 the extended version will briefly be used to clarify certain points.

An important example of balance is for the modular dynamics of the states, namely

$$(\sigma^{\mu},\mu)\,\omega\,(\sigma^{\nu},\nu)\,,$$

i.e.,

$$\omega(\sigma_t^{\mu}(a)\otimes b')=\omega(a\otimes\sigma_t^{\nu'}(b'))$$

for all $a\in M$, $b'=M'$ and $t\in\mathbb{R}$. Note that here we used the fact that $(\sigma_t^{\nu})'=\sigma_t^{\nu'}$, i.e., the modular group associated with ν'.

Our restricted sets of transport plans are then defined as follows:

Definition 3.5. The set of *transport plans from* $\mathbf{A}$ *to* $\mathbf{B}$ is

$$T(\mathbf{A},\mathbf{B}):=\{\omega\in T(\mu,\nu):\mathbf{A}\omega\mathbf{B}\}\,.$$

The set of *modular* transport plans from $\mathbf{A}$ to $\mathbf{B}$ is

$$T_{\sigma}(\mathbf{A},\mathbf{B}):=\{\omega\in T(\mathbf{A},\mathbf{B}):\mathbf{A}^{\sigma}\omega\mathbf{B}^{\sigma}\text{ and }(\sigma^{\mu},\mu)\omega(\sigma^{\nu},\nu)\}\,.$$

Note that we always have $\mu\odot\nu'\in T_{\sigma}(\mathbf{A},\mathbf{B})$. The conditions $\mathbf{A}^{\sigma}\omega\mathbf{B}^{\sigma}$ and $(\sigma^{\mu},\mu)\,\omega\,(\sigma^{\nu},\nu)$ will collectively be called the *modular balance conditions*.

Remark 3.6. If the modular dynamics are included in $\mathbf{A}$ and $\mathbf{B}$, at the same index value υ, then the condition $(\sigma^{\mu},\mu)\,\omega\,(\sigma^{\nu},\nu)$ becomes redundant in $T_{\sigma}(\mathbf{A},\mathbf{B})$. For any $*$-automorphism ς of M such that $\mu\circ\varsigma=\mu$, we have $\varsigma^{\sigma}=\varsigma^{-1}$. Hence the condition $(\varsigma,\mu)\omega(\tau,\nu)$ for $*$-automorphisms ς and τ, can also be written as $(\varsigma,\mu)^{\sigma}\omega(\tau,\nu)^{\sigma}$. This holds in particular for modular dynamics. If for each pair (υ,z), it holds that $\alpha_{\upsilon,z}$ and $\beta_{\upsilon,z}$ are either both $*$-automorphisms, or both *KMS-symmetric*, i.e., $\alpha_{\upsilon,z}^{\sigma}=\alpha_{\upsilon,z}$ and $\beta_{\upsilon,z}^{\sigma}=\beta_{\upsilon,z}$, then

$$T_{\sigma}(\mathbf{A},\mathbf{B})=T(\mathbf{A},\mathbf{B})$$

if the modular dynamics are included in $\mathbf{A}$ and $\mathbf{B}$ as above.

By [26, Theorem 4.1] we can express $\mathbf{A}\omega\mathbf{B}$ as

$$E_{\omega}\circ\alpha_{\upsilon,z}=\beta_{\upsilon,z}\circ E_{\omega}\tag{4}$$

for all $z\in Z_{\upsilon}$ and $\upsilon\in\Upsilon$, in terms of E_{ω} given by (2). This formulation of balance as a covariance is often useful, as it is in joinings [4].

Definition 3.7. Given $k_1,...,k_n\in M$, and writing $k=(k_1,...,k_n)$, the associated *transport cost function* I_k, which gives the cost of transport $I_k(\omega)$ from $\mu\in\mathfrak{F}(M)$ to $\nu\in\mathfrak{F}(M)$ for the transport plan $\omega\in T(\mu,\nu)$, is defined to be

$$I_k(\omega)=\sum_{l=1}^{n}\left[\mu(k_l^*k_l)+\nu(k_l^*k_l)-\nu(E_{\omega}(k_l)^*k_l)-\nu(k_l^*E_{\omega}(k_l))\right].\tag{5}$$

This formulation uses the ideas we have set up so far. An equivalent, but more suggestive formulation, in terms of the cyclic representations $(H_{\mu}^{\omega},\pi_{\mu}^{\omega},\Omega)$ and $(H_{\nu}^{\omega},\pi_{\nu}^{\omega},\Omega)$ of (M,μ) and (M,ν), respectively, inherited from the cyclic representation $(H_{\omega},\pi_{\omega},\Omega)$ of $(M\odot M',\omega)$, is

$$I_k(\omega) = \left\|\pi_\mu^\omega(k)\Omega - \pi_\nu^\omega(k)\Omega\right\|_{\oplus\omega}^2,$$

where we have written

$$\pi_\mu^\omega(k)\Omega \equiv \left(\pi_\mu^\omega(k_1)\Omega, ..., \pi_\mu^\omega(k_n)\Omega\right) \in \bigoplus_{l=1}^n H_\omega,$$

while $\|\cdot\|_{\oplus\omega}$ denotes the norm on $\bigoplus_{l=1}^n H_\omega$. I.e.,

$$I_k(\omega) = \sum_{l=1}^n \left\|\pi_\mu^\omega(k_l)\Omega - \pi_\nu^\omega(k_l)\Omega\right\|_\omega^2.$$

More detail regarding these representations can be found in [25]. From this formulation it is clear that $I_k(\omega) \geq 0$, but it can also be seen directly from (5): For all $a, b \in M$,

$$\begin{aligned}
&\mu(a^*a) + \nu(b^*b) - \nu(E_\omega(a)^*b) - \nu(b^*E_\omega(a)) \\
&= \nu\left(|b - E_\omega(a)|^2\right) + \mu\left(|a|^2\right) - \nu\left(|E_\omega(a)|^2\right) \\
&\geq \nu\left(|b - E_\omega(a)|^2\right) \\
&\geq 0
\end{aligned}$$

by Kadison's inequality.

In what follows, by a *distance function* on X, we simply mean a function $d : X \times X \to \mathbb{R}$. Such a function may be a metric. However, we are also interested in whether a distance function d is a *pseudometric*, which means that it satisfies the triangle inequality, is symmetric, and obeys $d(x, x) = 0$ and $d(x, y) \geq 0$, but could allow $d(x, y) = 0$ for $x \neq y$. Similarly for the asymmetric cases, to which we return at the end of this section.

The specific distance functions we study in this paper, will collectively be called *Wasserstein distances*. They are the distance functions W and W_σ on the space X of systems, given by our main definition below.

Definition 3.8. Given $k_1, ..., k_n \in A$, we define the associated *Wasserstein distance* W on X by

$$W(\mathbf{A}, \mathbf{B}) := \inf_{\omega \in T(\mathbf{A},\mathbf{B})} I_k(\omega)^{1/2},$$

and the associated *modular Wasserstein distance* W_σ on X by

$$W_\sigma(\mathbf{A}, \mathbf{B}) := \inf_{\omega \in T_\sigma(\mathbf{A},\mathbf{B})} I_k(\omega)^{1/2},$$

for all $\mathbf{A}, \mathbf{B} \in X$, in terms of Definition 3.7.

These are also called Wasserstein distances of order 2, or quadratic Wasserstein distances. Since we focus exclusively on the quadratic case in this paper, we do not include a subscript 2 as is standard notation in the classical case. More complete notation would be to include the k, say as $W^{(k)}$ and $W_\sigma^{(k)}$, but no confusion should arise.

Clearly

$$W(\mathbf{A}, \mathbf{B}) \leq W_\sigma(\mathbf{A}, \mathbf{B}) \tag{6}$$

for all $\mathbf{A},\mathbf{B}\in X$, while in specific examples it should be easier to determine $W_\sigma(\mathbf{A},\mathbf{B})$, as the additional balance conditions simplify finding the relevant (and smaller) set of transport plans $T_\sigma(\mathbf{A},\mathbf{B})$.

The central results of this section are the following two theorems, regarding the metric properties of W_σ and W respectively, with W_σ enjoying symmetry, but W not. The triangle inequality emerges from the natural M-M-bimodule structure of the GNS Hilbert spaces H_ω of the transport plans ω, and their relative tensor products $H_\omega\otimes_\nu H_\psi$. These products are treated in [50, Section IX.3], but see also [30] and the early work [49]. They go hand in hand with Connes' largely unpublished work on correspondences, though [18, Appendix V.B] gives a partial exposition.

Theorem 3.9. *Let W_σ be the modular Wasserstein distance on X associated to $k_1,...,k_n\in M$.*

(a) Then W_σ is a pseudometric and its value $W_\sigma(\mathbf{A},\mathbf{B})$ is reached by some transport plan $\omega\in T_\sigma(\mathbf{A},\mathbf{B})$, i.e., optimal modular transport plans always exist.

(b) If in addition we assume that $\{k_1^*,...,k_n^*\}=\{k_1,...,k_n\}$ and that M is generated by $k_1,...,k_n$, then W_σ is a metric.

Proof. (a) By its definition, W_σ is real-valued and never negative. Also note that $W_\sigma(\mathbf{A},\mathbf{A})=0$ from (5) with $\omega=\delta_\mu$, which is an element of $T_\sigma(\mathbf{A},\mathbf{A})$, since $E_\omega=\mathrm{id}_M$, trivially delivering all the balance requirements.

The triangle inequality. For $\omega\in T_\sigma(\mathbf{A},\mathbf{B})$ and $\psi\in T_\sigma(\mathbf{B},\mathbf{C})$, we set $\varphi=\omega\circ\psi$, defined via

$$E_{\omega\circ\psi}=E_\psi\circ E_\omega.$$

Note that $\varphi\in T_\sigma(\mathbf{A},\mathbf{C})$, since

$$E_\varphi\circ\alpha_{\upsilon,z}=E_\psi\circ E_\omega\circ\alpha_{\upsilon,z}=E_\psi\circ\beta_{\upsilon,z}\circ E_\omega=\gamma_{\upsilon,z}\circ E_\psi\circ E_\omega=\gamma_{\upsilon,z}\circ E_\varphi,$$

hence $\mathbf{A}\varphi\mathbf{C}$, while similarly $\mathbf{A}^\sigma\varphi\mathbf{C}^\sigma$ and $(\sigma^\mu,\mu)\varphi\left(\sigma^\xi,\xi\right)$. As in the proof of [25, Proposition 4.3] we have

$$I_k(\varphi)^{1/2}\leq I_k(\omega)^{1/2}+I_k(\psi)^{1/2},$$

by employing the triangle inequality in $\bigoplus_{l=1}^n(H_\omega\otimes_\nu H_\psi)$, and using properties of the relative tensor product of the M-M-bimodules H_ω and H_ψ (the required continuity properties making them M-M-bimodules, were shown to hold in [4, Theorem 3.3]). Now take the infimum on the left over all of $T_\sigma(\mathbf{A},\mathbf{C})$, which includes the compositions $\omega\circ\psi$ for all $\omega\in T_\sigma(\mathbf{A},\mathbf{B})$ and $\psi\in T_\sigma(\mathbf{B},\mathbf{C})$, followed in turn by the infima over all $\omega\in T_\sigma(\mathbf{A},\mathbf{B})$ and $\psi\in T_\sigma(\mathbf{B},\mathbf{C})$ on the right.

Symmetry. Recall from [25, Lemma 5.2] that a u.c.p. map $E:M\to M$ satisfying $\nu\circ E=\mu$, has a Hilbert space representation as a contraction $U:G\to G$ defined through $Ua\Lambda_\mu=E(a)\Lambda_\nu$ for all $a\in M$, such that the following equivalence holds:

$$\mu(aE^\sigma(b))=\nu(E(a)b)$$

for all $a,b\in M$, if and only if

$$J_\nu U=UJ_\mu.$$

We now apply this to E_ω and its Hilbert space representation U_ω, for $\omega\in T_\sigma(\mathbf{A},\mathbf{B})$. Since ω satisfies $(\sigma^\mu,\mu)\,\omega\,(\sigma^\nu,\nu)$, it follows from [26, Theorem 4.1] that $\Delta_\nu^{it}U_\omega=U_\omega\Delta_\mu^{it}$, where Δ_μ and Δ_ν are the modular operators associated to Λ_μ and Λ_ν respectively. Consequently $J_\nu U_\omega=U_\omega J_\mu$. It therefore follows that

$$\begin{aligned}&\mu(a^*a)+\nu(b^*b)-\nu(E_\omega(a)^*b)-\nu(b^*E_\omega(a))\\&=\nu(b^*b)+\mu(a^*a)-\mu(E_\omega^\sigma(b)^*a)-\mu(a^*E_\omega^\sigma(b))\\&=\nu(b^*b)+\mu(a^*a)-\mu(E_{\omega^\sigma}(b)^*a)-\mu(a^*E_{\omega^\sigma}(b))\end{aligned}$$

where $\omega^\sigma\in T(\nu,\mu)$ is determined by $E_{\omega^\sigma}=E_\omega^\sigma$ according to [26, Section 4]. Note that $\omega^\sigma\in T_\sigma(\mathbf{B},\mathbf{A})$, since $\mathbf{A}^\sigma\omega\mathbf{B}^\sigma$ is equivalent to $\mathbf{B}\omega^\sigma\mathbf{A}$, and $\mathbf{A}\omega\mathbf{B}$ to $\mathbf{B}^\sigma\omega^\sigma\mathbf{A}^\sigma$, according to [26, Corollary 4.6], while $(\sigma^\nu,\nu)\,\omega^\sigma\,(\sigma^\mu,\mu)$, since $E_{\omega^\sigma}\circ\sigma_t^\nu=(\sigma_{-t}^\nu\circ E_\omega)^\sigma=(E_\omega\circ\sigma_{-t}^\mu)^\sigma=\sigma_t^\mu\circ E_{\omega^\sigma}$, due to $(\sigma_t^\mu)^\sigma=\sigma_{-t}^\mu$.

The required symmetry $W_\sigma(\mathbf{A},\mathbf{B})=W_\sigma(\mathbf{B},\mathbf{A})$ now follows from (5) and Definition 3.8 of W_σ, since for each ω, every term in $I_k(\omega)$ is equal to the corresponding term in $I_k(\omega^\sigma)$, while $(\omega^\sigma)^\sigma=\omega$ because of $(E_\omega^\sigma)^\sigma=E_\omega$, giving a one-to-one correspondence between $T_\sigma(\mathbf{A},\mathbf{B})$ and $T_\sigma(\mathbf{B},\mathbf{A})$, which means we retain equality in the infima over $T_\sigma(\mathbf{A},\mathbf{B})$ and $T_\sigma(\mathbf{B},\mathbf{A})$ respectively.

Optimal transport plans exist. We know (see for example [23, Proposition 4.1]) that without loss we can view each element of $T_\sigma(\mathbf{A},\mathbf{B})$ as a state on the maximal C*-tensor product $A\otimes_{\max}B'$, from which it follows that $T_\sigma(\mathbf{A},\mathbf{B})$ is weakly* compact. By the definition of W_σ there is a sequence $\omega_q\in T_\sigma(\mathbf{A},\mathbf{B})$ such that $I_k(\omega_q)^{1/2}\to W_\sigma(\mathbf{A},\mathbf{B})$, which necessarily has a weak* cluster point $\omega\in T_\sigma(\mathbf{A},\mathbf{B})$. The existence of an optimal transport plan is now obtained by the same approximation argument as for [25, Lemma 6.2], namely $I_k(\omega)^{1/2}=W_\sigma(\mathbf{A},\mathbf{B})$.

(b) If $W_\sigma(\mu,\nu)=0$, then $\mu=\nu$ follows from the existence of an optimal transport plan combined with [25, Corollary 6.4]. □

Dropping the modular balance conditions, the same proof, with minor modifications (mostly simplifications) also delivers the corresponding result for W below, but without symmetry. We say that $d:X\times X\to[0,\infty)$ is an *asymmetric pseudometric*, if it satisfies the triangle inequality and $d(x,x)=0$. If in addition $d(x,y)=0$ implies that $x=y$, then we call d an *asymmetric metric*. The point being that we do not assume symmetry, $d(x,y)=d(y,x)$, in these definitions.

Theorem 3.10. *Let W be the Wasserstein distance on X associated to $k_1,...,k_n\in M$.*

(a) Then W is an asymmetric pseudometric and its value $W(\mathbf{A},\mathbf{B})$ is reached by some transport plan $\omega\in T(\mathbf{A},\mathbf{B})$, i.e., optimal transport plans always exist.

(b) If in addition we assume that $\{k_1^*,...,k_n^*\}=\{k_1,...,k_n\}$ and that M is generated by $k_1,...,k_n$, then W is an asymmetric metric.

In Section 5 it will indeed be seen by simple example that in general W does not possess symmetry.

We study both the *modular Wasserstein pseudometric* W_σ and the *asymmetric Wasserstein pseudometric* W in the sequel. It tends to be easier to prove results for W, as we do not need to take care of the modular balance conditions. Consequently, as above, we provide more detailed arguments regarding W_σ in the next section.

4. Reduced systems

The goal of this section is to illustrate that the formulation of Wasserstein distances between systems in the previous section, in terms of balance conditions, is natural. We do this in the context of reduced dynamics, where, loosely speaking, we study the dynamics of a system $\mathbf{S}$, which is interacting with another, called the *reservoir*. The interacting reservoir-system combination, will be called a *composite system* $\mathbf{A}$, with its dynamics due to the interaction called its *evolution*. In such cases $\mathbf{S}$ is referred to as an *open system*, but in this section we also call it a *reduced system*, as its dynamics is "reduced" from that of the composite system. We aim to determine how Wasserstein distances on the set of reduced systems relate to Wasserstein distances on the set of composite systems.

The literature on open systems is vast. Standard textbook treatments can be found in [10,19], for example. However, we presuppose as little as possible background from this field, in order to make this section accessible using only the framework set up so far.

We assume the reservoir and **S** to have fixed von Neumann algebras R and S respectively, but allow different states and dynamics to obtain a set of composite systems and a set of reduced systems (we do not need to define the reservoirs as separate systems). We then make use of a transport cost function (Definition 3.7) for which the k_l's are in $1_R \otimes S$, giving Wasserstein distances on the set Y of reduced systems. In this context we can also obtain natural Wasserstein distances on the set of composite systems via a reduction. This reduction is performed by including an appropriate conditional expectation of $R \otimes S$ onto $1_R \otimes S$ as dynamics (a *channel*) in the composite system.

In general the reduced dynamics does not satisfy a semigroup property, even if that of the composite system does. In this section we focus on the case where the evolution of **A** does have the semigroup property. The reason for this is solely to emphasize that even then the semigroup property in general does not hold for **S**.

If we include the modular dynamics, then both our composite system and reduced system will have $\Upsilon = \{1, 2\}$, with a semigroup $Z_1 = \Gamma$ for the evolution, and $Z_2 = \mathbb{R}$ for the modular dynamics. However, we also consider an augmented form of a composite system, with the dynamics supplemented by the conditional expectation mentioned above. In this case one has $\Upsilon = \{1, 2, 3\}$, with the same Z_1 and Z_2 as for the composite systems, and a one-point set Z_3 for the conditional expectation.

In this way, the present section demonstrates why the general setup of the previous section does not assume semigroup (or any other) structure on the sets Z_υ, and why it allows multiple dynamics (indexed by $\upsilon \in \Upsilon$). This section also emphasizes the balance conditions, in this case in particular for the conditional expectation, to illustrate that they play a natural role.

Remark 4.1. In this section we formulate everything with a view towards W_σ, that is, with the modular dynamics and the required balance conditions ensuring symmetry included. The same outline holds for W, but with simpler and shorter arguments, hence we do not present them separately. Short remarks at the end of this section will suffice.

To make the discussion above more precise, the setup in this section is as follows:
We set

$$M = R \bar{\otimes} S,$$

where R and S are both σ-finite von Neumann algebras in standard form, fixed throughout this section, and $\bar{\otimes}$ denotes the von Neumann algebraic tensor product. In this section we only consider product states

$$\mu = \mu_R \bar{\otimes} \mu_S$$

on M, where $\mu_R \in \mathfrak{F}(R)$ and $\mu_S \in \mathfrak{F}(S)$. We also fix a semigroup Γ to be used as set out below, for all systems appearing in this section.

In the previous section, the notation (α, μ) was used for a system, however, since there are several algebras involved in this section, we include the algebra as well in the notation for the system, namely (M, α, μ). We are interested in *composite systems* of the form

$$\mathbf{A} = (M, \alpha, \sigma^\mu, \mu) = \left(M, \alpha, \sigma^{\mu_R \bar{\otimes} \mu_S}, \mu_R \bar{\otimes} \mu_S\right), \tag{7}$$

but will also require their *augmented* form, given by

$$\mathbf{A}^{\mathfrak{p}} = \left(M, \alpha, \sigma^{\mu}, P_{1\otimes S}^{\mu_R}, \mu\right),$$

where:

α is a semigroup of u.c.p. maps $\alpha_g : M \to M$ such that $\mu \circ \alpha_g = \mu$ for all $g \in \Gamma$. Here the term semigroup means that $\alpha_{gh} = \alpha_g \circ \alpha_h$ for all $g, h \in \Gamma$. We call α the *evolution* of $\mathbf{A}$.

$P_{1\otimes S}^{\mu_R} : M \to M$ is the conditional expectation onto $1_R \otimes S$ defined by

$$P_{1\otimes S}^{\mu_R} = (\mu_R 1_R)\,\bar{\otimes}\,\mathrm{id}_S,$$

where 1_R is the unit of R. In particular, $P_{1\otimes S}^{\mu_R}(r \otimes s) = \mu_R(r)1_R \otimes s$ for any elementary tensor $r \otimes s$ in M.

Let

$$X_{\otimes}$$

be the space of all composite systems on M of the form in (7), as described above. Similarly,

$$X_{\otimes}^{\mathfrak{p}}$$

will denote the space of augmented systems as defined above.

Note that we explicitly include the modular dynamics in the systems to ensure symmetry of the Wasserstein distances, as seen in Section 3.

Remark 4.2. A standard physical situation is where $\Gamma = \mathbb{R}$, with α in addition being a one-parameter group of $*$-automorphisms, i.e., α_t is a $*$-automorphism for every $t \in \mathbb{R}$ and $\alpha_0 = \mathrm{id}_M$. On the other hand, one could in fact use more abstract generalized dynamics as in Section 3 instead of a semigroup α, since the semigroup property will not be used explicitly. But to clarify how the semigroup property can be spoiled by reduction, we use the setting where α is a semigroup.

Remark 4.3. The reason for introducing the augmented systems, is that balance between the conditional expectations $P_{1\otimes S}^{\mu_R}$ will play an essential role in connecting Wasserstein distances on the composite systems, to those on the reduced systems discussed below.

The *reduced system* on S, or the *reduction* of $\mathbf{A}$ to S, is defined to be the system

$$\mathbf{A}^{\mathfrak{r}} = (S, \alpha^{\mathfrak{r}}, \sigma^{\mu_S}, \mu_S),$$

where the reduced dynamics $\alpha^{\mathfrak{r}} : S \to S$ is given by

$$\alpha_g^{\mathfrak{r}} := P_S^{\mu_R} \circ \alpha_g \circ \iota_{S,M},$$

for all $g \in \Gamma$, in terms of

$$P_S^{\mu_R} = \mu_R \bar{\otimes}\,\mathrm{id}_S : M \to S$$

and

$$\iota_{S,M} : S \to M : s \mapsto 1_R \otimes s.$$

Note that indeed

$$\mu_S \circ \alpha_g^{\mathfrak{r}} = \mu \circ \alpha_g \circ \iota_{S,M} = \mu \circ \iota_{S,M} = \mu_S,$$

as is required of a system. One of the main points here, is that $\alpha^{\mathfrak{r}}$ clearly need not be a semigroup, despite the fact that α is, since α need not be product dynamics (due to interaction). This reduced system $\mathbf{A}^{\mathfrak{r}}$ is what we referred to as $\mathbf{S}$ in the preliminary discussion at the beginning of the section.

The set of reduced systems obtained in this way, will be denoted by

$$Y.$$

The reduced dynamics obtained as part of these reduced systems, can be viewed as a special class of reduced dynamics, as we assume that they are reduced from dynamics leaving $\mu_R \bar{\otimes} \mu_S$ invariant. This is not the most general situation one could consider, but this class is what fits into our framework for Wasserstein distances.

Although we are eventually interested only in systems on M, a key point regarding balance (see Proposition 4.5 below), and its proof, is clarified when generalized, by also considering analogous systems on a second von Neumann algebra N. So for σ-finite von Neumann algebras K and L in standard form, we set

$$N := K \bar{\otimes} L,$$
$$\mathbf{B} = (N, \beta, \sigma^{\nu}, \nu),$$

and

$$\mathbf{B}^{\mathfrak{p}} = \left(N, \beta, \sigma^{\nu}, P_{1\otimes L}^{\nu_K}, \nu\right),$$

where β is a semigroup of u.c.p. maps leaving

$$\nu = \nu_K \bar{\otimes} \nu_L \in \mathfrak{F}(N),$$

invariant, with $\nu_K \in \mathfrak{F}(K)$ and $\nu_L \in \mathfrak{F}(L)$, and the same sets Z_1, Z_2 and Z_3 as before. As mentioned in Remark 3.4, the notion of balance $\mathbf{A}\omega\mathbf{B}$ still makes sense.

In this setup, balance between the augmented composite systems, carries over to the reduced systems, as will be seen in Proposition 4.5 below, but first a lemma to get there:

Lemma 4.4. *Given $\omega \in T(\mu, \nu)$, with $\mu = \mu_R \bar{\otimes} \mu_S$ and $\nu = \nu_K \bar{\otimes} \nu_L$ as above, we set*

$$\omega^{\mathfrak{r}} := \omega \circ (\iota_{S,M} \odot \iota_{L',N'}) \in T(\mu_S, \nu_L).$$

Assuming the balance condition

$$\left(M, P_{1\otimes S}^{\mu_R}, \mu\right) \omega \left(N, P_{1\otimes L}^{\nu_K}, \nu\right),$$

it follows that

$$E_{\omega^{\mathfrak{r}}} = P_L^{\nu_K} \circ E_{\omega} \circ \iota_{S,M}.$$

Proof. From $\omega \in T(\mu, \nu)$ it is easily checked that $\omega^{\mathfrak{r}} \in T(\mu_S, \nu_L)$. Since, according to (2), $E_{\omega^{\mathfrak{r}}} : S \to L$ is determined by $\omega^{\mathfrak{r}} = \delta_{\nu_L} \circ (E_{\omega^{\mathfrak{r}}} \odot \mathrm{id}_{L'})$, we can verify this lemma as follows using (4): for all $s \in S$ and $l' \in L'$,

$$\begin{aligned}\omega^{\mathfrak{r}}(s\otimes l') &= \omega\left(\iota_{S,M}(s)\otimes\iota_{L',N'}(l')\right)\\ &= \delta_\nu\left(E_\omega(1_R\otimes s)\otimes\iota_{L',N'}(l')\right)\\ &= \delta_\nu\left(E_\omega\circ P^{\mu_R}_{1\otimes S}(1_R\otimes s)\otimes\iota_{L',N'}(l')\right)\\ &= \delta_\nu\left(P^{\nu_K}_{1\otimes L}\circ E_\omega(\iota_{S,M}(s))\otimes\iota_{L',N'}(l')\right)\\ &= \delta_{\nu_K}\odot\delta_{\nu_L}\left((1_K\otimes P^{\nu_K}_L\circ E_\omega\circ\iota_{S,M}(s))\otimes(1_{K'}\otimes l')\right)\\ &= \delta_{\nu_L}\left(P^{\nu_K}_L\circ E_\omega\circ\iota_{S,M}(s)\otimes l'\right),\end{aligned}$$

as required. □

I.e., the u.c.p. map $E_{\omega^{\mathfrak{r}}}$ corresponding to the *reduced* transport plan $\omega^{\mathfrak{r}}$, is given by the *reduction* of E_ω. Now, as promised, balance between the augmented composite systems, carries over to the reduced systems:

Proposition 4.5. *If* $\mathbf{A}^{\mathfrak{p}}\omega\mathbf{B}^{\mathfrak{p}}$*, then* $\mathbf{A}^{\mathfrak{r}}\omega^{\mathfrak{r}}\mathbf{B}^{\mathfrak{r}}$.

Proof. Using Lemma 4.4, the given balance conditions, as well as elementary relations of the form $\iota_{L,N}\circ P^{\nu_K}_L = P^{\nu_K}_{1\otimes L}$, $P^{\nu_K}_L\circ P^{\nu_K}_{1\otimes L} = P^{\nu_K}_L$ and $P^{\mu_R}_{1\otimes S}\circ\iota_{S,M} = \iota_{S,M}$, it follows that

$$\begin{aligned}\beta^{\mathfrak{r}}_g\circ E_{\omega^{\mathfrak{r}}} &= P^{\nu_K}_L\circ\beta_g\circ\iota_{L,N}\circ P^{\nu_K}_L\circ E_\omega\circ\iota_{S,M}\\ &= P^{\nu_K}_L\circ P^{\nu_K}_{1\otimes L}\circ\beta_g\circ P^{\nu_K}_{1\otimes L}\circ E_\omega\circ\iota_{S,M}\\ &= P^{\nu_K}_L\circ E_\omega\circ P^{\mu_R}_{1\otimes S}\circ\alpha_g\circ P^{\mu_R}_{1\otimes S}\circ\iota_{S,M}\\ &= P^{\nu_K}_L\circ E_\omega\circ\iota_{S,M}\circ P^{\mu_R}_S\circ\alpha_g\circ\iota_{S,M}\\ &= E_{\omega^{\mathfrak{r}}}\circ\alpha^{\mathfrak{r}}_g.\end{aligned}$$

Furthermore, since $\sigma^\mu_t = \sigma^{\mu_R}_t\bar{\otimes}\sigma^{\mu_S}_t$ and $\sigma^\nu_t = \sigma^{\nu_K}_t\bar{\otimes}\sigma^{\nu_L}_t$,

$$\begin{aligned}\sigma^{\nu_L}_t\circ E_{\omega^{\mathfrak{r}}} &= P^{\nu_K}_L\circ\sigma^\nu_t\circ E_\omega\circ\iota_{S,M}\\ &= P^{\nu_K}_L\circ E_\omega\circ\sigma^\mu_t\circ\iota_{S,M}\\ &= P^{\nu_K}_L\circ E_\omega\circ\iota_{S,M}\circ\sigma^{\mu_S}_t\\ &= E_{\omega^{\mathfrak{r}}}\circ\sigma^{\mu_S}_t,\end{aligned}$$

which completes the proof by (4). □

In the remainder of the section, we proceed with our main interest, namely the case

$$M = N,\ R = K \text{ and } S = L.$$

When using $T_\sigma(\mathbf{A},\mathbf{B})$ in Definition 3.5, we need to apply Proposition 4.5 to the KMS-duals of the systems as well, which leads us to the next proposition.

Proposition 4.6. *For every* $\mathbf{A}\in X_\otimes$*, we have* $(\mathbf{A}^{\mathfrak{p}})^\sigma = (\mathbf{A}^\sigma)^{\mathfrak{p}}$ *and* $(\mathbf{A}^{\mathfrak{r}})^\sigma = (\mathbf{A}^\sigma)^{\mathfrak{r}}$.

Proof. To show $(\mathbf{A}^{\mathfrak{p}})^\sigma = (\mathbf{A}^\sigma)^{\mathfrak{p}}$, we clearly just have to check that $\left(P^{\mu_R}_{1\otimes S}\right)^\sigma = P^{\mu_R}_{1\otimes S}$, i.e., that $P^{\mu_R}_{1\otimes S}$ is KMS-symmetric (w.r.t. μ), since the remaining conditions are trivially satisfied. First note that according to Definition 2.1,

$$\left(P^{\mu_R}_{1\otimes S}\right)' = P^{\mu'_R}_{1\otimes S'},$$

since for all $r \in R$, $r' \in R'$, $s \in S$ and $s' \in S'$,

$$\begin{aligned}
&\langle \Lambda_\mu, (r \otimes s)\left((\mu'_R 1_{R'}) \otimes \mathrm{id}_{S'}(r' \otimes s')\right)\Lambda_\mu\rangle \\
&= \mu_R(r)\mu'_R(r')\,\langle \Lambda_{\mu_S}, ss'\Lambda_{\mu_S}\rangle \\
&= \langle \Lambda_\mu, \left((\mu_R 1_R) \otimes \mathrm{id}_S(r \otimes s)\right)(r' \otimes s')\Lambda_\mu\rangle\,,
\end{aligned}$$

simply because $\Lambda_\mu = \Lambda_{\mu_R} \otimes \Lambda_{\mu_S}$. We therefore have

$$\begin{aligned}
\left(P^{\mu_R}_{1\bar{\otimes}S}\right)^\sigma &= j_\mu \circ P^{\mu'_R}_{1\bar{\otimes}S'} \circ j_\mu \\
&= (j_{\mu_R}\bar{\otimes} j_{\mu_S}) \circ ((\mu'_R 1_{R'})\,\bar{\otimes}\,\mathrm{id}_{S'}) \circ (j_{\mu_R}\bar{\otimes} j_{\mu_S}) \\
&= \left[(\mu'_R \circ j_{\mu_R})\, j_{\mu_R}(1_{R'})\right]\bar{\otimes}\,(j_{\mu_S} \circ \mathrm{id}_{S'} \circ j_{\mu_S}) \\
&= (\mu_R 1_R)\,\bar{\otimes}\,\mathrm{id}_S \\
&= P^{\mu_R}_{1\bar{\otimes}S}.
\end{aligned}$$

Next we show $(\mathbf{A}^{\mathfrak{r}})^\sigma = (\mathbf{A}^\sigma)^{\mathfrak{r}}$. In this case the only non-trivial part is to check that $(\alpha^{\mathfrak{r}})^\sigma = (\alpha^\sigma)^{\mathfrak{r}}$. We have

$$(\alpha^{\mathfrak{r}})^\sigma = (P^{\mu_R}_S \circ \alpha_g \circ \iota_{S,M})^\sigma = \iota^\sigma_{S,M} \circ \alpha^\sigma_g \circ (P^{\mu_R}_S)^\sigma$$

and

$$(\alpha^\sigma)^{\mathfrak{r}} = P^{\mu_R}_S \circ \alpha^\sigma_g \circ \iota_{S,M}.$$

To determine $\iota^\sigma_{S,A} : M \to S$, we calculate

$$\begin{aligned}
\langle \Lambda_\mu, \iota_{S,M}(s)\,(r' \otimes s')\,\Lambda_\mu\rangle &= \langle \Lambda_\mu, (r' \otimes (ss'))\,\Lambda_\mu\rangle \\
&= \langle \Lambda_{\mu_R}, r'\Lambda_{\mu_R}\rangle\,\langle \Lambda_{\mu_S}, ss'\Lambda_{\mu_S}\rangle \\
&= \langle \Lambda_{\mu_S}, s\mu'_R(r')s'\Lambda_{\mu_S}\rangle \\
&= \langle \Lambda_{\mu_S}, s\mu'_R\bar{\otimes}\,\mathrm{id}_{S'}(r' \otimes s')\Lambda_{\mu_S}\rangle\,,
\end{aligned}$$

which by σ-weak continuity of $\mu'_R\bar{\otimes}\,\mathrm{id}_{S'}$ (being the tensor product of normal maps), along with Definition 2.1, means that

$$\iota'_{S,M} = \mu'_R\bar{\otimes}\,\mathrm{id}_{S'} = P^{\mu'_R}_{S'} : M' \to S',$$

which in turns tells us that

$$\iota^\sigma_{S,M} = j_{\mu_s} \circ P^{\mu'_R}_{S'} \circ j_\mu = P^{\mu_R}_S,$$

similar to the calculation for $\left(P^{\mu_R}_{1\bar{\otimes}S}\right)^\sigma$ above. Thus, $(P^{\mu_R}_S)^\sigma = \left(\iota^\sigma_{S,M}\right)^\sigma = \iota_{S,M}$, hence $(\alpha^{\mathfrak{r}})^\sigma = (\alpha^\sigma)^{\mathfrak{r}}$. □

The next corollary, which now follows from Remark 3.6, is not needed in this section, but becomes relevant in the next. It implies that under the given conditions, the modular balance conditions for reduced systems, reduce to balance between the modular dynamics.

Corollary 4.7. *If Γ is a group and α a group representation as $*$-automorphisms of M, then $\left(\alpha^{\mathfrak{r}}_g\right)^\sigma = \alpha^{\mathfrak{r}}_{g^{-1}}$ for all $g \in \Gamma$.*

To connect the transport cost function for reduced systems to the cost for composite systems, we need the following:

Lemma 4.8. *Consider $\mu = \mu_R \bar{\otimes} \mu_S, \nu = \nu_R \bar{\otimes} \nu_S \in \mathfrak{F}(M)$ as before. For $\omega \in T(\mu, \nu)$, and assuming the balance condition*

$$\left(M, P_{1\otimes S}^{\mu_R}, \mu\right) \omega \left(M, P_{1\otimes S}^{\nu_R}, \nu\right),$$

it follows that

$$\nu\left(k^* E_\omega(k)\right) = \nu_S\left(w^* E_{\omega^{\mathfrak{r}}}(w)\right)$$

for any $w \in S$ and $k := 1 \otimes w \in M$.

Proof. Note that for $r \in R$ and $s_1, s_2 \in S$, one has

$$P_S^{\mu_R}\left((1 \otimes s_1)(r \otimes s_2)\right) = \mu_R(r) s_1 s_2 = s_1 P_S^{\mu_R}\left(r \otimes s_2\right).$$

Since $P_{1\otimes S}^{\mu_R}$ is normal (it is the tensor product of normal maps) and multiplication of operators is σ-weakly continuous in each factor separately, we know that $P_S^{\mu_R}\left((1 \otimes s)(\cdot)\right) : M \to S$ is normal. Hence

$$P_S^{\mu_R}\left(\iota_{S,M}(s) a\right) = s P_S^{\mu_R}(a)$$

for all $a \in M$ and $s \in S$. Consequently

$$\nu\left(\iota_{S,M}(s) a\right) = \nu_S \circ P_S^{\nu_R}\left(\iota_{S,M}(s) a\right) = \nu_S\left(s P_S^{\nu_R}(a)\right).$$

In particular, because of the balance assumption,

$$\nu\left(\iota_{S,M}(s) E_\omega \circ \iota_{S,M}(w)\right) = \nu_S\left(s P_S^{\nu_R} \circ E_\omega \circ \iota_{S,M}(w)\right) = \nu_S\left(s E_{\omega^{\mathfrak{r}}}(w)\right)$$

by Lemma 4.4, which for the special case $s = w^*$ gives $\nu\left(k^* E_\omega(k)\right) = \nu_S\left(w^* E_{\omega^{\mathfrak{r}}}(w)\right)$. □

The ideas of this section can now be brought together by the next definition and result, which show how Wasserstein distances on the set Y of reduced systems relate to Wasserstein distances on the set $X_\otimes$ of composite systems. Note that in the following definition we restrict the modular Wasserstein pseudometrics given by Theorem 3.9, to $X_\otimes$ and $X_\otimes^{\mathfrak{p}}$ respectively.

Definition 4.9. In terms of the modular Wasserstein pseudometric W_σ on $X_\otimes^{\mathfrak{p}}$ associated to given $k_1, ..., k_n \in M$, the corresponding *reduced* modular Wasserstein pseudometric $W_\sigma^{\mathfrak{r}}$ on $X_\otimes$, associated to $k_1, ..., k_n$, is defined by

$$W_\sigma^{\mathfrak{r}}(\mathbf{A}, \mathbf{B}) = W_\sigma\left(\mathbf{A}^{\mathfrak{p}}, \mathbf{B}^{\mathfrak{p}}\right)$$

for all $\mathbf{A}, \mathbf{B} \in X_\otimes$.

Theorem 4.10. *Consider the modular Wasserstein pseudometric W_σ on Y associated to $w_1, ..., w_n \in S$, and the reduced modular Wasserstein pseudometric $W_\sigma^{\mathfrak{r}}$ on $X_\otimes$ associated to $k_1, ..., k_n \in M$ given by*

$$k_l := 1_R \otimes w_l$$

for $l = 1, ..., n$. *Then*

$$W_\sigma\left(\mathbf{A}^{\mathfrak{r}}, \mathbf{B}^{\mathfrak{r}}\right) \leq W_\sigma^{\mathfrak{r}}\left(\mathbf{A}, \mathbf{B}\right)$$

for all $\mathbf{A}, \mathbf{B} \in X_\otimes$.

Proof. For any $\omega \in T\left(\mathbf{A}^{\mathfrak{p}}, \mathbf{B}^{\mathfrak{p}}\right)$, and writing $w = (w_1, ..., w_n)$ and $k = (k_1, ..., k_n)$, we have

$$I_w(\omega^{\mathfrak{r}}) = I_k(\omega)$$

in terms of (5), because of Lemma 4.8 and $\omega^{\mathfrak{r}} \in T(\mu_S, \nu_S)$.

In addition, $\mathbf{A}^{\mathfrak{r}}\omega^{\mathfrak{r}}\mathbf{B}^{\mathfrak{r}}$ by Proposition 4.5. Assuming $\left(\mathbf{A}^{\mathfrak{p}}\right)^\sigma \omega \left(\mathbf{B}^{\mathfrak{p}}\right)^\sigma$ as well, we know from Proposition 4.6 that $\left(\mathbf{A}^\sigma\right)^{\mathfrak{p}} \omega \left(\mathbf{B}^\sigma\right)^{\mathfrak{p}}$, hence $\left(\mathbf{A}^\sigma\right)^{\mathfrak{r}} \omega^{\mathfrak{r}} \left(\mathbf{B}^\sigma\right)^{\mathfrak{r}}$, giving $\left(\mathbf{A}^{\mathfrak{r}}\right)^\sigma \omega^{\mathfrak{r}} \left(\mathbf{B}^{\mathfrak{r}}\right)^\sigma$, by Propositions 4.5 and 4.6. Recalling Remark 3.6, this means that

$$\{\omega^{\mathfrak{r}} : \omega \in T_\sigma(\mathbf{A}^{\mathfrak{p}}, \mathbf{B}^{\mathfrak{p}})\} \subset T_\sigma(\mathbf{A}^{\mathfrak{r}}, \mathbf{B}^{\mathfrak{r}}).$$

In view of Definition 4.9, these two facts imply the result. □

Of course, assuming that $\{w_1, ..., w_n\}^* = \{w_1, ..., w_n\}$, and that this set generates S, the associated W_σ on Y in this theorem is in fact a metric according to Theorem 3.9.

An analogous result is achieved for W instead of W_σ. Simply consider the set $\mathsf{X}_\otimes$ of systems $\mathbf{A} = (M, \alpha, \mu)$ of the form (7), but with σ^μ dropped, and $\mathsf{X}_\otimes^{\mathfrak{p}}$ consisting of the corresponding augmented systems $\mathbf{A}^{\mathfrak{p}} = \left(M, \alpha, P_{1\otimes S}^{\mu_R}, \mu\right)$. Similarly, define Y as the set of corresponding reduced systems $\mathbf{A}^{\mathfrak{r}} = (S, \alpha^{\mathfrak{r}}, \mu_S)$. Define the *reduced* asymmetric Wasserstein pseudometric $W^{\mathfrak{r}}\left(\mathbf{A}, \mathbf{B}\right) := W\left(\mathbf{A}^{\mathfrak{p}}, \mathbf{B}^{\mathfrak{p}}\right)$ for all $\mathbf{A}, \mathbf{B} \in \mathsf{X}_\otimes$, from the asymmetric Wasserstein pseudometric W on $\mathsf{X}_\otimes^{\mathfrak{p}}$ associated to $k_1, ..., k_n \in M$, given by Theorem 3.10. Remove aspects related to KMS-duals from the preceding proof. Then we analogously have the following result:

Theorem 4.11. *Consider the asymmetric Wasserstein pseudometric* W *on* Y *associated to* $w_1, ..., w_n \in S$, *and the reduced modular Wasserstein pseudometric* $W^{\mathfrak{r}}$ *on* $\mathsf{X}_\otimes$ *associated to* $k_1, ..., k_n \in M$ *given by*

$$k_l := 1_R \otimes w_l$$

for $l = 1, ..., n$. *Then*

$$W\left(\mathbf{A}^{\mathfrak{r}}, \mathbf{B}^{\mathfrak{r}}\right) \leq W^{\mathfrak{r}}\left(\mathbf{A}, \mathbf{B}\right)$$

for all $\mathbf{A}, \mathbf{B} \in \mathsf{X}_\otimes$.

5. Examples

In this section we present a few simple but enlightening examples to provide some insight into the Wasserstein distances. As Wasserstein distances are usually defined on the states in the classical case (i.e., on probability measures), we are instead particularly interested in the behaviour of the Wasserstein distances of this paper with regards to the dynamics, including the modular dynamics.

We investigate how sharply these distances distinguish different dynamics, given the same pair of states. As will be seen, not necessarily very sharply, indicating that Wasserstein distance between systems chiefly measures distances between the states of systems, but in a way that is regulated by the dynamics. This may

be expected from the definition. An important related point is that balance conditions can cause jumps in values of the Wasserstein distances. In particular this can happen for modular balance, indicating that the latter condition may be unnatural in some ways when we are interested in distances between the states themselves, not between systems. On the other hand, it will be argued that these jumps should be of value in discerning qualitative differences between the dynamics of two systems.

The other point we make in this section, is that symmetry of a Wasserstein distance can indeed be lost if balance with respect to the modular dynamics is not included.

We divide the section into subsections. The first concerns a very simple conceptual point used in the examples, the last gives a brief summary of the conclusions and possible implications, while the others consider the examples in turn.

5.1. Equal distance

Given Υ and the Z_υ's, let D_μ denote the set of all α's appearing in Definition 3.1. Given $\alpha_1, \alpha_2 \in D_\mu$ and $\beta_1, \beta_2 \in D_\nu$, write

$$\mathbf{A}_l = (M, \alpha_l, \mu) \text{ and } \mathbf{B}_l = (M, \beta_l, \nu)$$

for $l = 1, 2$. Note that if

$$T_\sigma(\mathbf{A}_1, \mathbf{B}_1) = T_\sigma(\mathbf{A}_2, \mathbf{B}_2), \tag{8}$$

then

$$W_\sigma(\mathbf{A}_1, \mathbf{B}_1) = W_\sigma(\mathbf{A}_2, \mathbf{B}_2)$$

for any modular Wasserstein pseudometric on X. This is not a particularly powerful condition, but since it does not refer to details regarding the transport cost function, it can give us general conclusions independent of transport cost. It can also be comparatively straightforward to check for simple systems. The analogous remarks are of course also true for W.

5.2. Unitary dynamics in M_2

Consider the 2×2 complex matrices M_2. It is simpler to work directly in this representation, instead of the standard form $M = M_2 \otimes 1_2$ with 1_2 the 2×2 identity matrix. We use $\Upsilon = \{1\}$, and $Z_1 = \mathbb{Z}$ in the notation of Definition 3.1, and view the modular dynamics separately, as we did in Section 3. Given a fixed $\theta \in \mathbb{R}$, we define systems $\mathbf{A}_\theta = (\alpha_\theta, \mu)$ and $\mathbf{B}_\theta = (\alpha_\theta, \nu)$ on M_2 by the dynamics

$$\alpha_\theta(a) = U_\theta a U_\theta^*$$

acting through iteration (hence $Z_1 = \mathbb{Z}$ above) and the states

$$\mu(a) = \mathrm{Tr}(\zeta a) \text{ and } \nu(a) = \mathrm{Tr}(\eta a),$$

where U_θ is the unitary matrix

$$U_\theta = \begin{bmatrix} 1 & 0 \\ 0 & e^{i\theta} \end{bmatrix}$$

for every $\theta \in \mathbb{R}$, and ζ and η are the density matrices

$$\zeta = \begin{bmatrix} p_1 & 0 \\ 0 & p_2 \end{bmatrix} \text{ and } \eta = \begin{bmatrix} q_1 & 0 \\ 0 & q_2 \end{bmatrix}$$

with $0 < p_1, q_1 < 1$ and $p_1 + p_2 = q_1 + q_2 = 1$. Write any $\omega \in T(\mu, \nu)$ in terms of its 4×4 density matrix $\kappa = [\kappa_{kl}]$, that is,

$$\omega(c) = \mathrm{Tr}(\kappa c)$$

for all $c \in M_2 \odot M_2$. Note that $M_2 \odot M_2$ is the correct way to express $M \odot M'$ in our current representation.

We first want to determine the effect of the balance condition $\mathbf{A}_\phi \omega \mathbf{B}_\theta$ by itself on κ, without making any other assumptions about the matrix κ. For the moment we therefore assume that $\kappa = [\kappa_{lm}]$ is an arbitrary 4×4 complex matrix, defining a linear functional $\omega = \mathrm{Tr}(\kappa \cdot)$ on $M_2 \odot M_2$. Then it is elementary to check that the restrictions placed on κ by the balance condition $\mathbf{A}_\phi \omega \mathbf{B}_\theta$, are described by the following:

(a) $e^{i\phi} \neq 1$ implies $\kappa_{13} = \kappa_{31} = \kappa_{24} = \kappa_{42} = 0$, while $e^{i\phi} = 1$ has no implications for κ.
(b) $e^{i\theta} \neq 1$ implies $\kappa_{12} = \kappa_{21} = \kappa_{34} = \kappa_{43} = 0$, while $e^{i\theta} = 1$ has no implications for κ.
(c) $e^{i\phi} \neq e^{i\theta}$ implies $\kappa_{14} = \kappa_{41} = 0$, while $e^{i\phi} = e^{i\theta}$ has no implications for κ.
(d) $e^{i\phi} \neq e^{-i\theta}$ implies $\kappa_{23} = \kappa_{32} = 0$, while $e^{i\phi} = e^{-i\theta}$ has no implications for κ.

Since α_θ is a $*$-automorphism, $\mathbf{A}_\phi^\sigma \omega \mathbf{B}_\theta^\sigma$ is automatically satisfied if $\mathbf{A}_\phi \omega \mathbf{B}_\theta$ holds; see Remark 3.6.

The effect of modular balance $(\sigma^\mu, \mu)\omega(\sigma^\nu, \nu)$ is of exactly the same form. We have

$$\sigma_t^\mu(a) = \zeta^{it} a \zeta^{-it} = \begin{bmatrix} a_{11} & \left(\frac{p_2}{p_1}\right)^{-it} a_{12} \\ \left(\frac{p_2}{p_1}\right)^{it} a_{21} & a_{22} \end{bmatrix},$$

as opposed to

$$\alpha_\phi(a) = \begin{bmatrix} a_{11} & e^{-i\phi} a_{12} \\ e^{i\phi} a_{21} & a_{22} \end{bmatrix},$$

in terms of $a = [a_{lm}]$. It follows that modular balance has the same set of implications for κ as (a) to (d) above, with the corresponding inequality in each condition just having to hold for some value of t. Consequently, $e^{i\phi}$ can be replaced by p_2/p_1, and $e^{i\theta}$ by q_2/q_1, in (a) to (d). I.e., $p_2/p_1 \neq 1$, implies that $(p_2/p_1)^{it} \neq 1$ for some t, etc.

Consider the case where the inequality in each condition (a) to (d) for modular dynamics holds, namely $p_2/p_1 \neq 1$, $q_2/q_1 \neq 1$, $p_2/p_1 \neq q_2/q_1$ and $p_2/p_1 \neq q_1/q_2$. We intend to show that, irrespective of the transport cost function, the modular Wasserstein distance

$$W_\sigma(\mathbf{A}_\phi, \mathbf{B}_\theta)$$

is independent of ϕ and θ. To see this, note that in this case, κ is diagonal by (a) to (d) applied to modular balance. Therefore $\mathbf{A}_\phi \omega \mathbf{B}_\theta$ (and thus necessarily $\mathbf{A}_\phi^\sigma \omega \mathbf{B}_\theta^\sigma$) has no further effect on κ. To this we need to add the coupling property, which in our current representation reads $\omega(a \otimes 1) = \mu(a)$ and $\omega(1 \otimes a) = \nu(a)$, and the positivity of κ, neither of which depend on ϕ and θ. It follows that $T_\sigma(\mathbf{A}_\phi, \mathbf{B}_\theta)$, and thus $W_\sigma(\mathbf{A}_\phi, \mathbf{B}_\theta)$, are independent of ϕ and θ, as claimed. Since $\mu \neq \nu$, as $p_2/p_1 \neq q_2/q_1$, we of course have $W_\sigma(\mathbf{A}_\phi, \mathbf{B}_\theta) \neq 0$.

Other cases can be similarly explored, with similar conclusions. For example, assuming that $e^{i\phi} \neq 1$, $e^{i\theta} \neq 1$, $e^{i\phi} \neq e^{i\theta}$ and $e^{i\phi} \neq e^{-i\theta}$, but dropping the assumptions $p_2/p_1 \neq 1$, $q_2/q_1 \neq 1$, $p_2/p_1 \neq q_2/q_1$ and $p_2/p_1 \neq q_1/q_2$, we obtain diagonal κ from $\mathbf{A}_\phi \omega \mathbf{B}_\theta$. In this situation modular balance obviously holds no further implications for κ, since (a) to (d) will at most reproduce the conditions already covered by $\mathbf{A}_\phi \omega \mathbf{B}_\theta$. It follows that $W(\mathbf{A}_\phi, \mathbf{B}_\theta) = W_\sigma(\mathbf{A}_\phi, \mathbf{B}_\theta) \neq 0$, and that the value is independent of ϕ and θ. It is non-zero, even for $\mu = \nu$, simply because $\alpha_\phi \neq \alpha_\theta$.

5.3. Asymmetry of W

We proceed with the previous example. Here and in the next subsection we calculate some values of W and W_σ explicitly for a specific transport cost function. In particular, we use this example to show that without the modular balance conditions in Definition 3.5, symmetry can be lost. That is, W in Theorem 3.10, can indeed lack symmetry. In this example, the lack of symmetry is specifically due to an absence of balance between the modular dynamics, since the KMS duals are in balance with respect to all the allowed transport plans, as mentioned above.

We now calculate $W(\mathbf{A}_\theta, \mathbf{B}_\theta)$. For simplicity we assume that

$$0 < p \leq q < \frac{1}{2},$$

where we have written $p = p_1$ and $q = q_1$.

The transport cost function used in defining W, is chosen to be given by the following two self-adjoint generators for M_2:

$$k_1 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \text{ and } k_2 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

Then, assuming $e^{i\theta} \neq 1$ and $e^{i\theta} \neq e^{-i\theta}$,

$$W(\mathbf{A}_\theta, \mathbf{B}_\theta) = 2 + q - p - 2\sqrt{\frac{p}{q}}$$

and

$$W(\mathbf{B}_\theta, \mathbf{A}_\theta) = 2 + q - p - 2\sqrt{\frac{1-q}{1-p}}.$$

This is achieved by determining the set of linear functionals ω satisfying balance, restricting it further by the coupling property and then positivity, to obtain the set of all allowed transport plans, and finally using standard calculus to obtain the minimum cost. One way of doing this, is to use (5) directly. Another way, is to use a "cost matrix" c_k as given by [25, Equation (2)], and using the formula

$$I_k(\omega) = \omega(c_k),$$

which in finite dimensions is equivalent to (5), as can be seen from [25, Section 3].

In particular, we see that

$$W(\mathbf{A}_\theta, \mathbf{B}_\theta) > W(\mathbf{B}_\theta, \mathbf{A}_\theta)$$

when $p < q$. That is, W is not symmetric.

5.4. Jumps in W_σ

Including modular balance in Subsection 5.3 and recalling that $\mathbf{A}_\phi^\sigma \omega \mathbf{B}_\theta^\sigma$ is automatically satisfied, one finds that because of condition (c) above,

$$W_\sigma(\mathbf{A}_\theta, \mathbf{B}_\theta) = \begin{cases} 2 + q - p \text{ if } p < q \\ 0 \text{ if } p = q. \end{cases} \tag{9}$$

Note the jump in value from $p < q$ to $p = q$. The reason this happens is as follows: The condition $\mathbf{A}_\theta \omega \mathbf{B}_\theta$ forces κ to be diagonal, except for κ_{14} and κ_{41}, which are not forced to be zero. On the other hand, for $p < q$, modular balance does force κ to be diagonal (so $\kappa_{14} = \kappa_{41} = 0$), whereas for $p = q$, it does not. Thus, for $p < q$, the allowed set of transport plans is smaller, leading to the larger value in (9) for W_σ, compared to W. But when $p = q$, modular balance does not add further restrictions on the allowed set of transport plans beyond those due to $\mathbf{A}_\theta \omega \mathbf{B}_\theta$, hence in this case we have $W_\sigma(\mathbf{A}_\theta, \mathbf{B}_\theta) = W(\mathbf{A}_\theta, \mathbf{B}_\theta) = W(\mathbf{B}_\theta, \mathbf{A}_\theta) = 0$.

This jump in $W_\sigma(\mathbf{A}_\theta, \mathbf{B}_\theta)$ indicates that the inclusion of modular balance may not be natural in all respects if we want to consider distances between states. For example, one would prefer $W_\sigma(\mathbf{A}_\theta, \mathbf{B}_\theta)$ to converge to zero as μ and ν approach one another in a natural way, as in this case where p and q approach each other in the usual metric on $\mathbb{R}$.

The latter is indeed what happens in the classical analogy of this example, for a two point metric space and the expectations with respect to the probability measures on this space being the states on the von Neumann algebra $\mathbb{C}^2$ of functions on it. No jumps as above occur in this case, since even in a much more general setting, the classical Wasserstein metric on the states, metrizes the weak topology on the states [52, Remark 7.13(iii)].

This could mean that for states, the asymmetric Wasserstein metrics in Theorem 3.10, are more natural in the bimodule approach. As mentioned in [25, Section 7], this may be sensible, as the direction of transport would then appear to be reflected to some extent in the asymmetric metric.

As the remaining subsections illustrate, jumps can occur in W as well, for essentially the same reasons, due to other dynamics. For systems (as to opposed to states) jumps in W and W_σ could be viewed as feature, not a drawback, as will be explained.

5.5. Reduced systems on M_2

We proceed in a similar vein as the previous subsections, but in the context of Section 4.

We set $R = S = M_2$ in the notation of Section 4, but in terms of the same approach to representations as in Subsection 5.2. Consider dynamics, including an interaction between the reservoir and the open system, described by a Hamiltonian of the form

$$h = \Theta \otimes 1_2 + 1_2 \otimes \Phi + \lambda u \otimes v,$$

where $\Theta, \Phi, u, v \in M_2$ are diagonal real matrices, and $\lambda \geq 0$ controls the strength of the interaction.

Let the states μ_R and μ_S be given by diagonal density matrices. The dynamics on $R \bar{\otimes} S = R \odot S$ is given by $\alpha_t = e^{iht}(\cdot)e^{-iht}$ for all $t \in \mathbb{R}$, which clearly leaves μ invariant as required. Using

$$d^\mu = \begin{bmatrix} d_1^\mu & 0 \\ 0 & d_2^\mu \end{bmatrix}$$

with $0 < d_1^\mu < 1$ as the density matrix of μ_R, the reduction of α to S gives

$$\alpha_t^\lambda(s) := \alpha_t^{\mathfrak{r}}(s) = \begin{bmatrix} s_{11} & \Xi_{\mu,t}(\lambda)s_{12} \\ \Xi_{\mu,t}^*(\lambda)s_{21} & s_{22} \end{bmatrix}$$

for all $s = [s_{lm}] \in M_2$, in terms of

$$\begin{aligned}\Xi_{\mu,t}(\lambda) &:= d_1^\mu e^{i(h_1-h_2)t} + d_2^\mu e^{i(h_3-h_4)t} \\ &= \left[d_1^\mu e^{i\lambda u_1(v_1-v_2)t} + d_2^\mu e^{i\lambda u_2(v_1-v_2)t}\right] e^{i(\Phi_1-\Phi_2)t}\end{aligned}$$

and its complex conjugate $\Xi_{\mu,t}^*$, where $h_1, ..., h_4$, Φ_1, Φ_2, u_1, u_2 and v_1, v_2 are the (diagonal) entries of h, Φ, u and v respectively. Note that α^λ depends on μ_R, but not on the Hamiltonian Θ used for R, or on μ_S. We have thus obtained a reduced system

$$\mathbf{A}^\lambda = \left(\alpha^\lambda, \mu_S\right)$$

on S, with α^λ independent of the state μ_S. Note that α^λ does not have the semigroup property, unless $\lambda(u_l - u_2)(v_1 - v_2) = 0$. The latter condition boils down to no interaction.

Following exactly the same procedure, we can obtain another reduced system on S,

$$\mathbf{B}^\lambda = \left(\beta^\lambda, \nu_S\right),$$

by using another state $\nu = \nu_R \otimes \nu_S$ given by diagonal density matrices, while keeping the rest of the specifications the same.

Using the corresponding restrictions provided by (a) to (d) in Subsection 5.2, we can draw analogous conclusions. For example, assume that $\Phi_1 \neq \Phi_2$, $u_1, u_2 > 0$, $u_1 \neq u_2$ and $v_1 \neq v_2$. For the case $\mu_R \neq \nu_R$ it then follows that

$$W\left(\mathbf{A}^{\lambda_1}, \mathbf{B}^{\lambda_2}\right)$$

is independent of the two parameters λ_1 and λ_2, as long as at least one of them is not zero. If $\mu_R = \nu_R$, on the other hand, then $W\left(\mathbf{A}^{\lambda_1}, \mathbf{B}^{\lambda_2}\right)$ is independent of λ_1 and λ_2, when $\lambda_1 \neq \lambda_2$, even if one of them is zero, and alternatively when $\lambda_1 = \lambda_2$. Both these cases hold simply because the allowed κ are diagonal when at least one of λ_1 or λ_2 is not zero (respectively, when $\lambda_1 \neq \lambda_2$), whereas κ_{14} and κ_{41} are allowed to be non-zero otherwise.

Moreover, using the same transport cost as in Subsection 5.3, and ζ and η appearing in Subsection 5.2 as the density matrix of μ_S and ν_S respectively, we consequently find similar values for the Wasserstein distance as found in Subsection 5.3. More precisely, again writing $p = p_1$ and $q = q_1$, and assuming $0 < p \leq q < \frac{1}{2}$, we find:

For $\mu_R \neq \nu_R$,

$$W(\mathbf{A}^{\lambda_1}, \mathbf{B}^{\lambda_2}) = W(\mathbf{B}^{\lambda_2}, \mathbf{A}^{\lambda_1}) = 2 + q - p \tag{10}$$

if at least one of λ_1 or λ_2 is not zero, simply because in this case for W, the same sets of transport plans (one set for each of the two orders in which the systems appear in W) are allowed as for W_σ in the first line of (9), even for $p = q$, while

$$W(\mathbf{A}^0, \mathbf{B}^0) = 2 + q - p - 2\sqrt{\frac{p}{q}}$$

and

$$W(\mathbf{B}^0, \mathbf{A}^0) = 2 + q - p - 2\sqrt{\frac{1-q}{1-p}}.$$

Analogous to W_σ in Subsection 5.3, we now see a jump in W when the point $\lambda_1 = \lambda_2 = 0$ is reached, since this is where the sets of allowed transport plans change. The same values are achieved for $\mu_R = \nu_R$, by the same arguments, but for the cases $\lambda_1 \neq \lambda_2$ and $\lambda_1 = \lambda_2$ respectively, i.e., $W(\mathbf{A}^\lambda, \mathbf{B}^\lambda) = 2 + q - p - 2\sqrt{p/q}$, etc.

In this example it appears that W is not sensitive to the interaction strength λ, except that for $\mu_R \neq \nu_R$ it distinguishes between the case where non-zero interactions are involved, and the case where there are no interactions, while for $\mu_R = \nu_R$ it discerns the case where the interaction strengths differ, from the case where they do not.

When $\mu_S \neq \nu_S$, the distance W_σ distinguishes even less well between the different cases for the reduced dynamics in this example, since balance between the modular dynamics is then already enough to force κ to be diagonal. Keeping in mind the implications of Corollary 4.7, it follows as for (10), that

$$W_\sigma(\mathbf{A}^{\lambda_1}, \mathbf{B}^{\lambda_2}) = 2 + q - p$$

for all $\lambda_1, \lambda_2 \geq 0$, whether $\mu_R \neq \nu_R$ or not. For $\mu_S = \nu_S$, on the other hand, balance between the modular dynamics does not force κ_{14} and κ_{41} to be zero, hence this balance condition becomes redundant, thus $W_\sigma(\mathbf{A}^{\lambda_1}, \mathbf{B}^{\lambda_2}) = W(\mathbf{A}^{\lambda_1}, \mathbf{B}^{\lambda_2})$ for all $\lambda_1, \lambda_2 \geq 0$, returning the various values above, but with $p = q$.

Therefore, when we want to distinguish qualitative differences between dynamics more sharply, W has the advantage, while W_σ gives a coarser view, as is to be expected from (6).

5.6. Translations on quantum tori

For our last example we make use of disjointness as it appears in the theory of joinings. Systems are *disjoint* exactly when the set of transport plans is trivial, that is, consists solely of the product coupling. In this situation Wasserstein distances are simple to compute.

In short, we consider actions of $\mathbb{R}^2$ on the von Neumann algebraic quantum torus (or irrational rotation algebra) M_θ for a given irrational number θ, represented in standard form on the Hilbert space $G = L^2(\mathbb{T}^2)$, with respect to the normalized Haar measure on $\mathbb{T}^2$, where $\mathbb{T}^2 = \mathbb{R}^2/\mathbb{Z}^2$ is the classical torus. We have such an action $\alpha^{p,q}$ for every pair of real numbers p and q, given by

$$\alpha^{p,q}_{s,t} = \tau_{ps,qt}$$

for all $(s,t) \in \mathbb{R}^2$, where τ is the natural translation on M_θ.

The quantum torus is a standard example in operator algebra, but for clarity we briefly outline its construction. With every $\theta \in \mathbb{R}$, we associate the von Neumann algebra M_θ generated by the unitary operators $u, v \in \mathcal{B}(G)$ defined by

$$(uf)(x,y) := e^{2\pi i x} f(x, y + \theta/2)$$

and

$$(vf)(x,y) := e^{2\pi i y} f(x - \theta/2, y)$$

for all $f \in G$ and $(x,y) \in \mathbb{T}^2$. Then $\Lambda := 1 \in G$ is a cyclic and separating vector for M_θ, which defines a faithful normal trace Tr on M_θ by

$$\mathrm{Tr}(a) := \langle \Lambda, a\Lambda \rangle$$

for all $a \in M_\theta$.

A natural action τ of $\mathbb{R}^2$ on M_θ, as $*$-automorphisms which leave Tr invariant, can be obtained via the classical translations

$$T_{s,t}(x,y) := (x+s, y+t) \in \mathbb{T}^2$$

for all $(x,y) \in \mathbb{T}^2$, by defining

$$U_{s,t} : G \to G : f \mapsto f \circ T_{s,t}$$

and setting

$$\tau_{s,t}(a) := U_{s,t} a U_{s,t}^*$$

for all $a \in M_\theta$, and all $(s,t) \in \mathbb{R}^2$. We thus have a system $\mathbf{A}_{p,q} = (M_\theta, \alpha^{p,q}, \mathrm{Tr})$ for any real numbers p and q.

According to [27, Example 3.14] and Remark 3.6, we have

$$T_\sigma(\mathbf{A}_{p,q}, \mathbf{A}_{c,d}) = T(\mathbf{A}_{p,q}, \mathbf{A}_{c,d}) = \{\mathrm{Tr} \odot \mathrm{Tr}'\}$$

for all $p, q, c, d \in \mathbb{R}\backslash\{0\}$ with p/c or q/d irrational. Consequently, for a transport cost function given by any self-adjoint set of generators of M_θ, it follows that $W_\sigma(\mathbf{A}_{p,q}, \mathbf{A}_{c,d}) = W(\mathbf{A}_{p,q}, \mathbf{A}_{c,d}) \neq 0$ is independent of p, q, c and d, as long as the mentioned irrationality condition holds. In particular, these Wasserstein distances do not depend on how far (p,q) is from (c,d) in $\mathbb{R}^2$. As an obvious example, using $k = (u, v, u^*, v^*)$ in Definition 3.7, one easily calculates $W_\sigma(\mathbf{A}_{p,q}, \mathbf{A}_{c,d}) = W(\mathbf{A}_{p,q}, \mathbf{A}_{c,d}) = 8$, since $E_{\mathrm{Tr} \odot \mathrm{Tr}'} = \mathrm{Tr}(\cdot)1_{M_\theta}$ because of (2).

For p, q, c and d for which the above mentioned irrationality condition does not hold, disjointness could fail, that is, the set of transport plans could be larger than $\{\mathrm{Tr} \odot \mathrm{Tr}'\}$, potentially leading to jumps in the Wasserstein distances. In particular, we of course have $W_\sigma(\mathbf{A}_{p,q}, \mathbf{A}_{p,q}) = W(\mathbf{A}_{p,q}, \mathbf{A}_{p,q}) = 0$ for any k in Definition 3.7.

5.7. Summary of conclusions

We considered systems parametrized by constants appearing in their dynamics or states. Varying these parameters, gives families of related systems. For pairs of systems, this includes ranges where the dynamics of the two systems are in some relation, for example having different interactions with a reservoir or when the state of the reservoir is the same for both reduced systems. Both W and W_σ are often independent of such parameters, except for jumps at special points. These jumps are the result of changes in the allowed set of transport plans. Such a change tends to occur in a "discrete" way when the two systems' dynamics reaches some threshold, rather than just because of any small perturbations in the dynamics.

From this limited range of examples, it appears that W and W_σ, and indeed also the sets of transport plans T and T_σ, tend to be able to discern qualitative or structural differences between dynamics, but are less sensitive to finer details or small perturbations in the dynamics. This is in line with what may be expected from the definition, which adapts the definition for states (or measures in the classical case) by still considering a cost function on transport plans between states, rather than cost directly related to dynamics, but regulates the allowed sets of transport plans using the dynamics. We expect this behaviour of W and W_σ to be of value in the intended applications mentioned in the introduction, namely questions related to balance, detailed balance and joinings, where qualitative differences between systems are relevant.

In Subsection 5.4 we also saw a case where W_σ varies continuously with the states (while keeping the dynamics the same), except for a discontinuity when the states become equal. This is due to the balance condition for the modular dynamics. As mentioned, this could be an indication that W is a more natural distance function than W_σ for states themselves (as opposed to systems), though at the cost of losing symmetry of the distance function.

It would be of interest to investigate the ideas of this section in more general terms, to determine to what extent the conclusions drawn here remain valid and to build further theory around them.

Acknowledgment

I thank the referee for suggestions to improve aspects of the presentation.

References

[1] L. Accardi, C. Cecchini, Conditional expectations in von Neumann algebras and a theorem of Takesaki, J. Funct. Anal. 45 (1982) 245–273.
[2] J. Agredo, F. Fagnola, On quantum versions of the classical Wasserstein distance, Stochastics 89 (2017) 910–922.
[3] H. Araki, Some properties of modular conjugation operator of von Neumann algebras and a non-commutative Radon-Nikodym theorem with a chain rule, Pac. J. Math. 50 (1974) 309–354.
[4] J.P. Bannon, J. Cameron, K. Mukherjee, On noncommutative joinings, Int. Math. Res. Not. (2018) 4734–4779.
[5] J. Bannon, J. Cameron, K. Mukherjee, Noncommutative joinings II, Groups Geom. Dyn. 15 (2021) 553–575.
[6] J.P. Bannon, J. Cameron, K. Mukherjee, On noncommutative joinings III, Oper. Matrices 14 (2020) 469–490.
[7] J.-D. Benamou, Y. Brenier, A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem, Numer. Math. 84 (2000) 375–393.
[8] P. Biane, D. Voiculescu, A free probability analogue of the Wasserstein metric on the trace-state space, Geom. Funct. Anal. 11 (2001) 1125–1138.
[9] O. Bratteli, D.W. Robinson, Operator Algebras and Quantum Statistical Mechanics 1, second edition, Springer-Verlag, New York, 1987.
[10] H.-P. Breuer, F. Petruccione, The Theory of Open Quantum Systems, Oxford University Press, New York, 2002.
[11] J.W. Brown, Approximation theorems for Markov operators, Pac. J. Math. 16 (1966) 13–23.
[12] E.A. Carlen, J. Maas, An analog of the 2-Wasserstein metric in non-commutative probability under which the fermionic Fokker-Planck equation is gradient flow for the entropy, Commun. Math. Phys. 331 (2014) 887–926.
[13] Y. Chen, T.T. Georgiou, A. Tannenbaum, Matrix optimal mass transport: a quantum mechanical approach, IEEE Trans. Autom. Control 63 (2018) 2612–2619.
[14] M.-D. Choi, Completely positive linear maps on complex matrices, Linear Algebra Appl. 10 (1975) 285–290.
[15] F. Cipriani, Dirichlet forms and Markovian semigroups on standard forms of von Neumann algebras, J. Funct. Anal. 147 (1997) 259–300.
[16] S. Cole, M. Eckstein, S. Friedland, K. Życzkowski, Quantum optimal transport, arXiv:2105.06922.
[17] A. Connes, Caractérisation des espaces vectoriels ordonnés sous-jacents aux algèbres de von Neumann, Ann. Inst. Fourier (Grenoble) 24 (1974) 121–155.
[18] A. Connes, Noncommutative Geometry, Academic Press, San Diego, CA, 1994.
[19] E.B. Davies, Quantum Theory of Open Systems, Academic Press, London-New York, 1976.
[20] G. De Palma, M. Marvian, D. Trevisan, S. Lloyd, The quantum Wasserstein distance of order 1, IEEE Trans. Inf. Theory 67 (2021) 6627–6643.
[21] G. De Palma, D. Trevisan, Quantum optimal transport with quantum channels, Ann. Henri Poincaré 22 (2021) 3199–3234.
[22] R. Duvenhage, Joinings of W*-dynamical systems, J. Math. Anal. Appl. 343 (2008) 175–181.
[23] R. Duvenhage, Ergodicity and mixing of W*-dynamical systems in terms of joinings, Ill. J. Math. 54 (2010) 543–566.
[24] R. Duvenhage, Relatively independent joinings and subsystems of W*-dynamical systems, Stud. Math. 209 (2012) 21–41.
[25] R. Duvenhage, Quadratic Wasserstein metrics for von Neumann algebras via transport plans, J. Oper. Theory 88 (2022) 289–308.
[26] R. Duvenhage, M. Snyman, Balance between quantum Markov semigroups, Ann. Henri Poincaré 19 (2018) 1747–1786.
[27] R. Duvenhage, A. Ströh, Disjointness of C*-dynamical systems, Houst. J. Math. 42 (2016) 223–247.
[28] F. Fagnola, R. Rebolledo, Entropy production for quantum Markov semigroups, Commun. Math. Phys. 335 (2015) 547–570.
[29] F. Fagnola, V. Umanità, Generators of KMS symmetric Markov semigroups on B(h) symmetry and quantum detailed balance, Commun. Math. Phys. 298 (2010) 523–547.
[30] T. Falcone, L^2-von Neumann modules, their relative tensor products and the spatial derivative, Ill. J. Math. 44 (2000) 407–437.
[31] D. Feliciangeli, A. Gerolin, L. Portinale, A non-commutative entropic optimal transport approach to quantum composite systems at positive temperature, arXiv:2106.11217.
[32] W. Gangbo, D. Jekel, K. Nam, D. Shlyakhtenko, Duality for optimal couplings in free probability, Commun. Math. Phys. 396 (2022) 903–981.

[33] L. Gao, C. Rouzé, Ricci curvature of quantum channels on non-commutative transportation metric spaces, arXiv:2108.10609.
[34] E. Glasner, Ergodic Theory via Joinings, Mathematical Surveys and Monographs, vol. 101, American Mathematical Society, Providence, RI, 2003.
[35] S. Goldstein, J.M. Lindsay, Beurling-Deny conditions for KMS-symmetric dynamical semigroups, C. R. Acad. Sci., Sér. 1 Math. 317 (1993) 1053–1057.
[36] S. Goldstein, J.M. Lindsay, KMS-symmetric Markov semigroups, Math. Z. 219 (1995) 591–608.
[37] F. Golse, C. Mouhot, T. Paul, On the mean field and classical limits of quantum mechanics, Commun. Math. Phys. 343 (2016) 165–205.
[38] U. Haagerup, The standard form of von Neumann algebras, Math. Scand. 37 (1975) 271–283.
[39] F. Hiai, D. Petz, Y. Ueda, Free transportation cost inequalities via random matrix approximation, Probab. Theory Relat. Fields 130 (2004) 199–221.
[40] D.F. Hornshaw, L^2-Wasserstein distances of tracial W*-algebras and their disintegration problem, arXiv:1806.01073.
[41] R.M. Gray, D.L. Neuhoff, P.C. Shields, A generalization of Ornstein's $\bar{d}$ distance with applications to information theory, Ann. Probab. 3 (1975) 315–328.
[42] K. Ikeda, Foundation of quantum optimal transport and applications, Quantum Inf. Process. 19 (1) (2020) 25.
[43] P. Mikusiński, M.D. Taylor, Markov operators and n-copulas, Ann. Pol. Math. 96 (2009) 75–95.
[44] L. Ning, T. Georgiou, A. Tannenbaum, On matrix-valued Monge-Kantorovich optimal mass transport, IEEE Trans. Autom. Control 60 (2015) 373–382.
[45] M. Ohya, D. Petz, Quantum Entropy and Its Use, Texts and Monographs in Physics, Springer-Verlag, Berlin, 1993.
[46] D. Ornstein, An application of ergodic theory to probability theory, Ann. Probab. 1 (1973) 43–65.
[47] D. Petz, A dual in von Neumann algebras with weights, Quart. J. Math. Oxford Ser. (2) 35 (1984) 475–483.
[48] L. Rüschendorf, T. Sei, On optimal stationary couplings between stationary processes, Electron. J. Probab. 17 (2012) 17.
[49] J.-L. Sauvageot, Sur le produit tensoriel relatif d'espaces de Hilbert, J. Oper. Theory 9 (1983) 237–252.
[50] M. Takesaki, Theory of Operator Algebras. II, Encyclopaedia of Mathematical Sciences: Operator Algebras and Non-commutative Geometry, 6, vol. 125, Springer-Verlag, Berlin, 2003.
[51] L.N. Vasershtein, Markov processes over denumerable products of spaces describing large system of automata, Probl. Inf. Transm. 5 (1969) 47–52.
[52] C. Villani, Topics in Optimal Transportation, Graduate Studies in Mathematics, vol. 58, American Mathematical Society, Providence, RI, 2003.
[53] M. Wirth, A noncommutative transport metric and symmetric quantum Markov semigroups as gradient flows of the entropy, arXiv:1808.05419.