

QUANTUM WASSERSTEIN DISTANCE OF ORDER 1 BETWEEN CHANNELS

ROCCO DUVENHAGE AND MATHUMO MAPAYA

ABSTRACT. We set up a general theory leading to a quantum Wasserstein distance of order 1 between channels in an operator algebraic framework. This gives a metric on the set of channels from one composite system to another, which is deeply connected to reductions of the channels. The additivity and stability properties of this metric are studied.

Keywords: quantum optimal transport; quantum Wasserstein distance of order 1; quantum channels; composite systems

CONTENTS

1. Introduction	2
2. Outline in finite dimensions	5
2.1. A representation of channels	5
2.2. The W_1 norm	6
2.3. Wasserstein distance of order 1 between channels	8
3. Elementary background	9
4. Pointed spaces and absorption	10
4.1. Pointed spaces	11
4.2. The composite setup	12
4.3. Absorption and W_1 seminorms	14
5. Algebras, unital maps and ray-wise boundedness	16
5.1. Ray-wise boundedness in an algebraic framework	16
5.2. The composite algebraic framework and W_1 norms	19
6. C*-algebras and complete positivity	19
7. Subsystems and additivity	23
7.1. Pointed W_1 substructures and their W_1 seminorms	23
7.2. Reducible pointed W_1 structures	25
7.3. Additivity	29
7.4. C*-algebras	30
8. Outlook	33
References	34

Date: 2023-3-8.

2020 Mathematics Subject Classification. 81P47, 81S22, 49Q22, 46L60.

1. INTRODUCTION

This paper is devoted to devising a distance between quantum channels which is natural in the context of composite quantum systems. Our approach is motivated by a recently introduced quantum Wasserstein distance of order 1 between states of a composite system with finite dimensional state space [22], as a natural and very effective distinguishability measure with a number of desirable properties.

The method that [22] uses to set up their quantum Wasserstein distance of order 1, will be referred to as the DMTL approach, and the resulting metric can be viewed as a quantum version of Ornstein's \bar{d} distance [43]. While we use some of the core ideas of the DMTL approach, the extension to infinite dimensions and to channels, involves significant changes in the approach, including concepts and techniques not present in the finite dimensional case for states. Although the term Wasserstein distance is typically used for states (or probability distributions in the classical case), we nevertheless continue to use the term in the case of channels as well.

The infinite dimensional setting of this paper is expressed in terms of C*-algebras as a generalization of the matrix algebras M_d . We obtain a metric on the set of all channels from one composite system to another, where both consist of n systems. The case of states is obtained when the channels are taken to map to the complex numbers. In particular, the DMTL approach is recovered when restricting to finite dimensions, specifically to the algebras M_d .

The DMTL approach starts with the concept of neighbouring states, a form of which was introduced earlier in [1]. Broadly speaking, our approach is to extend the idea of neighbouring states to neighbouring channels. This is done through the reduction of channels. Given a channel from one composite system to another, choose the j 'th system in each, and reduce the channel to the remaining systems. If two channels give the same channel after this reduction for some j is applied to both, we consider them to be neighbouring channels. I.e., two channels are called neighbouring if they coincide after the removal of two corresponding systems from the two composite systems respectively. The neighbouring channels are then used to build a unit ball leading to a norm on a certain space of linear maps and ultimately to a distance between any two channels. More precisely, we obtain a metric on the set of channels from the one composite system to the other. All this is illustrated concretely in finite dimensions in the next section.

One of our intentions is to highlight the general mathematical structure behind the quantum Wasserstein distance of order 1, through an abstract approach. This involves first setting up the theory in general vector spaces, and subsequently algebras, without reference to any positivity conditions on the maps between these spaces. The C*-algebraic

framework, with the maps taken as channels (unital completely positive maps between C^* -algebras), is obtained as a special case of the abstract setup. This clarifies the overall structure of our approach, in particular, where and how complete positivity is used, and potentially allows for cases other than C^* -algebras and channels between them.

We nevertheless also discuss the finite dimensional case, particularly in the next section, to clarify the basic approach and ideas. The C^* -algebraic framework is a generalization of the finite dimensional case, and very natural and relevant from the view of quantum physics. The paper is written in such a way that the main thread and results in the C^* -algebraic case can be followed without going through the more abstract approach mentioned above. The finite dimensional situation is in turn obtained as a simple special case of the C^* -algebraic framework. The proofs of our results do depend on the abstract development, though.

The next section gives an outline of our approach in finite dimensions, expressed in a form that allows clear comparison to the DMTL approach. It can be viewed as an extension of this introduction, explaining some of the goals and motivation for this paper, but it also serves as an overview for readers whose main interest is the finite dimensional case. Section 3 reviews basic background related to norms and unit balls, and takes an initial step in developing the theory behind the Wasserstein distance of order 1 between channels. Sections 4 and 5 develop the abstract theory to obtain our Wasserstein distance of order 1. The reader can in fact skip these two sections upon initial reading, and go directly to Section 6. There the definition and main results leading to the Wasserstein distance of order 1 between channels in the C^* -algebraic framework, are presented with no reference to Sections 4 and 5, although the proofs depend on those two sections. Note that we'll usually simply say "Wasserstein distance" instead of "quantum Wasserstein distance", since our theory contains the special case of abelian C^* -algebras.

The paper then proceeds to the behaviour of the Wasserstein distance of order 1 with respect to subsystems of the composite systems, i.e., when we consider smaller composite systems consisting of a subset of the original systems. The main result here is that this Wasserstein distance is additive over tensor products of channels between such subsystems, with stability as a special case. Again this result is first approached abstractly in general vector spaces, before the C^* -algebraic case is presented. As before, the reader can page directly to Subsection 7.4 after Section 6, to see the C^* -algebraic results with a minimum of direct reference to the abstract theory, although the proofs again rely on the latter.

To conclude this introduction, we briefly discuss previous work.

The study of quantum channels, including mathematical techniques to analyze them theoretically, remains undeniably important. Refer for example to [11, 34, 39] for an overview of a variety of aspects of quantum channels and their significance. In particular, there have been other approaches to distances between channels, in relation to channel discrimination. For example, the diamond norm, also called the completely bounded trace norm; see [38, 3] for early work in connection to quantum computation, and [50, Section 3.3] for a more general finite dimensional overview in the context of quantum information. Refer to [29] for a broader perspective on distances between channels.

On the other hand, there has also been much effort to obtain quantum (or noncommutative) Wasserstein distances between states. Papers obtaining Wasserstein distances (with a focus on order 2) which are actual metrics on sets of quantum states, include [6, 8, 9, 10, 12, 14, 16, 26, 35, 51]. These papers follow different approaches from this paper and [22], and to a large extent from one another. Broadly speaking there have been two main approaches, in analogy to the classical case: a coupling (or transport plan) approach [6, 16, 26] and a dynamical approach [8, 9, 10, 12, 14, 35, 51]. The book [48] includes a nice introduction to the former, while [4] is the origin of the latter. The DMTL approach and this paper, can roughly be classified as part of the coupling approach. Other papers using the coupling approach, though not obtaining all the usual metric properties, include [23, 30, 31, 32]. Papers treating various other approaches than [22] to noncommutative or quantum Wasserstein distances of order 1, include [2, 10, 13, 28, 42, 45, 47, 49]. Some of these use a dual formulation of the coupling approach, which is closely related to Connes' spectral distance in noncommutative geometry, first introduced in [17], and studied further in [18, 19, 20, 44], among others.

The history of the classical theory of optimal transport goes back much further, and is too long and involved to recount here. We restrict ourselves to brief remarks regarding terminology. Although the term Wasserstein distance has become fairly standard in the literature, which is the reason we use it, it does not fully reflect the history of optimal transport. Indeed, other terms, in particular Kantorovich distance, Kantorovich-Rubinstein distance and the earth mover's distance, are also used. These refer to earlier pioneers of these types of distances, and to the historical origins of the field.

It seems that a theory of Wasserstein distances, by any approach, has not before been extended from states to channels in the literature. This is of course aside from attempting to apply the former directly to the latter via the Choi-Jamiołkowski duality between states and channels. This duality holds under quite general conditions and could be applied beyond just the finite dimensional case (see in particular [27, Section 3]), though to the best of our knowledge not in the generality of this

paper's framework. Furthermore, the intention of this paper is rather to systematically build the theory directly for channels. As is outlined in Section 8, even in finite dimensions this is not equivalent to merely translating via the Choi-Jamiołkowski duality.

2. OUTLINE IN FINITE DIMENSIONS

In this section we briefly outline the paper's approach and core results in finite dimensions, using some notation and concepts from physics, to make the basic ideas as accessible as possible to a wide audience. To clarify how this relates to the DMTL approach, we use a formulation analogous to theirs.

To do this, we in particular make use of the Choi-Jamiołkowski duality between channels and states. To avoid confusion, however, it is important to note from the outset that the latter states will not be employed as the states in the DMTL approach. That is to say, we do not simply apply their results to states representing channels via the Choi-Jamiołkowski duality. Rather, the DMTL approach is systematically adapted directly to channels, while the representation of channels via Choi-Jamiołkowski duality simply emphasizes the connection to the DMTL approach, as will be seen below.

In the general theory presented in later sections, on the other hand, it will be more convenient to avoid the Choi-Jamiołkowski duality, instead working with the channels directly. The finite dimensional setup discussed in this section emerges as a special case of that general theory. The reader who wants to focus on the general mathematical development, can skip ahead to Section 3.

2.1. A representation of channels. Consider a channel \mathcal{E} from the density matrices of a system \mathcal{B} to that of a system \mathcal{A} . Fix an orthonormal basis $|1^{\mathcal{B}}\rangle, \dots, |r^{\mathcal{B}}\rangle$ for the Hilbert space $H_{\mathcal{B}}$ of the system \mathcal{B} , and define a density matrix $\kappa_{\mathcal{E}}$ of the composite system \mathcal{BA} by

$$(1) \quad \kappa_{\mathcal{E}} = \frac{1}{r} \sum_{i=1}^r \sum_{j=1}^r |i^{\mathcal{B}}\rangle \langle j^{\mathcal{B}}| \otimes \mathcal{E}(|i^{\mathcal{B}}\rangle \langle j^{\mathcal{B}}|).$$

This density matrix reduces to the maximally mixed state of \mathcal{B} . The Choi-Jamiołkowski duality [15] tells us that there is a one-to-one correspondence between the set of such channels \mathcal{E} , and the set of density matrices of \mathcal{BA} reducing to the maximally mixed state of \mathcal{B} . (Also see [25, 36], but Choi's approach [15] forms the basis for our use of the Choi-Jamiołkowski duality in this section.)

Represent the observable algebra of \mathcal{B} in terms of the given basis as the matrix algebra $B = M_r$, and similarly use $A = M_q$ for \mathcal{A} . To keep the distinction between states and observables conceptually and notationally clear, we also write $S_{\mathcal{A}} = M_q$ for the matrix algebra containing the density matrices of \mathcal{A} , and similarly $S_{\mathcal{B}} = M_r$ for \mathcal{B} . In

particular, the channel \mathcal{E} is then a completely positive trace preserving linear map

$$\mathcal{E} : S_{\mathcal{B}} \rightarrow S_{\mathcal{A}}$$

from $S_{\mathcal{B}}$ to $S_{\mathcal{A}}$. This channel's dual representation in terms of the observable algebras is the unital completely positive linear map

$$E : A \rightarrow B$$

defined via

$$\text{Tr}[E(a)b] = \text{Tr}[a\mathcal{E}(b)]$$

for all $a \in A$ and $b \in S_{\mathcal{B}}$.

Consider the set $K(A, B)$ of channels $E : A \rightarrow B$, and the set

$$C(\mathcal{B}\mathcal{A}) = \{\delta \in B \otimes S_{\mathcal{A}} : \delta \geq 0 \text{ and } \text{Tr}_{\mathcal{A}} \delta = 1_B\},$$

(not $\delta \in S_{\mathcal{B}} \otimes S_{\mathcal{A}}$; see below) where $\text{Tr}_{\mathcal{A}}$ denotes the partial trace over the system \mathcal{A} , and 1_B is the identity matrix in $B = M_r$. Implicit to the Choi-Jamiołkowski duality, and fairly straightforward to extract from it, we have a one-to-one correspondence between $K(A, B)$ and $C(\mathcal{B}\mathcal{A})$ given by

$$(2) \quad E(a) = \text{Tr}_{\mathcal{A}}[\delta(1_B \otimes a)].$$

Denoting the dual of E by $\mathcal{E} : S_{\mathcal{B}} \rightarrow S_{\mathcal{A}}$, we note that in terms of the Choi-Jamiołkowski duality as stated above, one has

$$\delta = r\kappa_{\mathcal{E}}^{\text{T}_{\mathcal{B}}}$$

in this one-to-one correspondence, where $\text{T}_{\mathcal{A}}$ is the partial transposition over \mathcal{B} with respect to the chosen basis for $H_{\mathcal{B}}$. Because of the usual interpretation of the Choi-Jamiołkowski duality, one might be tempted to rather view the elements of $C(\mathcal{B}\mathcal{A})$ as $\delta \in S_{\mathcal{B}} \otimes S_{\mathcal{A}}$ such that δ/r is a density matrix of the composite system $\mathcal{B}\mathcal{A}$ reducing to the maximally mixed state of \mathcal{B} , but strictly speaking $E(a)$ as given by (2) would then be in $S_{\mathcal{B}}$ instead of B . When $r > 1$, it is therefore in fact conceptually better if we do not view δ/r as a density matrix representing some state. On the other hand, in the special case where \mathcal{B} is a trivial system, i.e., $r = 1$ and $B = \mathbb{C}$, the set $C(\mathcal{B}\mathcal{A})$ is precisely all the density matrices of \mathcal{A} , and we simply recover the usual representation of expectation values of a state, $E(a) = \text{Tr}(\delta a)$, in terms of the density matrix δ .

2.2. The W_1 norm. Using this representation of channels as elements of $C(\mathcal{B}\mathcal{A})$, we can formulate an extension of the Wasserstein distance of order 1 between states in the DMTL approach, to channels. This is a distance between channels acting from one composite system, \mathcal{A} , to another, \mathcal{B} . Here we assume that

$$A = A_1 \otimes \dots \otimes A_n \text{ and } B = B_1 \otimes \dots \otimes B_n$$

with $A_j = M_{q_j}$ and $B_j = M_{r_j}$. The latter are simply the observable algebras of systems \mathcal{A}_j and \mathcal{B}_j respectively. One then defines the real vector space

$$(3) \quad \mathcal{O} = \{X \in \text{span}_{\mathbb{R}} C(\mathcal{BA}) : \text{Tr}_{\mathcal{A}} X = 0\},$$

where $\text{span}_{\mathbb{R}}$ refers to finite linear combinations with real coefficients. The goal is to define a certain norm $\|\cdot\|_{W_1}$ on \mathcal{O} , which when applied to differences $\delta - \varepsilon$ of elements of $C(\mathcal{BA})$, will in turn define a metric on $C(\mathcal{BA})$. This metric will be the Wasserstein distance of order 1 on $C(\mathcal{BA})$, or equivalently on $K(A, B)$ via the one-to-one correspondence above, extending the construction in the DMTL approach to channels.

The definition of the norm $\|\cdot\|_{W_1}$ entails extending the idea of neighbouring states in the DMTL approach to channels. We write

$$A_{\hat{j}} = A_1 \otimes \dots \otimes \widehat{A_j} \dots \otimes A_n \quad \text{and} \quad B_{\hat{j}} = B_1 \otimes \dots \otimes \widehat{B_j} \dots \otimes B_n,$$

i.e., these are A and B with A_j and B_j respectively left out of the tensor products. Similarly we set

$$A_{\leq j} = A_1 \otimes \dots \otimes A_j \quad \text{and} \quad A_{\geq j} = A_j \otimes \dots \otimes A_n,$$

and likewise for $B_{\leq j}$ and $B_{\geq j}$. In terms of the following notation (which is chosen to fit in with that of later sections),

$$\nu_j = \frac{1}{r_j} \text{Tr}$$

where this Tr is the usual trace on $B_j = M_{r_j}$, we can then reduce a channel $E : A \rightarrow B$ to the channel

$$E_{\hat{j}} = (\text{id}_{B_{\leq j-1}} \otimes \nu_j \otimes \text{id}_{B_{\geq j+1}}) \circ E|_{A_{\hat{j}}}$$

from $A_{\hat{j}}$ to $B_{\hat{j}}$, where id denotes the identity map on the indicated algebra, and with $E|_{A_{\hat{j}}}$ defined via

$$E|_{A_{\hat{j}}}(a_1 \otimes \dots \otimes \widehat{a_j} \dots \otimes a_n) = E(a_1 \otimes \dots \otimes a_{j-1} \otimes 1_{A_j} \otimes a_{j+1} \otimes \dots \otimes a_n).$$

I.e., we restrict E to $A_{\hat{j}}$, and evaluate the ‘‘partial expectation’’ of the result over B_j . Note that via the one-to-one correspondence given by (2), this reduction of E is equivalent to the reduction

$$\delta_{\hat{j}} = (\text{id}_{B_{\leq j-1}} \otimes \nu_j \otimes \text{id}_{B_{\geq j+1}}) \otimes \text{Tr}_{\mathcal{A}_j}$$

of the corresponding $\delta \in C(\mathcal{BA})$, where $\text{Tr}_{\mathcal{A}_j}$ denotes the partial trace on $S_{\mathcal{A}} = M_{q_1} \otimes \dots \otimes M_{q_n}$ over $S_{\mathcal{A}_j} = M_{q_j}$. In terms of this notation, we view $\delta, \varepsilon \in C(\mathcal{BA})$ as representing neighbouring channels when $\delta_{\hat{j}} = \varepsilon_{\hat{j}}$ for some j . We define

$$(4) \quad \mathcal{N}_j = \{\delta - \varepsilon : \delta, \varepsilon \in C(\mathcal{BA}) \text{ with } \delta_{\hat{j}} = \varepsilon_{\hat{j}}\}$$

and

$$\mathcal{N} = \bigcup_{j=1}^n \mathcal{N}_j.$$

We then define $\|\cdot\|_{W_1}$ as the norm on \mathcal{O} which has the convex hull

$$\mathcal{C} = \text{conv } \mathcal{N}$$

of \mathcal{N} as its unit ball. I.e.,

$$(5) \quad \|X\|_{W_1} = \inf\{t \geq 0 : X \in t\mathcal{C}\}$$

where $t\mathcal{C} = \{tX : X \in \mathcal{C}\}$ for any real number t . We write this as an infimum, rather than a minimum, to conform to our more general approach later on. The norm $\|\cdot\|_{W_1}$ will be referred to as the W_1 norm. To prove that this is a norm on \mathcal{O} of course requires some work, which will be done in a more general context in the sequel.

2.3. Wasserstein distance of order 1 between channels. Given this norm, we can define a metric W_1 on $K(A, B)$ via

$$W_1(E_\delta, E_\varepsilon) = \|\delta - \varepsilon\|_{W_1}$$

with E_δ denoting the channel E corresponding to $\delta \in C(\mathcal{BA})$ in (2). This metric W_1 is the generalization of the Wasserstein distance of order 1 between states to the case of channels. We consequently refer to it as the *Wasserstein distance of order 1* on $K(A, B)$.

The rough intuition behind this metric follows from δ and ε in the definition of \mathcal{N}_j above being neighbouring channels. This condition tells us that ε and δ coincide when reduced to $A_{\hat{j}} \rightarrow B_{\hat{j}}$ for some j , i.e., with one system, \mathcal{A}_j and \mathcal{B}_j respectively, removed from each of the composite systems \mathcal{A} and \mathcal{B} . In this way the ‘‘local differences’’ between two channels E_δ and E_ε are picked up by W_1 , where ‘‘local’’ here is simply in relation to the systems $\mathcal{A}_1, \dots, \mathcal{A}_n$ and $\mathcal{B}_1, \dots, \mathcal{B}_n$ composing \mathcal{A} and \mathcal{B} . A typical case is the dynamics of an open composite system \mathcal{A} , where we take $\mathcal{B}_j = \mathcal{A}_j$ for all j . We then expect W_1 to naturally take into account the differences between two dynamical processes in the individual systems \mathcal{A}_j .

A basic property of W_1 is additivity with respect to tensor products. If we partition the set $[n] = \{1, \dots, n\}$ into m non-empty and sequential parts, i.e.,

$$\begin{aligned} P(1) &= 1, \dots, n_1 \\ P(2) &= n_1 + 1, \dots, n_2 \\ &\vdots \\ P(m) &= n_{m-1} + 1, \dots, n, \end{aligned}$$

then we can consider the subsystems of \mathcal{A} and \mathcal{B} with observable algebras

$$A_{P(k)} = \bigotimes_{j \in P(k)} A_j \quad \text{and} \quad B_{P(k)} = \bigotimes_{j \in P(k)} B_j$$

respectively. For any channels $D_k, E_k : A_{P(k)} \rightarrow B_{P(k)}$ we then have

$$W_1(D_1 \otimes \dots \otimes D_m, E_1 \otimes \dots \otimes E_m) = \sum_{k=1}^m W_1(D_k, E_k),$$

with a resulting stability property when $D_j = E_j$ for some of the j 's. This can be refined by dropping the assumption that the partition is sequential, but the form above is for the moment notationally clearer. Keep in mind that W_1 on $K(A_{P(k)}, B_{P(k)})$ is of course defined by the same procedure as for $K(A, B)$.

When $\mathcal{B}_j = \mathbb{C}$, we recover the case of states on \mathcal{A} , and indeed our W_1 above then specializes to the Wasserstein distance of order 1 in the DMTL approach, with $\delta, \varepsilon \in C(\mathcal{B}\mathcal{A})$ becoming density matrices ρ and σ of the composite system \mathcal{A} . They specifically considered the case $q_1 = \dots = q_n = d$.

In the subsequent general theory, the formulation will be in the Heisenberg picture $E : A \rightarrow B$ from the outset. The Choi-Jamiołkowski duality will also be side-stepped, with the formulation expressed directly in terms of the channels themselves. In connection to this, note that the basic condition $\text{Tr}_{\mathcal{A}} X = 0$ in (3) can equivalently be expressed as $\lambda(1_A) = 0$, with $\lambda : A \rightarrow B$ defined by

$$\lambda(a) = \text{Tr}_{\mathcal{A}}[X(1_B \otimes a)]$$

in terms of the given X . The latter extends the formula in (2). The condition $\lambda(1_A) = 0$ relates to the unitality of channels, namely $E(1_A) = 1_B$, as will be seen in abstract form in Sections 4 and 5.

The reader may now turn directly to Section 6 and 7.4 to see the general C^* -algebraic version of this section. However, the proofs that the Wasserstein distance of order 1 treated there is indeed a metric and satisfies additivity, rely on Sections 4 and 5 as well as the rest of 7.

3. ELEMENTARY BACKGROUND

This section reviews and sets up some basic notions and elementary results to be used later on. This is for easy reference and since references providing this material in exactly the form we need appear to be scarce. Presentations of this material, though in somewhat different forms than what we use, can be found in [5, Section I.1] and [41, Chapters 4 and 5]. The results in this section are stated for real vector spaces, since this is exactly what we need later on, as can already be seen from (3). Note that we do not assume any of the vector spaces to be finite dimensional.

The basic result is how a norm arises from a given set which is to serve as a unit ball for the norm.

Definition 3.1. Consider a subset \mathcal{C} of a real vector space \mathcal{X} .

(1) The set \mathcal{C} is called *absorbing (for \mathcal{X})* if for every $x \in \mathcal{X}$ there is a number $t \geq 0$ such that $x \in t\mathcal{C}$, where $t\mathcal{C} = \{tx : x \in \mathcal{C}\}$. The *gauge* $\|\cdot\|_{\mathcal{C}} : \mathcal{X} \rightarrow \mathbb{R}_+$ of such an absorbing set \mathcal{C} is the function on \mathcal{X} defined by

$$\|x\|_{\mathcal{C}} := \inf\{t \geq 0 : x \in t\mathcal{C}\}$$

for every $x \in \mathcal{X}$, where $\mathbb{R}_+ := \{t \in \mathbb{C} : t \geq 0\}$.

(2) The set \mathcal{C} is called *symmetric* if $\mathcal{C} = -\mathcal{C}$.

(3) The set \mathcal{C} is called *ray-wise (or radially) bounded (in \mathcal{X})* if for every non-zero $x \in \mathcal{X}$ there is a number $s_0 > 0$ such that $sx \notin \mathcal{C}$ for all $s > s_0$.

(4) If \mathcal{C} is absorbing, convex and symmetric, then we call it a *semi unit ball for \mathcal{X}* . If in addition \mathcal{C} is ray-wise bounded, then we call it a *unit ball for \mathcal{X}* .

The following simple result is the method by which we intend to define our norms, as was seen in the special case (5).

Proposition 3.2. *Let \mathcal{C} be a semi unit ball for a real vector space \mathcal{X} . Then its gauge $\|\cdot\|_{\mathcal{C}}$ is a seminorm on \mathcal{X} . If \mathcal{C} is a unit ball for \mathcal{X} , then $\|\cdot\|_{\mathcal{C}}$ is a norm on \mathcal{X} .*

Results of this type are well known and straightforward to prove. That \mathcal{C} is absorbing, ensures that $\|\cdot\|_{\mathcal{C}}$ is well defined and finite at every point of \mathcal{X} , convexity implies the triangle inequality, symmetry leads to $\|\alpha x\|_{\mathcal{C}} = |\alpha| \|x\|_{\mathcal{C}}$ for all $\alpha \in \mathbb{R}$ (these three properties making $\|\cdot\|_{\mathcal{C}}$ a seminorm), while ray-wise boundedness enforces $x = 0$ when $\|x\|_{\mathcal{C}} = 0$. The harder problem is to define an appropriate set \mathcal{C} in an operator algebraic context (generalizing the previous section) and prove that it is indeed a unit ball, which is exactly the focus in the next three sections. In the meantime we note an elementary result which will be relevant in the next section when proving the absorbing property.

Lemma 3.3. *Let \mathcal{X} be a real vector space. Consider any symmetric subset \mathcal{N} of \mathcal{X} such that $\mathcal{X} = \text{span } \mathcal{N}$. Then its convex hull $\mathcal{C} = \text{conv } \mathcal{N}$ is absorbing for \mathcal{X} .*

Proof. For any non-zero $x \in \mathcal{X}$ we can write $x = s_1x_1 + \dots + s_kx_k$ for some $k < \infty$, $x_j \in \mathcal{N}$ and $s_j > 0$. Let $t = s_1 + \dots + s_k$ and $p_j = s_j/t$ to have $x = ty$ with $y = p_1x_1 + \dots + p_kx_k \in \mathcal{C}$. The case $x = 0$ is trivial. \square

4. POINTED SPACES AND ABSORPTION

The absorbing property in Definition 3.1 is needed for the function $\|\cdot\|_{\mathcal{C}}$ on \mathcal{X} in Proposition 3.2 to be well defined and finite at every point of \mathcal{X} , and is in that sense the most basic property. In this section we show how it arises under general assumptions for the set to be

used as a unit ball for the W_1 norm. We are ultimately interested in the situation where \mathcal{X} is a space of maps containing the channels between two composite systems. In this section \mathcal{X} is more generally taken as a vector space \mathcal{O} of linear maps between two vector spaces, each with a distinguished point. The maps will be required to map the one distinguished point to the other, as an abstraction of unitality. Actual unitality, in the case of unital algebras as the pointed spaces, will be treated in the next section, while the C*-algebraic case, for channels, follows in Section 6.

4.1. Pointed spaces. By a *pointed space* (A, u_A) we simply mean a real or complex vector space A with a distinguished non-zero element $u_A \in A$. For simplicity of notation, a pointed space (A, u_A) will be denoted as A , with u_A assumed as the notation for the distinguished point.

Keep in mind that eventually (Section 6) the pointed spaces will be taken to be (complex) unital C*-algebras, with the units serving as the distinguished points. These C*-algebras will generalize the matrix algebras from Section 2.

In the remainder, all pointed spaces involved are assumed to be over the same scalars, either real or complex. However, in either case certain constructions will involve the span of a subset over real scalars, leading to a real vector space, and such spans will be indicated by $\text{span}_{\mathbb{R}}$.

For any two vector spaces A and B over the same scalars, the space of all linear maps $\lambda : A \rightarrow B$ will be denoted by $L(A, B)$. When A and B are pointed spaces, we define

$$L_u(A, B) = \{\eta \in L(A, B) : \eta(u_A) = u_B\},$$

which is the set of *pointed maps* from A to B . This is of course an abstract version of unital maps in the case where A and B are unital algebras.

The following result will shortly be used in tandem with Lemma 3.3 to prove the absorbing property of certain sets.

Lemma 4.1. *Let A and B be pointed spaces and consider any subset \mathcal{L} of $L_u(A, B)$. Set*

$$\mathcal{O} := \{\lambda \in \text{span}_{\mathbb{R}} \mathcal{L} : \lambda(u_A) = 0\} \quad \text{and} \quad \mathcal{V} := \{\eta - \theta : \eta, \theta \in \mathcal{L}\}.$$

Then it follows that $\mathcal{O} = \text{span}_{\mathbb{R}} \mathcal{V}$.

Proof. Clearly $\text{span}_{\mathbb{R}} \mathcal{V} \subset \mathcal{O}$. Conversely, take any non-zero $\lambda \in \mathcal{O}$ and write $\lambda_1 := \lambda$. Since $\lambda_1(u_A) = 0$ and $u_B \neq 0$, the coefficients in λ_1 's expansion as a linear combination of \mathcal{L} 's elements, can not all be strictly positive or all strictly negative. Without loss of generality we can therefore rewrite the expansion in the form

$$\lambda = r_1^{(1)} \eta_1^{(1)} + \dots + r_k^{(1)} \eta_k^{(1)} - s_1^{(1)} \theta_1^{(1)} - \dots - s_k^{(1)} \theta_k^{(1)}$$

where $\eta_j^{(1)}, \theta_j^{(1)} \in \mathcal{L}$ and $r_j^{(1)}, s_j^{(1)} > 0$ for some (finite) $k^{(1)}$, where the superscript (1) is merely an index, and with the proviso that $\lambda(u_A) = 0$. Note that $k^{(1)}$ is simply the biggest of the number of positive coefficients and the number of negative coefficients in λ_1 's initial expansion, while the coefficients in smaller number are split to increase their number to $k^{(1)}$. Writing $\lambda'_1 = r_1^{(1)}(\eta_1^{(1)} - \theta_1^{(1)}) + \dots + r_{k^{(1)}}^{(1)}(\eta_{k^{(1)}}^{(1)} - \theta_{k^{(1)}}^{(1)}) \in \text{span}_{\mathbb{R}} \mathcal{V}$ and $\lambda_2 = (r_1^{(1)} - s_1^{(1)})\theta_1^{(1)} + \dots + (r_{k^{(1)}}^{(1)} - s_{k^{(1)}}^{(1)})\theta_{k^{(1)}}^{(1)}$, we have $\lambda_1 = \lambda'_1 + \lambda_2$, implying that $\lambda_2 \in \mathcal{O}$. If $\lambda_2 \neq 0$, then repeat this procedure for λ_2 instead of λ_1 , noting that we can now analogously write

$$\lambda_2 = r_1^{(2)}\eta_1^{(2)} + \dots + r_{k^{(2)}}^{(2)}\eta_{k^{(2)}}^{(2)} - s_1^{(2)}\theta_1^{(2)} - \dots - s_{k^{(2)}}^{(2)}\theta_{k^{(2)}}^{(2)},$$

but with $k^{(2)} < k^{(1)}$. This delivers $\lambda'_2 \in \text{span}_{\mathbb{R}} \mathcal{V}$ and $\lambda_3 \in \mathcal{O}$. If $\lambda_3 \neq 0$, then repeat for λ_3 , etc. Since $k^{(j+1)} < k^{(j)}$, this process must stop to deliver $\lambda = \lambda'_1 + \dots + \lambda'_m \in \text{span}_{\mathbb{R}} \mathcal{V}$. \square

Note that the lemma is geared towards spaces such as that appearing in (3), but in the Heisenberg picture.

4.2. The composite setup. Here we set up a framework which will serve as an abstraction of composite systems in terms of pointed spaces. We also define the set for which the absorbing property will be proven in the next subsection, leading to it being a semi unit ball giving a W_1 seminorm.

Consider pointed spaces A_1, \dots, A_n and B_1, \dots, B_n . Let

$$A = A_1 \odot \dots \odot A_n$$

be the algebraic tensor product of the vector spaces A_1, \dots, A_n , which is itself a pointed space with distinguished point

$$u_A := u_{A_1} \otimes \dots \otimes u_{A_n}.$$

However, as we want to allow for completions of A , in particular in the C^* -algebraic framework of Section 6 (where A_1, \dots, A_n will be unital C^* -algebras), we need to allow for completions of the tensor product of B_1, \dots, B_n from the outset. To emphasize this, we write

$$B = B_1 \otimes \dots \otimes B_n,$$

which may be the algebraic tensor product, or some completion of it with respect to a specified norm. In all cases B is a pointed space with $u_B = u_{B_1} \otimes \dots \otimes u_{B_n}$. The particular tensor product B remains fixed throughout this section, however. The point of this is that in the theory developed here, linear maps $A \rightarrow B$ are then not constrained to have their image contained in an uncompleted tensor product.

Along the lines of Section 2, we define

$$\begin{aligned} A_{\hat{j}} &= A_1 \odot \dots \widehat{A_j} \dots \odot A_n \quad \text{and} \quad B_{\hat{j}} = B_1 \otimes \dots \widehat{B_j} \dots \otimes B_n, \\ A_{\leq j} &= A_1 \odot \dots \odot A_j \quad \text{and} \quad A_{\geq j} = A_j \odot \dots \odot A_n, \end{aligned}$$

and

$$B_{\leq j} = B_1 \otimes \dots \otimes B_j \quad \text{and} \quad B_{\geq j} = B_j \otimes \dots \otimes B_n,$$

for $j = 1, \dots, n$, using the same completion (if relevant) for tensor products of the B_j 's as for B .

Let ν_j be a linear functional on B_j such that

$$\nu_j(u_{B_j}) = 1$$

for $j = 1, \dots, n$. These functionals will act as reference functionals relative to which linear maps $A \rightarrow B$ will be reduced to $A_{\widehat{j}} \rightarrow B_{\widehat{j}}$. Later on (Section 6), in the case of unital C^* -algebras, the ν_j 's will be taken as states, generalizing the normalized traces used in Section 2.

If B is indeed a completion with respect to some norm, rather than just an algebraic tensor product, we also assume that algebraic tensor products of the ν_j 's with themselves as well as with the identity maps $\text{id}_{B_{\leq j-1}}$ and $\text{id}_{B_{\geq j+1}}$ are continuous with respect to this norm and therefore uniquely extendible to the completed tensor products. The relevant reductions of a linear map $\eta \in L(A, B)$ from A to B can then be defined as

$$(6) \quad \eta_{\widehat{j}} = (\text{id}_{B_{\leq j-1}} \otimes \nu_j \otimes \text{id}_{B_{\geq j+1}}) \circ \eta|_{A_{\widehat{j}}} : A_{\widehat{j}} \rightarrow B_{\widehat{j}}$$

$$(7) \quad \eta_{\leq j} = (\text{id}_{B_{\leq j}} \otimes \nu_{> j}) \circ \eta|_{A_{\leq j}} : A_{\leq j} \rightarrow B_{\leq j}$$

$$(8) \quad \eta_{\geq j} = (\nu_{< j} \otimes \text{id}_{B_{\geq j}}) \circ \eta|_{A_{\geq j}} : A_{\geq j} \rightarrow B_{\geq j}$$

for $j = 1, \dots, n$, with $\nu_{< j} = \nu_1 \otimes \dots \otimes \nu_{j-1}$ (empty when $j = 1$) and $\nu_{> j} = \nu_{j+1} \otimes \dots \otimes \nu_n$ (empty when $j = n$), where all the indicated tensor product maps are algebraic when the tensor product B is, or (uniquely) continuously extended to B as assumed possible above, when B is completed. Here the restriction $\eta|_{A_{\widehat{j}}} : A_{\widehat{j}} \rightarrow B$ is defined via

$$\eta|_{A_{\widehat{j}}}(a_1 \otimes \dots \otimes \widehat{a_j} \dots \otimes a_n) = \eta(a_1 \otimes \dots \otimes a_{j-1} \otimes u_{A_j} \otimes a_{j+1} \otimes \dots \otimes a_n)$$

for $a_1 \in A_1, \dots, a_n \in A_n$ (using the universal property), replacing a_j by u_{A_j} , and analogously for $\eta|_{A_{\leq j}} : A_{\leq j} \rightarrow B$ and $\eta|_{A_{\geq j}} : A_{\geq j} \rightarrow B$, where the notation $\widehat{a_j}$ indicates the absence of a_j in the elementary tensor.

Note in particular that when $\eta \in L_u(A, B)$, it follows directly from these definitions that

$$\eta_{\widehat{j}} \in L_u(A_{\widehat{j}}, B_{\widehat{j}}), \quad \eta_{\leq j} \in L_u(A_{\leq j}, B_{\leq j}) \quad \text{and} \quad \eta_{\geq j} \in L_u(A_{\geq j}, B_{\geq j}),$$

where the property $\nu_j(u_{B_j}) = 1$ has been used. Note that for $\eta, \theta \in L_u(A, B)$ this implies that

$$(9) \quad \theta_{\leq j} \odot \eta_{\geq j+1} \in L_u(A, B)$$

for $j = 1, \dots, n-1$, which will implicitly play a role in Lemma 4.2 below.

We also fix any subset

$$\mathcal{L} \subset L_u(A, B)$$

and let

$$(10) \quad \mathcal{O} := \{\lambda \in \text{span}_{\mathbb{R}} \mathcal{L} : \lambda(u_A) = 0\}$$

as in Lemma 4.1. Our ultimate goal, given sufficient additional structure and assumptions, is to define a metric, namely the Wasserstein distance of order 1, on \mathcal{L} . This will be done by first obtaining a seminorm, and in the next section under further assumptions a norm, on \mathcal{O} .

Let

$$(11) \quad \mathcal{N}_j := \{\eta - \theta : \eta, \theta \in \mathcal{L} \text{ such that } \eta_{\hat{j}} = \theta_{\hat{j}}\},$$

generalizing (4), though now in the Heisenberg picture, and set

$$(12) \quad \mathcal{N} := \bigcup_{j=1}^n \mathcal{N}_j \quad \text{and} \quad \mathcal{C} := \text{conv } \mathcal{N},$$

where \mathcal{C} is an abstract version of the set we ultimately want to use as a unit ball defining the W_1 norm. In this section we settle for a seminorm.

4.3. Absorption and W_1 seminorms. In terms of the setup of the previous subsection, we now show that \mathcal{C} is absorbing for \mathcal{O} . The abstract assumption (13) below, is made in lieu of complete positivity. Keep in mind that because of (9), $\theta_{\leq j} \odot \eta_{\geq j+1}$ in (13) is already a pointed map (which is an abstraction of a unital map). The main technical step is the following lemma.

Lemma 4.2. *Assume that*

$$(13) \quad \theta_{\leq j} \odot \eta_{\geq j+1} \in \mathcal{L}$$

for all $\eta, \theta \in \mathcal{L}$ and $j = 1, \dots, n-1$. For any $\eta, \theta \in \mathcal{L}$ it then follows that $\eta - \theta = \lambda_1 + \dots + \lambda_n$ for some $\lambda_j \in \mathcal{N}_j$. In particular,

$$\mathcal{O} = \text{span}_{\mathbb{R}} \mathcal{N}.$$

Proof. In line with Lemma 4.1, we set $\mathcal{V} := \{\eta - \theta : \eta, \theta \in \mathcal{L}\}$. The case $n = 1$ indeed follows immediately from Lemma 4.1, as we then have $\mathcal{N} = \mathcal{V}$. We can therefore assume $n > 1$. For any $\eta, \theta \in \mathcal{L}$, set

$$\begin{aligned} \lambda_1 &= \eta - \theta_{\leq 1} \odot \eta_{\geq 2} \\ \lambda_2 &= \theta_{\leq 1} \odot \eta_{\geq 2} - \theta_{\leq 2} \odot \eta_{\geq 3} \\ &\vdots \\ \lambda_{n-1} &= \theta_{\leq n-2} \odot \eta_{\geq n-1} - \theta_{\leq n-1} \odot \eta_{\geq n} \\ \lambda_n &= \theta_{\leq n-1} \odot \eta_{\geq n} - \theta. \end{aligned}$$

Then $\eta - \theta = \lambda_1 + \dots + \lambda_n$. If indeed $\lambda_j \in \mathcal{N}_j$, it follows that $\eta - \theta \in \text{span}_{\mathbb{R}} \mathcal{N}$, hence $\mathcal{V} \subset \text{span}_{\mathbb{R}} \mathcal{N}$. From $\mathcal{N} \subset \mathcal{V}$ and Lemma 4.1 we can conclude that $\mathcal{O} = \text{span}_{\mathbb{R}} \mathcal{N}$.

It remains to show that $\lambda_j \in \mathcal{N}_j$. Because of (13), we simply have to check the equality of the reductions as required in (11), i.e., that $(\lambda_j)_{\hat{j}} = 0$ for $j = 1, \dots, n$. To handle all cases at once, set $A_{\leq 0} = \mathbb{C}$, $A_{\geq n+1} = \mathbb{C}$, $\theta_{\leq 0} = \text{id}_{\mathbb{C}}$ and $\eta_{\geq n+1} = \text{id}_{\mathbb{C}}$, making $\eta = \theta_{\leq 0} \odot \eta_{\geq 1}$ and $\theta = \theta_{\leq n} \odot \eta_{\geq n+1}$. For $j = 1, \dots, n$, and any $a_{<j} \in A_{\leq j-1}$ and $a_{>j} \in A_{\geq j+1}$, one has the following direct calculation:

$$\begin{aligned} & (\theta_{\leq j-1} \odot \eta_{\geq j})_{\hat{j}}(a_{<j} \otimes a_{>j}) \\ &= (\text{id}_{B_{\leq j-1}} \otimes \nu_j \otimes \text{id}_{B_{\geq j+1}}) \circ (\theta_{\leq j-1} \odot \eta_{\geq j})(a_{<j} \otimes u_{A_j} \otimes a_{>j}) \\ &= \theta_{\leq j-1}(a_{<j}) \otimes ((\nu_j \otimes \text{id}_{B_{\geq j+1}}) \circ \eta_{\geq j}(u_{A_j} \otimes a_{>j})) \end{aligned}$$

and

$$\begin{aligned} & (\nu_j \otimes \text{id}_{B_{\geq j+1}}) \circ \eta_{\geq j}(u_{A_j} \otimes (\cdot)) \\ &= (\nu_j \otimes \text{id}_{B_{\geq j+1}}) \circ (\nu_{<j} \otimes \text{id}_{B_{\geq j}}) \circ \eta|_{A_{\geq j}}(u_{A_j} \otimes (\cdot)) \\ &= (\nu_{<j+1} \otimes \text{id}_{B_{\geq j+1}}) \circ \eta|_{A_{\geq j+1}} \\ &= \eta_{\geq j+1}, \end{aligned}$$

hence $(\theta_{\leq j-1} \odot \eta_{\geq j})_{\hat{j}} = \theta_{\leq j-1} \odot \eta_{\geq j+1}$.

Similarly $(\theta_{\leq j} \odot \eta_{\geq j+1})_{\hat{j}} = \theta_{\leq j-1} \odot \eta_{\geq j+1}$, thus $(\theta_{\leq j-1} \odot \eta_{\geq j})_{\hat{j}} = (\theta_{\leq j} \odot \eta_{\geq j+1})_{\hat{j}}$, as required. \square

Using this lemma we can now show that \mathcal{C} is indeed absorbing.

Proposition 4.3. *Assuming (13), the set \mathcal{C} in (12) is a semi unit ball (Definition 3.1) for \mathcal{O} given by (10).*

Proof. From (11) and (12) it is clear that $-\mathcal{N} = \mathcal{N}$, i.e., \mathcal{N} is symmetric, hence so is \mathcal{C} , which is also convex by definition. In addition, because of \mathcal{N} 's symmetry, Lemmas 4.2 and 3.3 imply that \mathcal{C} is absorbing for \mathcal{O} . \square

Because of Proposition 3.2, this is sufficient to deliver a seminorm. We summarize this as follows.

Definition 4.4. The structure set up in Subsection 4.2, with \mathcal{L} assumed to satisfy (13), is called a *pointed (W_1, n) structure* or *pointed W_1 structure* (if n is clear from context), and is denoted by the shorthand (A, B, ν, \mathcal{L}) , where $\nu := (\nu_1, \dots, \nu_n)$. The rest of the notation in Subsection 4.2 is then implied. For clarity the space \mathcal{O} and set \mathcal{C} in (10) and (12) can respectively be denoted by

$$\mathcal{O}_{\mathcal{L}} \text{ and } \mathcal{C}_{\mathcal{L}}$$

in this context.

Corollary 4.5. *Given a pointed (W_1, n) structure (A, B, ν, \mathcal{L}) , the function $\|\cdot\|_{W_1}$ on $\mathcal{O}_{\mathcal{L}}$ defined by*

$$\|\lambda\|_{W_1} = \inf\{t \geq 0 : \lambda \in t\mathcal{C}_{\mathcal{L}}\},$$

for every $\lambda \in \mathcal{O}_{\mathcal{L}}$, is a seminorm referred to as the W_1 seminorm associated to (A, B, ν, \mathcal{L}) . Consequently, the function $W_1 : \mathcal{L} \times \mathcal{L} \rightarrow \mathbb{R}$ defined by

$$W_1(\eta, \theta) = \|\eta - \theta\|_{W_1},$$

is a pseudometric on \mathcal{L} .

Although W_1 is at this stage only a pseudometric, which means we may have $W_1(\eta, \theta) = 0$ for $\eta \neq \theta$, it will nevertheless be called the *Wasserstein distance of order 1* associated to (A, B, ν, \mathcal{L}) . We still need ray-wise boundedness to achieve a norm and make W_1 a metric, which is what we turn to next.

5. ALGEBRAS, UNITAL MAPS AND RAY-WISE BOUNDEDNESS

To obtain a norm from the W_1 seminorm in Corollary 4.5, we need ray-wise boundedness. In order to achieve this, we make use of a more specialized algebraic framework as well as assumptions complementary to those made in the previous section. We are going to work in the context of unital algebras. This means that we do not assume the presence of an involution (an adjoint operation) in the algebras. Consequently, as in the previous section, positivity plays no role here, though we again make the abstract assumption (13) which will be implied by the complete positivity of channels in the next section. On the other hand, the unitality of maps will be used.

5.1. Ray-wise boundedness in an algebraic framework. For the moment we step away from the tensor product setup of the previous section, and consider a simple algebraic setting. The core structure (10) will remain in place, however. This allows us to obtain the remaining building block required by Proposition 3.2 in the next result. We return to the tensor product setup in the next subsection.

Note that in the proposition below, all elements of \mathcal{L} are unital maps.

Proposition 5.1. *Let A and B be any unital algebras (both of them real, or both of them complex), with their units 1_A and 1_B respectively serving as the distinguished points making A and B pointed spaces. Fix any subset \mathcal{L} of $L_u(A, B)$ and set*

$$\mathcal{O} := \{\lambda \in \text{span}_{\mathbb{R}} \mathcal{L} : \lambda(1_A) = 0\} \quad \text{and} \quad \mathcal{V} := \{\eta - \theta : \eta, \theta \in \mathcal{L}\}.$$

Consider any subset \mathcal{N} of \mathcal{V} and set $\mathcal{C} := \text{conv } \mathcal{N}$. Then \mathcal{C} is ray-wise bounded in \mathcal{O} .

Proof. In order to decide whether an element of \mathcal{O} lies outside \mathcal{C} , it is going to be convenient to attach a quantitative measure to the element, which when too large, will imply that the element is not in \mathcal{O} . To attain this, we follow a construction inspired by the representation theory of C^* -algebras.

For every $\eta \in \mathcal{L}$, define a bilinear map $\langle \cdot, \cdot \rangle_\eta : A \times A \rightarrow B$ by

$$\langle x, y \rangle_\eta = \eta(xy)$$

for all $x, y \in A$. We write G_η for A equipped with this bilinear map. Define $\pi_\eta : A \rightarrow L(G_\eta)$, with $L(G_\eta)$ the space of linear maps from G_η to itself, through

$$\pi_\eta(a)x = ax$$

for all $a \in A$ and $x \in G_\eta$, with ax simply being the product in A . Then

$$\eta(a) = \langle 1_A, \pi_\eta(a)1_A \rangle_\eta$$

in analogy to a cyclic representation obtained from the GNS construction, where 1_A serves as the ‘‘cyclic vector’’ (indeed, $\pi_\eta(A)1_A = G_\eta$). We proceed to consider the direct sum

$$(G, \langle \cdot, \cdot \rangle, \pi) = \bigoplus_{\eta \in \mathcal{L}} (G_\eta, \langle \cdot, \cdot \rangle_\eta, \pi_\eta).$$

I.e., every $x \in G$ is of the form $x = (x_\eta)_{\eta \in \mathcal{L}}$ with $x_\eta \in G_\eta$ and $\{\eta \in \mathcal{L} : x_\eta \neq 0\}$ a finite set. Furthermore, $\pi : A \rightarrow L(G)$ is defined by $\pi(a)x = (\pi_\eta(a)x_\eta)_{\eta \in \mathcal{L}}$ for all such x and every $a \in A$. Lastly, $\langle x, y \rangle := \sum_{\eta \in \mathcal{L}} \langle x_\eta, y_\eta \rangle_\eta$ for all $x = (x_\eta)_{\eta \in \mathcal{L}}, y = (y_\eta)_{\eta \in \mathcal{L}} \in G$.

For any $\eta \in \mathcal{L}$, define $\hat{\eta} \in G$ by

$$\hat{\eta}_\theta = \begin{cases} 1_A & \text{for } \theta = \eta \\ 0 & \text{for } \theta \neq \eta \end{cases}$$

for all $\theta \in \mathcal{L}$. Note that $\langle \hat{\eta}, \pi(a)\hat{\eta} \rangle = \eta(a)$ for all $\eta \in \mathcal{L}$ and $a \in A$.

Consider any $\lambda \in \mathcal{O}$, which using Lemma 4.1, we can write as

$$\lambda = \sum_{j=1}^l r_j (\eta_j - \theta_j)$$

where $r_j \geq 0$ and $\eta_j, \theta_j \in \mathcal{L}$. Then

$$\lambda = \sum_{j=1}^l r_j \left(\langle \hat{\eta}_j, \pi(\cdot)\hat{\eta}_j \rangle - \langle \hat{\theta}_j, \pi(\cdot)\hat{\theta}_j \rangle \right).$$

We can use this to lift $\lambda : A \rightarrow B$ to a linear map $\bar{\lambda} : L(G) \rightarrow B$ defined by

$$\bar{\lambda}(T) = \sum_{j=1}^l r_j \left(\langle \hat{\eta}_j, T\hat{\eta}_j \rangle - \langle \hat{\theta}_j, T\hat{\theta}_j \rangle \right)$$

for all $T \in L(G)$. This lifting may not be unique (it may depend on the choice of η_j 's and θ_j 's), but for our purposes any such lifting will do. In particular, any $\gamma \in \mathcal{C}$ can be lifted to $\bar{\gamma} : L(G) \rightarrow B$ in the form

$$\bar{\gamma}(T) = \sum_{j=1}^m p_j \left(\langle \hat{\alpha}_j, T\hat{\alpha}_j \rangle - \langle \hat{\beta}_j, T\hat{\beta}_j \rangle \right)$$

for some $\alpha_j, \beta_j \in \mathcal{L}$, with $p_j \geq 0$ and $p_1 + \dots + p_m = 1$.

According to the Hahn-Banach theorem there is a linear functional f on B such that $f(1_B) \neq 0$ (in the case where $1_B = 0$ and therefore $B = \{0\}$, Proposition 5.1 is trivial). Normalize it to obtain

$$\nu = \frac{1}{f(1_B)} f.$$

For any $x, y \in G_\zeta$ we use this to define $x \bowtie y : G_\zeta \rightarrow G_\zeta$ by $(x \bowtie y)z = x\nu(\langle y, z \rangle_\zeta)$ for all $z \in G_\zeta$, where the notation $x \bowtie y$ is inspired by the Dirac notation $|x\rangle\langle y|$. For $x, y \in G$ this in turn allows us to define

$$x \bowtie_\oplus y = \bigoplus_{\zeta \in \mathcal{L}} x_\zeta \bowtie y_\zeta \in L(G),$$

i.e., $(x \bowtie_\oplus y)z = (x_\zeta \nu(\langle y_\zeta, z \rangle_\zeta))_{\zeta \in \mathcal{L}}$ for $z \in G$. For all $\eta, \theta \in \mathcal{L}$ we then have

$$\begin{aligned} \langle \hat{\theta}, (\hat{\eta} \bowtie_\oplus \hat{\eta}) \hat{\theta} \rangle &= \sum_{\zeta \in \mathcal{L}} \left\langle \hat{\theta}_\zeta, \hat{\eta}_\zeta \nu \left(\langle \hat{\eta}_\zeta, \hat{\theta}_\zeta \rangle_\zeta \right) \right\rangle_\zeta \\ &= \begin{cases} 1_B & \text{for } \theta = \eta \\ 0 & \text{for } \theta \neq \eta. \end{cases} \end{aligned}$$

Thus, in terms of $\bar{\gamma}$ above,

$$\bar{\gamma}(\hat{\eta} \bowtie_\oplus \hat{\eta}) = \sum_{j=1}^m p_j b_j$$

where $b_j \in \{-1_B, 0, 1_B\}$ for all j , and consequently

$$(14) \quad |\nu(\bar{\gamma}(\hat{\eta} \bowtie_\oplus \hat{\eta}))| \leq 1$$

for all $\eta \in \mathcal{L}$. This is therefore a condition satisfied by all $\gamma \in \mathcal{C}$, for any lifting $\bar{\gamma}$ of the form above.

On the other hand, for any non-zero $\lambda \in \mathcal{O}$ and a lifting $\bar{\lambda}$ as above, we see that $s\bar{\lambda} := s\bar{\lambda}$ correspondingly lifts $s\lambda$ for any $s > 0$, using sr_j instead of r_j . To simplify the notation in the remainder of the proof, rewrite $\bar{\lambda}$ above as

$$\bar{\lambda}(T) = \sum_{i=1}^k q_i \langle \hat{\zeta}_i, T \hat{\zeta}_i \rangle,$$

where $k = 2l$, $\zeta_j = \eta_j$, $\zeta_{j+k} = \theta_j$, $q_j = r_j$ and $q_{j+k} = -r_j$ for $j = 1, \dots, k$. In this form we may as well assume without loss that $\zeta_i \neq \zeta_{i'}$ when $i \neq i'$ by collecting terms with $\zeta_i = \zeta_{i'}$ if needed. As $\lambda \neq 0$, we have $q_i \neq 0$ for some i . In terms of this we have

$$s\bar{\lambda}(\hat{\zeta}_i \bowtie_\oplus \hat{\zeta}_i) = sq_i 1_B,$$

hence

$$\left| \nu \left(s\bar{\lambda}(\hat{\zeta}_i \bowtie_\oplus \hat{\zeta}_i) \right) \right| = s|q_i|,$$

for all $s > 0$. It follows that there is an $s_0 > 0$ such that

$$\left| \nu \left(s\bar{\lambda}(\hat{\zeta}_i \bowtie_\oplus \hat{\zeta}_i) \right) \right| > 1$$

i.e., $s\lambda \notin \mathcal{C}$ because of (14), for all $s > s_0$, proving that \mathcal{C} is ray-wise bounded according to Definition 3.1(3). \square

With this proposition we have all the elements of the abstract theory in place, which will now allow us to formulate an abstract version of Wasserstein distance of order 1 as a metric, rather than just a pseudo-metric.

5.2. The composite algebraic framework and W_1 norms. We return to the setup in Subsection 4.2, but specialize it as follows.

Definition 5.2. An algebraic (W_1, n) structure (A, B, ν, \mathcal{L}) is a pointed (W_1, n) structure as in Definition 4.4, where the pointed spaces A_1, \dots, A_n and B_1, \dots, B_n are unital algebras (all of them real, or all of them complex), with their units serving as their distinguished points,

$$u_{A_j} = 1_{A_j} \quad \text{and} \quad u_{B_j} = 1_{B_j}$$

for $j = 1, \dots, n$.

From the preceding development we immediately conclude the following.

Theorem 5.3. Consider an algebraic (W_1, n) structure (A, B, ν, \mathcal{L}) . Then $\|\cdot\|_{W_1}$ defined by

$$\|\lambda\|_{W_1} = \inf\{t \geq 0 : \lambda \in t\mathcal{C}_{\mathcal{L}}\}$$

for all $\lambda \in \mathcal{O}_{\mathcal{L}}$, is a norm on $\mathcal{O}_{\mathcal{L}}$, called the W_1 norm associated to (A, B, ν, \mathcal{L}) .

Proof. From Propositions 4.3 and 5.1 we know that $\mathcal{C}_{\mathcal{L}}$ is a unit ball for $\mathcal{O}_{\mathcal{L}}$, as defined in Definition 3.1. By Proposition 3.2 we are done. \square

Corollary 5.4. In terms of Theorem 5.3, the function $W_1 : \mathcal{L} \times \mathcal{L} \rightarrow \mathbb{R}$ defined by

$$W_1(\eta, \theta) = \|\eta - \theta\|_{W_1}$$

is a metric on \mathcal{L} , called the Wasserstein distance of order 1 on \mathcal{L} associated to the algebraic (W_1, n) structure (A, B, ν, \mathcal{L}) .

This theorem and its corollary are the main results up to this point and completes the development without the presence of any form of positivity assumed of the $A \rightarrow B$ maps in \mathcal{L} . So far the maps in \mathcal{L} were only assumed to be linear and unital. In the next section we add complete positivity in a C^* -algebraic framework.

6. C^* -ALGEBRAS AND COMPLETE POSITIVITY

For unital C^* -algebras A and B , let

$$K(A, B)$$

be the set of all channels $E : A \rightarrow B$ from A to B , where a channel is a unital completely positive linear (u.c.p.) map. The goal of this

section is to define a Wasserstein distance of order 1 on $K(A, B)$. Conventionally the term Wasserstein distance applies to states, including (integrals with respect to) probability measures in the classical case, i.e., the case $B = \mathbb{C}$. However, here we use the same terminology for channels as well, as already indicated in Section 2.

It is a fairly straightforward matter to apply Theorem 5.3 in a C*-algebraic framework, essentially taking \mathcal{L} in the previous two sections to be $K(A, B)$, though there are some technical points regarding this which will be made clear in the proof of Theorem 6.2 below. In order to make this section as directly accessible as possible, however, we formulate the definitions and results without reference to Sections 4 and 5. References to these two sections will only appear in the proof of Theorem 6.2. We start with some notation and conventions.

The tensor products of C*-algebras are not merely algebraic, but are completed in some norm. Specifically, the tensor products in this section are either all minimal tensor products or all maximal tensor products. These tensor products will simply be indicated by the symbol \otimes . Accordingly for tensor products of maps on C*-algebras. Some standard background regarding complete positivity and tensor products of C*-algebras can be reviewed in [7, Subsections II.6.9 and II.9.7]. In particular we note that tensor products of channels are again channels, for both the minimal and maximal tensor products.

For easy reference, we highlight the main structure which will be used:

Definition 6.1. Consider unital C*-algebras A_1, \dots, A_n and B_1, \dots, B_n , as well as a state ν_j on B_j for $j = 1, \dots, n$. This will be referred to as an *n-composite C*-system*, denoted $(A_j, B_j, \nu_j : j = 1, \dots, n)$, with the notation

$$A = A_1 \otimes \dots \otimes A_n \quad \text{and} \quad B = B_1 \otimes \dots \otimes B_n$$

being implied.

Given such an *n-composite C*-system*, the following notation, along the lines of Section 2, will be used to set up the Wasserstein norm and subsequent Wasserstein distance of order 1:

$$A_{\hat{j}} := A_1 \otimes \dots \otimes \widehat{A_j} \otimes \dots \otimes A_n \quad \text{and} \quad B_{\hat{j}} := B_1 \otimes \dots \otimes \widehat{B_j} \otimes \dots \otimes B_n,$$

and

$$B_{\leq j} := B_1 \otimes \dots \otimes B_j \quad \text{and} \quad B_{\geq j} := B_j \otimes \dots \otimes B_n,$$

for $j = 1, \dots, n$. Keep in mind that as mentioned above, either all of these tensor products are minimal, or all are maximal. A channel $E : A \rightarrow B$ can be reduced to a channel $E_{\hat{j}} : A_{\hat{j}} \rightarrow B_{\hat{j}}$ by

$$E_{\hat{j}} := (\text{id}_{B_{\leq j-1}} \otimes \nu_j \otimes \text{id}_{B_{\geq j+1}}) \circ E|_{A_{\hat{j}}}$$

for $j = 1, \dots, n$. Here the restriction $E|_{A_{\widehat{j}}} : A_{\widehat{j}} \rightarrow B$ is defined via

$$E|_{A_{\widehat{j}}}(a_1 \otimes \dots \widehat{a_j} \dots \otimes a_n) = E(a_1 \otimes \dots \otimes a_{j-1} \otimes 1_{A_j} \otimes a_{j+1} \otimes \dots \otimes a_n)$$

for $a_1 \in A_1, \dots, a_n \in A_n$, where 1_{A_j} denotes the unit of A_j .

Set

$$\begin{aligned} \mathcal{O}_{A,B} &:= \{\lambda \in \text{span}_{\mathbb{R}} K(A, B) : \lambda(1_A) = 0\}, \\ \mathcal{N}_j &:= \{D - E : D, E \in K(A, B) \text{ such that } D_{\widehat{j}} = E_{\widehat{j}}\}, \end{aligned}$$

and

$$\mathcal{N} := \bigcup_{j=1}^n \mathcal{N}_j \quad \text{and} \quad \mathcal{C} := \text{conv } \mathcal{N}.$$

Here of course two channels $D, E \in K(A, B)$ are said to be *neighbouring* if $D_{\widehat{j}} = E_{\widehat{j}}$ for some $j \in \{1, \dots, n\}$.

This allows us to state the main result of this section, which is one of the main results of the paper.

Theorem 6.2. *Let $(A_j, B_j, \nu_j : j = 1, \dots, n)$ be an n -composite C^* -system. Then in both the minimal and maximal tensor product setup, $\|\cdot\|_{W_1}$ defined by*

$$\|\lambda\|_{W_1} = \inf\{t \geq 0 : \lambda \in t\mathcal{C}\}$$

for all $\lambda \in \mathcal{O}_{A,B}$, is a norm on $\mathcal{O}_{A,B}$, called the W_1 norm associated to $(A_j, B_j, \nu_j : j = 1, \dots, n)$.

Proof. We are going to obtain this from the algebraic (rather than C^* -algebraic) setup of the previous section. Therefore we have to convert between algebraic and completed tensor products as needed. Write $A_{\odot} = A_1 \odot \dots \odot A_n$ and define

$$K(A, B)|_{\odot} = \{E|_{A_{\odot}} : E \in K(A, B)\}.$$

Since a channel $E \in K(A, B)$ is necessarily continuous, it is uniquely determined by its restriction $E|_{A_{\odot}}$, the latter being the usual restriction of the map E to the subset A_{\odot} of A . Hence $K(A, B)|_{\odot}$ and $K(A, B)$ are in one-to-one correspondence.

Also define

$$A_{\odot, \leq j} = A_1 \odot \dots \odot A_j \quad \text{and} \quad A_{\odot, \geq j} = A_j \odot \dots \odot A_n,$$

and

$$A_{\leq j} = A_1 \otimes \dots \otimes A_j \quad \text{and} \quad A_{\geq j} = A_j \otimes \dots \otimes A_n,$$

for $j = 1, \dots, n$, as well as

$$\begin{aligned} E_{\leq j} &= (\text{id}_{B_{\leq j}} \otimes \nu_{> j}) \circ E|_{A_{\leq j}} : A_{\leq j} \rightarrow B_{\leq j} \\ E_{\geq j} &= (\nu_{< j} \otimes \text{id}_{B_{\geq j}}) \circ E|_{A_{\geq j}} : A_{\geq j} \rightarrow B_{\geq j} \end{aligned}$$

for any $E \in K(A, B)$.

Note that $(A_j, B_j, \nu_j : j = 1, \dots, n)$ gives an algebraic (W_1, n) structure $(A_{\odot}, B, \nu, K(A, B)|_{\odot})$ as in Definition 5.2. This follows from the

automatic continuity of tensor products of states and identity maps in the C*-algebraic framework, along with the fact that condition (13) is satisfied. The latter, in terms of (7) and (8), being

$$(E|_{A_\odot})_{\leq j} \odot (D|_{A_\odot})_{\geq j+1} = E_{\leq j}|_{A_\odot, \leq j} \odot D_{\geq j+1}|_{A_\odot, \geq j+1} \in K(A, B)|_\odot$$

for all $D, E \in K(A, B)$, where $|_{A_\odot}$, $|_{A_\odot, \leq j}$ and $|_{A_\odot, \geq j+1}$ are the usual restrictions to the indicated algebraic tensor products. This fact in turn is true, since $E_{\leq j} \odot D_{\geq j+1}$ uniquely extends to an element $E_{\leq j} \otimes D_{\geq j+1}$ of $K(A, B)$, as $E_{\leq j}$ and $D_{\geq j+1}$ themselves are channels (being the composition of u.c.p. maps), hence indeed

$$E_{\leq j}|_{A_\odot, \leq j} \odot D_{\geq j+1}|_{A_\odot, \geq j+1} = (E_{\leq j} \otimes D_{\geq j+1})|_{A_\odot} \in K(A, B)|_\odot.$$

By Theorem 5.3 and the one-to-one correspondence between $K(A, B)|_\odot$ and $K(A, B)$ mentioned above, we are done. \square

This leads to the following key conclusion.

Corollary 6.3. *Given an n -composite C*-system $(A_j, B_j, \nu_j : j = 1, \dots, n)$, then in both the minimal and the maximal tensor product setup we obtain a metric W_1 on $K(A, B)$ defined by*

$$W_1(D, E) = \|D - E\|_{W_1}$$

for all $D, E \in K(A, B)$, called the Wasserstein distance of order 1 associated to $(A_j, B_j, \nu_j : j = 1, \dots, n)$.

Note that Section 2 emerges as a special case of this section, albeit directly in the Heisenberg picture, by simply setting

$$A_j = M_{q_j} \quad \text{and} \quad B_j = M_{r_j}$$

and letting ν_j be the normalized trace on B_j .

Another special case is $B_1 = \dots = B_n = \mathbb{C}$, with ν_1, \dots, ν_n becoming trivial and irrelevant, but with general unital C*-algebras A_1, \dots, A_n . In this case $K(A, B)$ is the set of all states on $A = A_1 \otimes \dots \otimes A_n$, hence W_1 is now the Wasserstein distance of order 1 between states on A . For $A_1 = \dots = A_n = M_d$ this reduces to the (quantum) Wasserstein distance of order 1 studied in [22], as can be seen from Section 2, keeping in mind that a state on A is exactly a normalized positive linear functional μ , which in this finite dimensional case can be uniquely represented as $\mu(a) = \text{Tr}(\rho a)$ for all $a \in A$ in terms of some density matrix ρ .

We have focussed on the composite system aspect of the framework. From a single system point of view, note that in the finite dimensional setup for states, and setting $n = 1$, for any states ψ and ω on $A_1 = M_d$, we have

$$W_1(\psi, \omega) = \frac{1}{2} \text{Tr} |\rho_\psi - \rho_\omega|,$$

with ρ_ψ and ρ_ω being the density matrices representing ψ and ω respectively, according to [22, Proposition 2]. In the general C*-algebraic case

for states (i.e., $B = \mathbb{C}$) with $n = 1$, we can therefore view W_1 as an abstract version of the trace distance between states, despite the fact that no canonical trace is specified on A_1 in this setup. Refer to Section 8 for further remarks related to this.

This ends our development of the Wasserstein distance of order 1. Next we study its behaviour in relation to subsystems.

7. SUBSYSTEMS AND ADDITIVITY

A core idea behind W_1 is that it is built to reflect the composite structure of systems. It is therefore natural to study its basic properties in relation to subsystems of the composite systems, i.e., smaller tensor products. This is what is done in the current section, first in terms of the W_1 seminorms obtained for pointed (W_1, n) -structures in Section 4, and subsequently for the C^* -algebraic framework of the previous section. We focus on the additivity of W_1 with respect to tensor products (see Theorems 7.9 and 7.11) and the resulting stability of W_1 (see Corollary 7.12). The additivity results of this section generalize those of [22, Section IV.C], though the techniques to achieve them are necessarily different, as [22] makes use of trace norms, which are not available in our context. The reader who wants to see the main results in the C^* -algebraic context, can turn directly to Subsection 7.4, but the proofs and some notation rely on Subsections 7.1, 7.2 and 7.3.

7.1. Pointed W_1 substructures and their W_1 seminorms. Let (A, B, ν, \mathcal{L}) be a pointed (W_1, n) structure as defined in Definition 4.4, again writing $\mathcal{C}_{\mathcal{L}}$ for the semi unit ball in $\mathcal{O}_{\mathcal{L}}$, as given by Proposition 4.3. For simplicity of notation, particularly in the following subsections, however, we continue to write \mathcal{N}_j and \mathcal{N} as in Section 4, rather than, say, $\mathcal{N}_{\mathcal{L},j}$ and $\mathcal{N}_{\mathcal{L}}$. To describe the related subsystems, we need some further notation.

Write

$$[n] := \{1, \dots, n\},$$

and \mathcal{P}_n for the collection of non-empty proper subsets J of $[n]$; by “proper” we mean that $J \neq [n]$. The complement of $J \in \mathcal{P}_n$ will be written as

$$J' := [n] \setminus J = \{j \in [n] : j \notin J\} \in \mathcal{P}_n.$$

We define

$$A_J = \bigodot_{j \in J} A_j \quad \text{and} \quad B_J = \bigotimes_{j \in J} B_j$$

for any $J \in \mathcal{P}_n$, with the same tensor product conventions as in Subsection 4.2. Here the order of the A_j 's in A_J is taken to be the same as in $A = A_1 \odot \dots \odot A_n$, for example $A_{\{2,5\}} = A_2 \odot A_5$ rather than $A_5 \odot A_2$.

Similarly for B_J and correspondingly for $\nu_J := (\nu_j)_{j \in J}$. Elementary tensors in A_J can be denoted as

$$\otimes_{j \in J} a_j$$

for $a_j \in A_j$, and similarly for B_J , for any $J \in \mathcal{P}_n$. In particular A_J is a pointed space with distinguished point

$$u_{A_J} := \otimes_{j \in J} u_{A_j}.$$

Similarly for B_J .

We need to define corresponding reductions of maps. Given $\eta \in L(A, B)$, its *reduction*

$$\eta^J := \left(\bigotimes_{j=1}^n \varphi_j \right) \circ \eta|_{A_J} \in L(A_J, B_J)$$

to $J \in \mathcal{P}_n$ (or over J') is defined as an obvious generalization of the reductions considered in Subsection 4.2, where

$$\varphi_j = \begin{cases} \text{id}_{B_j} & \text{for } j \in J \\ \nu_j & \text{for } j \in J' \end{cases}$$

and with $\eta|_{A_J} : A_J \rightarrow B$ given via

$$\eta|_{A_J}(\otimes_{j \in J} a_j) = \eta(a_1 \otimes \dots \otimes a_n)$$

for $\otimes_{j \in J} a_j$ an elementary tensor in A_J , by setting $a_j = u_{A_j}$ for $j \in J'$. For any subset \mathcal{S} of $L(A, B)$, let

$$(15) \quad \mathcal{S}^J := \{\eta^J : \eta \in \mathcal{S}\}.$$

Note that one can of course reduce any $\theta \in \mathcal{S}^J$ to θ^I for any non-empty proper $I \subset J$, by the obvious adjustment of the method above to this case.

Then we have the next basic fact in terms of Definition 4.4.

Proposition 7.1. *Let (A, B, ν, \mathcal{L}) be a pointed (W_1, n) structure and consider any $J \in \mathcal{P}_n$. The pointed spaces A_j and B_j for $j \in J$, along with $\nu_J = (\nu_j)_{j \in J}$ and \mathcal{L}^J , then form a pointed $(W_1, |J|)$ structure $(A_J, B_J, \nu_J, \mathcal{L}^J)$, called a pointed W_1 substructure of (A, B, ν, \mathcal{L}) .*

Proof. Clearly $\eta^J(u_{A_J}) = u_{B_J}$ for every $\eta \in \mathcal{L}$ by η^J 's definition, hence $\mathcal{L}^J \subset L(A_J, B_J)$, the tensor products of ν_j 's and identity maps restrict those of (A, B, ν, \mathcal{L}) and are therefore still continuous, while the analogue of (13) is easily seen to hold in this context by simply reducing it to $L(A_J, B_J)$. The latter is confirmed by a direct calculation similar to that in Lemma 4.2's proof. \square

By Corollary 4.5, the pointed $(W_1, |J|)$ structure $(A_J, B_J, \nu_J, \mathcal{L}^J)$ provides us with a W_1 seminorm on

$$(16) \quad \mathcal{O}_{\mathcal{L}^J} := \{\lambda \in \text{span}_{\mathbb{R}} \mathcal{L}^J : \lambda(u_{A_J}) = 0\}$$

for every $J \in \mathcal{P}_n$, still denoted as

$$\|\cdot\|_{W_1},$$

as well as the resulting pseudometric W_1 on \mathcal{L}^J . This is simply a case of Section 4, but now of course using

$$(17) \quad \mathcal{N}_{J,j} := \{\eta - \theta : \eta, \theta \in \mathcal{L}^J \text{ such that } \eta_{\hat{j}} = \theta_{\hat{j}}\} \text{ for } j \in J$$

instead of $\mathcal{N}_1, \dots, \mathcal{N}_n$, where the latter led to the W_1 seminorm on $\mathcal{O}_{\mathcal{L}}$. Here

$$\eta_{\hat{j}} := \eta^{J \setminus \{j\}}$$

for $\eta \in \mathcal{L}^J$, which can equivalently be defined by (6), but using A_i, B_i and ν_i only for $i \in J$ when setting up Section 4, rather than for the entire range $i = 1, \dots, n$. The semi unit ball leading to this W_1 seminorm is

$$(18) \quad \mathcal{C}_{\mathcal{L}^J} := \text{conv } \mathcal{N}_J,$$

where

$$(19) \quad \mathcal{N}_J := \bigcup_{j \in J} \mathcal{N}_{J,j}.$$

7.2. Reducible pointed W_1 structures. A natural question is whether the semi unit ball $\mathcal{C}_{\mathcal{L}^J}$ above can be obtained as the reduction $\mathcal{C}_{\mathcal{L}}^J$ of the original semi unit ball $\mathcal{C}_{\mathcal{L}}$. Similarly for $\mathcal{O}_{\mathcal{L}^J}$ and the sets $\mathcal{N}_{J,j}$. These questions will in fact become relevant in the next subsection, when we reach the main goal of this section, namely to prove additivity properties.

In order to answer these questions positively, assumptions beyond those in Definition 4.4 need to be made. We note that these assumptions will automatically be satisfied in the C*-algebraic framework.

As in the previous subsection we consider a pointed (W_1, n) structure (A, B, ν, \mathcal{L}) and any $J \in \mathcal{P}_n$.

To avoid any mismatches and ambiguities, we always need to preserve the ordering of the A_j 's in any tensor product of them. Similarly for the B_j 's. Therefore the notation

$$A_I \vec{\odot} A_J := A_{I \cup J} \quad \text{and} \quad B_I \vec{\otimes} B_J := B_{I \cup J}$$

will be used for any $I, J \in \mathcal{P}_n$ with no points in common, i.e., $I \cap J = \emptyset$. But then the tensor product of maps $\eta \in L(A_I, B_I)$ and $\theta \in L(A_J, B_J)$ for $I, J \in \mathcal{P}_n$ with $I \cap J = \emptyset$ need to be defined correspondingly. This is indeed possible. Note that transpositions of adjacent A_j 's in any tensor product of A_j 's are linear bijections, compositions of which in particular give us a natural unique well defined *ordering map*

$$\alpha_{IJ} : A_I \odot A_J \rightarrow A_I \vec{\odot} A_J,$$

such that

$$\alpha_{IJ}((\otimes_{i \in I} a_i) \otimes (\otimes_{j \in J} a_j)) = \otimes_{j \in I \cup J} a_j$$

for arbitrary $a_j \in A_j$. As a simple example to clarify the meaning of this, suppose $I = \{1, 3, 5\}$ and $J = \{2, 4\}$, then for $a_j \in A_j$, we have $\alpha_{IJ}(a_1 \otimes a_3 \otimes a_5 \otimes a_2 \otimes a_4) = a_1 \otimes a_2 \otimes a_3 \otimes a_4 \otimes a_5$. Note that α_{IJ} can be viewed as a *pointed space isomorphism*, i.e., a bijection $\alpha_{IJ} \in L_u(A_I \odot A_J, A_I \vec{\odot} A_J)$.

Similarly we have the ordering map

$$\beta_{IJ}^0 : B_I \odot B_J \rightarrow B_I \vec{\odot} B_J.$$

Clearly $\beta_{IJ}^0 \in L_u(B_I \odot B_J, B_I \vec{\odot} B_J)$.

However, if the tensor products B_J for $J \subset [n]$ are indeed completed in some norm, we need to *assume* that β_{IJ}^0 is continuous in this norm, and therefore extends uniquely to a continuous bijection

$$\beta_{IJ} \in L_u(B_I \otimes B_J, B_I \vec{\otimes} B_J),$$

which is the uniquely defined pointed space isomorphism serving as the ordering map on $B_I \otimes B_J$.

Given this, we can define

$$\eta \vec{\odot} \theta : A_I \vec{\odot} A_J \rightarrow B_I \vec{\otimes} B_J$$

as

$$\eta \vec{\odot} \theta := \beta_{IJ} \circ (\eta \odot \theta) \circ \alpha_{IJ}^{-1}.$$

It has the following expected property.

Proposition 7.2. *In terms of the notation and assumptions so far in this subsection,*

$$\eta \vec{\odot} \theta = \theta \vec{\odot} \eta,$$

for all $\eta \in L(A_I, B_I)$ and $\theta \in L(A_J, B_J)$, where $I, J \in \mathcal{P}_n$ with $I \cap J = \emptyset$.

Proof. Note that by the definition of the ordering maps, we have $\beta_{IJ}(c \otimes d) = \beta_{JI}(d \otimes c)$ for elements in the algebraic tensor products, $c \in \odot_{i \in I} B_i$ and $d \in \odot_{j \in J} B_j$. Now, for arbitrary $a_j \in A_j$,

$$(\eta \odot \theta) \circ \alpha_{IJ}^{-1}(\otimes_{j \in I \cup J} a_j) = \eta(\otimes_{i \in I} a_i) \otimes \theta(\otimes_{j \in J} a_j)$$

$$(\theta \odot \eta) \circ \alpha_{JI}^{-1}(\otimes_{j \in I \cup J} a_j) = \theta(\otimes_{j \in J} a_j) \otimes \eta(\otimes_{i \in I} a_i)$$

Approximate $\eta(\otimes_{i \in I} a_i)$ and $\theta(\otimes_{j \in J} a_j)$ by sequences (c_l) and (d_l) in the algebraic tensor products $\odot_{i \in I} B_i$ and $\odot_{j \in J} B_j$ respectively. Since $\beta_{IJ}(c_l \otimes d_l) = \beta_{JI}(d_l \otimes c_l)$ and β_{IJ} and β_{JI} are assumed to be continuous, it follows that

$$\beta_{IJ} \circ (\eta \odot \theta) \circ \alpha_{IJ}^{-1}(\otimes_{j \in I \cup J} a_j) = \beta_{JI} \circ (\theta \odot \eta) \circ \alpha_{JI}^{-1}(\otimes_{j \in I \cup J} a_j),$$

as required. \square

We also need to strengthen (13) in Lemma 4.2 to the following: *Assume that*

$$(20) \quad \eta^J \vec{\odot} \theta^{J'} \in \mathcal{L}$$

for all $\eta, \theta \in \mathcal{L}$ and $J \in \mathcal{P}_n$.

Definition 7.3. The pointed (W_1, n) structure (A, B, ν, \mathcal{L}) is called *reducible* if both the above mentioned assumptions are indeed satisfied, namely (20) and the existence of the continuous ordering maps β_{IJ} in the case of completed B .

Note that for a non-empty $I \subset J'$, it follows from (20) and the definition of reduction in the previous subsection, that

$$(21) \quad \eta^J \vec{\circ} \theta^I = (\eta^J \vec{\circ} \theta^{J'})^{J \cup I} \in \mathcal{L}^{J \cup I}.$$

In particular, this gives the following simple result.

Proposition 7.4. *If the pointed (W_1, n) structure (A, B, ν, \mathcal{L}) is reducible, then so is its pointed W_1 substructures.*

The reason for the terminology “reducible” in Definition 7.3, is that the semi unit ball of $(A_J, B_J, \nu_J, \mathcal{L}^J)$ is then obtained from that of (A, B, ν, \mathcal{L}) by reduction. This and related facts are shown below.

In terms of the setup and notation of this section, we have the following.

Lemma 7.5. *Assume that (A, B, ν, \mathcal{L}) is a reducible pointed (W_1, n) structure and consider any $J \in \mathcal{P}_n$. For all $\lambda \in \mathcal{N}_{J,j}$ with $j \in J$, and $\zeta \in \mathcal{L}^{J'}$, it follows that $\lambda \vec{\circ} \zeta \in \mathcal{N}_j$.*

Proof. Note that $(\lambda \vec{\circ} \zeta)_{\hat{j}} = \lambda_{\hat{j}} \vec{\circ} \zeta = 0$. In terms of $\lambda = \eta - \theta$ with $\eta, \theta \in \mathcal{L}^J$, this means that $\lambda \vec{\circ} \zeta = \eta \vec{\circ} \zeta - \theta \vec{\circ} \zeta$, where $\eta \vec{\circ} \zeta, \theta \vec{\circ} \zeta \in \mathcal{L}$ by (20) and $(\eta \vec{\circ} \zeta)_{\hat{j}} = (\theta \vec{\circ} \zeta)_{\hat{j}}$, as needed in (11). \square

Proposition 7.6. *Assume that (A, B, ν, \mathcal{L}) is a reducible pointed (W_1, n) structure and consider any $J \in \mathcal{P}_n$. Then $\mathcal{N}_j^J = \{0\}$ when $j \in J'$, while*

$$\mathcal{N}_j^J = \mathcal{N}_{J,j}$$

for $j \in J$.

Proof. By (15), $\mathcal{N}_j^J := \{\eta^J : \eta \in \mathcal{N}_j\}$. For $j \in J'$ the reduction over j is included in the reduction over J' , hence $\lambda_{\hat{j}} = 0$ for $\lambda \in L(A, B)$ implies that $\lambda^J = 0$, directly from the definitions of $\lambda_{\hat{j}}$ and λ^J . Thus $\mathcal{N}_j^J = \{0\}$ when $j \in J'$. Now assume that $j \in J$. Consider any $\lambda^J \in \mathcal{N}_j^J$, i.e., we take $\lambda = \eta - \theta$ with $\eta, \theta \in \mathcal{L}$ and $\eta_{\hat{j}} = \theta_{\hat{j}}$. Since $j \in J$, both sides of $(\eta^J)_{\hat{j}} = (\eta_{\hat{j}})^J$ are well defined, and indeed equal by the definitions of these reductions. Similarly for θ , which means that $(\eta^J)_{\hat{j}} = (\theta^J)_{\hat{j}}$, hence $\lambda^J \in \mathcal{N}_{J,j}$ by (17). This shows that $\mathcal{N}_j^J \subset \mathcal{N}_{J,j}$, even if (A, B, ν, \mathcal{L}) is not assumed reducible. Conversely, consider any $\lambda \in \mathcal{N}_{J,j}$. For any $\zeta \in \mathcal{L}^{J'}$ it then follows from Lemma 7.5 that $\lambda \vec{\circ} \zeta \in \mathcal{N}_j$. Consequently, $\lambda = (\lambda \vec{\circ} \zeta)^J \in \mathcal{N}_j^J$, proving that $\mathcal{N}_{J,j} \subset \mathcal{N}_j^J$. \square

In particular this tells us that the reductions \mathcal{N}_j^J of \mathcal{N}_j for $j \in J$, play the same role for $(A_J, B_J, \nu_J, \mathcal{L}^J)$ as $\mathcal{N}_1, \dots, \mathcal{N}_n$ play for the reducible pointed W_1 structure (A, B, ν, \mathcal{L}) .

Corollary 7.7. *In Proposition 7.6 we have*

$$\mathcal{O}_{\mathcal{L}^J} = \mathcal{O}_{\mathcal{L}}^J, \quad \mathcal{N}_J = \mathcal{N}^J \quad \text{and} \quad \mathcal{C}_{\mathcal{L}^J} = \mathcal{C}_{\mathcal{L}}^J$$

for (16), (19) and (18).

Proof. From Proposition 7.6, (12) and (19) one has $\mathcal{N}_J = \mathcal{N}^J$, hence

$$\mathcal{C}_{\mathcal{L}^J} = \text{conv } \mathcal{N}^J = (\text{conv } \mathcal{N})^J = \mathcal{C}_{\mathcal{L}}^J$$

and

$$\mathcal{O}_{\mathcal{L}^J} = \text{span}_{\mathbb{R}} \mathcal{N}^J = (\text{span}_{\mathbb{R}} \mathcal{N})^J = \mathcal{O}_{\mathcal{L}}^J,$$

because of Lemma 4.2. \square

As one may expect, $\mathcal{O}_{\mathcal{L}^J} = \mathcal{O}_{\mathcal{L}}^J$ can alternatively be proved along the lines of the proof of Proposition 7.6.

We can also use Lemma 7.5 along with Corollary 7.7 to obtain the following property of $\|\cdot\|_{W_1}$.

Proposition 7.8. *Assume that (A, B, ν, \mathcal{L}) is a reducible pointed (W_1, n) structure and consider any $I, J \in \mathcal{P}_n$ with $I \cap J = \emptyset$. Then*

$$\|(\eta - \theta)\vec{\odot}\zeta\|_{W_1} = \|\eta - \theta\|_{W_1}$$

for all $\eta, \theta \in \mathcal{L}^I$ and $\zeta \in \mathcal{L}^J$.

Proof. For any $\eta, \theta \in \mathcal{L}^I$ and $\zeta \in \mathcal{L}^J$, set $\lambda := (\eta - \theta)\vec{\odot}\zeta \in \mathcal{L}^{I \cup J}$, according to (21), then $\lambda^I = \eta - \theta$.

For any $\gamma \in \mathcal{C}_{\mathcal{L}^{I \cup J}}$ and $t \geq 0$ such that $\lambda = t\gamma$, one has that $\lambda^I = t\gamma^I$. By Corollary 7.7, but applied to the reducible (Proposition 7.4) pointed W_1 structure $(A_{I \cup J}, B_{I \cup J}, \nu_{I \cup J}, \mathcal{L}^{I \cup J})$, we know that $\gamma^I \in \mathcal{C}_{\mathcal{L}^I}$, from which it follows that $\|\lambda^I\|_{W_1} \leq t$, hence $\|\lambda^I\|_{W_1} \leq \|\lambda\|_{W_1}$ by the definition of $\|\cdot\|_{W_1}$ in Corollary 4.5 (via Proposition 3.2).

Conversely, consider any $\gamma \in \mathcal{C}_{\mathcal{L}^I}$ and $t \geq 0$ such that $\lambda^I = t\gamma$. By (18) we have

$$\gamma = \sum_{i=1}^l p_i \gamma_i$$

for some $p_1, \dots, p_l > 0$ with $p_1 + \dots + p_l = 1$, and $\gamma_i \in \mathcal{N}_{I, j_i}$ for some $j_i \in I$. Because of Lemma 7.5 applied to $(A_{I \cup J}, B_{I \cup J}, \nu_{I \cup J}, \mathcal{L}^{I \cup J})$, it follows that

$$\gamma\vec{\odot}\zeta = \sum_{i=1}^l p_i \gamma_i \vec{\odot}\zeta \in \mathcal{C}_{\mathcal{L}^{I \cup J}}.$$

Since $\lambda = t\gamma\vec{\odot}\zeta$, we conclude that $\|\lambda\|_{W_1} \leq t$, thus $\|\lambda\|_{W_1} \leq \|\lambda^I\|_{W_1}$. \square

These results will be applied in the next subsection to prove the additivity of W_1 with respect to tensor products.

7.3. Additivity. We now arrive at this section's main results in the abstract pointed space setup, which will be applied to the C^* -algebras in the next subsection. Consider a reducible pointed (W_1, n) structure (A, B, ν, \mathcal{L}) and any m -partition P of $[n]$, by which we mean a function $P : [m] \rightarrow \mathcal{P}_n$ such that $P(1) \cup \dots \cup P(m) = [n]$ and $P(k) \cap P(l) = \emptyset$ for $k \neq l$. Our goal is to determine how $\|\cdot\|_{W_1}$ for (A, B, ν, \mathcal{L}) relates to $\|\cdot\|_{W_1}$ for the $(A_{P(k)}, B_{P(k)}, \nu_{P(k)}, \mathcal{L}^{P(k)})$'s via reduction. Similarly for W_1 , but specifically for product maps. Using the results of the previous subsection, these relationships can be stated as a form of “reductive superadditivity” of $\|\cdot\|_{W_1}$ and an additivity property of W_1 .

Theorem 7.9. *Let (A, B, ν, \mathcal{L}) be a reducible pointed (W_1, n) structure and P any m -partition of $[n]$. Then*

$$(22) \quad \|\lambda\|_{W_1} \geq \sum_{k=1}^m \|\lambda^{P(k)}\|_{W_1}$$

for all $\lambda \in \mathcal{O}_{\mathcal{L}}$, while

$$(23) \quad W_1(\eta_1 \vec{\odot} \dots \vec{\odot} \eta_m, \theta_1 \vec{\odot} \dots \vec{\odot} \theta_m) = \sum_{k=1}^m W_1(\eta_k, \theta_k)$$

for all $\eta_k, \theta_k \in \mathcal{L}^{P(k)}$, for $k = 1, \dots, m$.

Proof. Let $\lambda \in \mathcal{O}_{\mathcal{L}}$. Consider any $\gamma \in \mathcal{C}_{\mathcal{L}}$ and $t \geq 0$ such that $\lambda = t\gamma$. By $\mathcal{C}_{\mathcal{L}} = \text{conv } \mathcal{N}$,

$$\gamma = \sum_{i=1}^l p_i \alpha_i$$

for some $p_1, \dots, p_l > 0$ with $p_1 + \dots + p_l = 1$, and $\alpha_i \in \mathcal{N} = \cup_{j \in [n]} \mathcal{N}_j$. Set

$$q_k := \sum_{i \in R_k} p_i \quad \text{and} \quad \gamma_k := \sum_{i \in R_k} \frac{p_i}{q_k} \alpha_i \in \mathcal{C}_{\mathcal{L}},$$

where $R_k := \{i \in [l] : \alpha_i \in \cup_{j \in P(k)} \mathcal{N}_j\} \setminus (R_1 \cup \dots \cup R_{k-1})$ for $k = 1, \dots, m$, with $R_1 \cup \dots \cup R_{k-1} = \emptyset$ for $k = 1$. Then by Proposition 7.6,

$$\gamma = \sum_{k=1}^m q_k \gamma_k \quad \text{and} \quad \gamma^{P(k)} = q_k \gamma_k^{P(k)}$$

with $\gamma_k^{P(k)} \in \mathcal{C}_{\mathcal{L}^{P(k)}} = \mathcal{C}_{\mathcal{L}^{P(k)}}$ by Corollary 7.7. This tells us that $\lambda^{P(k)} = t\gamma^{P(k)} = tq_k \gamma_k^{P(k)} \in tq_k \mathcal{C}_{\mathcal{L}^{P(k)}}$, hence $\|\lambda^{P(k)}\|_{W_1} \leq tq_k$, thus

$$\sum_{k=1}^m \|\lambda^{P(k)}\|_{W_1} \leq t,$$

implying (22), by $\|\cdot\|_{W_1}$'s definition in Corollary 4.5.

In particular, for $\eta, \theta \in \mathcal{L}$ we have

$$\|\eta - \theta\|_{W_1} \geq \sum_{k=1}^m \|\eta^{P(k)} - \theta^{P(k)}\|_{W_1}.$$

On the other hand, given $\eta_k, \theta_k \in \mathcal{L}^{P(k)}$, by Proposition 7.8 we have

$$\begin{aligned} & \|\eta_1 \vec{\otimes} \dots \vec{\otimes} \eta_m - \theta_1 \vec{\otimes} \dots \vec{\otimes} \theta_m\|_{W_1} \\ & \leq \|(\eta_1 - \theta_1) \vec{\otimes} \eta_2 \vec{\otimes} \dots \vec{\otimes} \eta_m\|_{W_1} + \|\theta_1 \vec{\otimes} (\eta_2 \vec{\otimes} \dots \vec{\otimes} \eta_m - \theta_2 \vec{\otimes} \dots \vec{\otimes} \theta_m)\|_{W_1} \\ & = \|\eta_1 - \theta_1\|_{W_1} + \|\eta_2 \vec{\otimes} \dots \vec{\otimes} \eta_m - \theta_2 \vec{\otimes} \dots \vec{\otimes} \theta_m\|_{W_1} \\ & \vdots \\ & \leq \sum_{k=1}^m \|\eta_k - \theta_k\|_{W_1}. \end{aligned}$$

These two inequalities prove (23). \square

With this result in hand, we can turn to the final aim of this section, namely additivity in the C*-algebraic framework.

7.4. C*-algebras. The question being studied at the moment, is how W_1 between product channels from one compound system to another, relate to W_1 between the channels composing the product channels. This was answered in a more abstract form in Theorem 7.9 of the previous subsection, namely they are related in a simple additive way. Now we essentially just translate this additivity of W_1 to the C*-algebraic framework of Section 6.

Given an n -composite C*-system $(A_j, B_j, \nu_j : j = 1, \dots, n)$ as in Definition 6.1, let $[n] := \{1, \dots, n\}$, let \mathcal{P}_n be the collection of non-empty proper subsets of $[n]$, and define an m -partition P of $[n]$, as a function $P : [m] \rightarrow \mathcal{P}_n$ such that $P(1) \cup \dots \cup P(m) = [n]$ and $P(k) \cap P(l) = \emptyset$ for $k \neq l$. Furthermore, we set

$$A_J := \bigotimes_{j \in J} A_j \quad \text{and} \quad B_J := \bigotimes_{j \in J} B_j$$

for any $J \in \mathcal{P}_n$, both being minimal tensor products or both maximal tensor products. As in Section 6, either all tensor products here are minimal, or all are maximal. For any $I, J \in \mathcal{P}_n$ with no points in common, i.e., $I \cap J = \emptyset$, consider the ordered tensor products

$$A_I \vec{\otimes} A_J := A_{I \cup J} \quad \text{and} \quad B_I \vec{\otimes} B_J := B_{I \cup J}$$

and the corresponding ordered tensor product

$$\eta \vec{\otimes} \theta : A_I \vec{\otimes} A_J \rightarrow B_I \vec{\otimes} B_J$$

of bounded linear maps $\eta : A_I \rightarrow B_I$ and $\theta : A_J \rightarrow B_J$, defined as in Subsection 7.2, but via the minimal or maximal tensor product $\eta \otimes \theta$, instead of the algebraic tensor product, and where the ordering map

α_{IJ} is of course extended to the completed tensor product, like β_{IJ} is. In the C^* -algebraic case, this extension of the ordering maps is automatically possible (see the proof of Proposition 7.10 below). Note that the reduction η^J of $\eta \in L(A, B)$ is defined analogously to Subsection 7.1, the only difference being that A and A_J are now completed tensor products.

The next result is the key to convert Theorem 7.9 to the C^* -algebraic setup.

Proposition 7.10. *Consider the setup of this subsection. For any $J \in \mathcal{P}_n$, it follows that*

$$K(A, B)^J = K(A_J, B_J).$$

In addition, for any $D \in K(A_I, B_I)$ and $E \in K(A_J, B_J)$ with $I, J \in \mathcal{P}_n$ such that $I \cap J = \emptyset$, one has

$$D \vec{\otimes} E \in K(A_I \vec{\otimes} A_J, B_I \vec{\otimes} B_J).$$

Proof. Starting with the latter statement, note that in this C^* -algebraic setup the ordering maps from Subsection 7.2 are $*$ -isomorphisms $\alpha_{IJ} : A_I \otimes A_J \rightarrow A_I \vec{\otimes} A_J$ and $\beta_{IJ} : B_I \otimes B_J \rightarrow B_I \vec{\otimes} B_J$ (and therefore extended to the completions and u.c.p.) from the outset, since they are compositions of the transposition maps mentioned in Subsection 7.2, which are indeed $*$ -isomorphisms (see for example [7, II.9.2.6]). Hence $D \vec{\otimes} E := \beta_{IJ} \circ (D \otimes E) \circ \alpha_{IJ}^{-1}$ is a composition of u.c.p. maps, since $D \vec{\otimes} E$ is a channel. Consequently $D \vec{\otimes} E \in K(A_I \vec{\otimes} A_J, B_I \vec{\otimes} B_J)$, as required.

In particular, for $I = J'$, it follows that $E = (D \vec{\otimes} E)^J \in K(A, B)^J$. Hence $K(A_J, B_J) \subset K(A, B)^J$. Conversely, since the reduction of a channel is again a channel, we have $K(A, B)^J \subset K(A_J, B_J)$. \square

Now we can answer the above mentioned question as follows, along with the “reductive superadditivity” of $\|\cdot\|_{W_1}$.

Theorem 7.11. *Let $(A_j, B_j, \nu_j : j = 1, \dots, n)$ be an n -composite C^* -system and P any m -partition of $[n]$. Then in both the minimal and the maximal tensor product setup,*

$$W_1(D_1 \vec{\otimes} \dots \vec{\otimes} D_m, E_1 \vec{\otimes} \dots \vec{\otimes} E_m) = \sum_{k=1}^m W_1(D_k, E_k)$$

for all $D_k, E_k \in K(A_{P(k)}, B_{P(k)})$, for $k = 1, \dots, m$. In addition,

$$\|\lambda\|_{W_1} \geq \sum_{k=1}^m \|\lambda^{P(k)}\|_{W_1}$$

for all $\lambda \in \mathcal{O}_{A, B}$ as defined in Section 6.

Proof. If the algebraic (W_1, n) structure $(A_\odot, B, \nu, K(A, B)|_\odot)$, obtained as in the proof of Theorem 6.2 from $(A_j, B_j, \nu_j : j = 1, \dots, n)$, is a reducible pointed (W_1, n) structure, and

$$(24) \quad (K(A, B)|_\odot)^J = K(A_J, B_J)|_\odot,$$

for $J \in \mathcal{P}_n$, then the theorem follows directly from Theorem 7.9. This is because of the continuity of all the maps involved and the resulting one-to-one correspondence between $K(A, B)|_\odot$ and $K(A, B)$ explained in the proof of Theorem 6.2, and similarly for $K(A_J, B_J)|_\odot$ and $K(A_J, B_J)$, which allows us to translate directly between the algebraic and completed tensor products. Condition (24) is needed to ensure that for $\mathcal{L} = K(A, B)|_\odot$ in Theorem 7.9, we have $\mathcal{L}^{P(k)} = K(A_{P(k)}, B_{P(k)})|_\odot$. Note that (24) indeed holds because of Proposition 7.10 and $(K(A, B)|_\odot)^J = K(A, B)^J|_\odot$, the latter being true since the only difference between the reductions on the two sides is that they respectively involve restrictions to the algebraic and a completed tensor product of A_j 's. We are simply left to verify reducibility.

As mentioned in the proof of Proposition 7.10, the β_{IJ} 's are $*$ -isomorphisms and therefore continuous, verifying the one condition for reducibility in Definition 7.3. The other condition is guaranteed by the special case of Proposition 7.10 with $I \cup J = [n]$. \square

Recall that the special case in finite dimensions of the first part of this theorem was already mentioned in Section 2, in a special form where the ordering of the tensor products was unnecessary.

As an immediate consequence of this theorem, we obtain the following stability result for W_1 .

Corollary 7.12. *In Theorem 7.11, let P a 2-partition of $[n]$. Then*

$$W_1(D_1 \vec{\otimes} F_2, E_1 \vec{\otimes} F_2) = W_1(D_1, E_1)$$

for all $D_1, E_1 \in K(A_{P(1)}, B_{P(1)})$ and any $F_2 \in K(A_{P(2)}, B_{P(2)})$. Similarly,

$$W_1(F_1 \vec{\otimes} D_2, F_1 \vec{\otimes} E_2) = W_1(D_2, E_2)$$

for all $D_2, E_2 \in K(A_{P(2)}, B_{P(2)})$ and any $F_1 \in K(A_{P(1)}, B_{P(1)})$.

Proof. This is simply because $W_1(F_2, F_2) = 0$ in the first case. Similarly for the second. \square

Stability for other distances between channels have been presented and discussed in for example [3] and [29]. The latter in particular emphasizes the utility of stability. These references only treated the case where the F_k 's in the corollary above were identity maps.

Of course, stability for pointed W_1 structures similarly follow from Theorem 7.9.

8. OUTLOOK

In this paper our focus has been the mathematical development of a Wasserstein distance of order 1 between channels from one composite system to another, denoted by W_1 . We have not yet investigated any relations or comparisons of W_1 with the diamond norm (mentioned in the introduction) or other distances between channels. See [29] for a critical assessment of various such distances. A logical first step is to do this in finite dimensions on matrix algebras M_d , where one has a simple canonical trace which plays an important role in the diamond (aka, completely bounded trace) norm. Indeed, [22] extensively explored the relation between W_1 for states and the trace norm, including the characterization of W_1 for the case of $n = 1$, i.e., for single systems rather than composed systems. This can analogously be explored for the case of channels.

As mentioned at the end of the introduction, one could in principle attempt to translate the DMTL approach for the case of states to channels via the Choi-Jamiołkowski duality, in situations where the latter holds. But note that this duality gives a correspondence between channels and a specific subset of states, namely those which reduce to the maximally mixed state of the system to which the channels map in the Heisenberg picture. (For simplicity, in this discussion we are only considering the finite dimensional case.) Although we have not studied the precise relationship between our approach and this possible Choi-Jamiołkowski translation approach, it appears intuitively clear that they are not equivalent. To appreciate this, note that since DMTL's setup refers to all states (not just those corresponding to channels), it follows that the unit ball we obtain in our approach for channels will typically be strictly contained in the one obtained by the DMTL approach. In this regard, recall from Subsection 2.2 that the reduction processes for the channels and their corresponding states are equivalent. In the case of channels our W_1 can consequently be expected to be larger than what would be obtained by DMTL via Choi-Jamiołkowski, though the exact relationship between these two distances remains to be explored. Given that our W_1 appears to be the larger, it is natural to ask whether it provides a sharper way to distinguish channels than using DMTL via Choi-Jamiołkowski.

We have also not yet explored applications of this distance. Applications of Wasserstein distance of order 1 between states in the finite dimensional case developed in [22], have already been treated and proposed in [22] itself, as well as in a number of papers [21, 24, 33, 37, 40, 46] in various contexts. We expect that the approach of this paper to the case of channels should be similarly applicable. In addition, our abstract approach has the potential to allow for applications in

other contexts than quantum channels and for further mathematical development.

At the end of Subsection 7.4 we pointed out that W_1 satisfies stability. However, there are other properties that one may also want a distance between channels to satisfy, depending on the application. See for example [29]. Since W_1 is a metric, a number of basic properties are already satisfied. One property that we have not discussed in this paper is chaining or bounds on the overall difference between composed channels; see [29] and [3, Subsection 5.4]. Whether or not this or similar and other properties hold for W_1 , certainly warrants further investigation.

Acknowledgement. We thank the referee for a careful reading and remarks to improve aspects of the exposition. This work is based on research supported in part by the National Research Foundation of South Africa (Grant Number 119899).

REFERENCES

- [1] S. Aaronson, G. N. Rothblum, Gentle measurement of quantum states and differential privacy, in *Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing*, STOC 2019, pages 322–333, Association for Computing Machinery, New York, NY, USA, 2019.
- [2] J. Agredo, A Wasserstein-type distance to measure deviation from equilibrium of quantum Markov semigroups, *Open Syst. Inf. Dyn.* **20** (2013), 1350009.
- [3] D. Aharonov, A. Kitaev, N. Nisan, Quantum circuits with mixed states, in *Proceedings of the 30th Annual ACM Symposium on Theory of Computing*, STOC '98, pages 20–30, Association for Computing Machinery, New York, NY, USA, 1998.
- [4] J.-D. Benamou, Y. Brenier, A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem, *Numer. Math.* **84** (2000), 375–393.
- [5] F. F. Bonsall, J. Duncan, *Complete normed algebras*, Ergebnisse der Mathematik und ihrer Grenzgebiete Band 80, Springer -Verlag, Berlin, Heidelberg, New York, 1973.
- [6] P. Biane, D. Voiculescu, A free probability analogue of the Wasserstein metric on the trace-state space, *Geom. Funct. Anal.* **11** (2001), 1125–1138.
- [7] B. Blackadar, *Operator Algebras: Theory of C^* -Algebras and von Neumann Algebras*, Encyclopaedia of Mathematical Sciences Volume 122, Springer-Verlag, Berlin, Heidelberg, 2006.
- [8] 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, *Comm. Math. Phys.* **331** (2014), 887–926.
- [9] E. A. Carlen, J. Maas, Gradient flow and entropy inequalities for quantum Markov semigroups with detailed balance, *J. Funct. Anal.* **273** (2017), 1810–1869.
- [10] E. A. Carlen, J. Maas, Non-commutative calculus, optimal transport and functional inequalities in dissipative quantum systems, *J. Stat. Phys.* **178** (2020), 319–378.
- [11] F. Caruso, V. Giovannetti, C. Lupo, S. Mancini, Quantum channels and memory effects, *Rev. Mod. Phys.* **86** (2014), 1203–1259.

- [12] Y. Chen, W. Gangbo, T. T. Georgiou, A. Tannenbaum, On the matrix Monge-Kantorovich problem, *European J. Appl. Math.* **31** (2020), 574–600.
- [13] Y. Chen, T. T. Georgiou, L. Ning, A. Tannenbaum, Matricial Wasserstein-1 distance, *IEEE Control Syst. Lett.* **1** (2017), 14–19.
- [14] Y. Chen, T.T. Georgiou, A. Tannenbaum, Matrix optimal mass transport: a quantum mechanical approach, *IEEE Trans. Automat. Control* **63** (2018), 2612–2619.
- [15] M.-D. Choi, Completely positive linear maps on complex matrices, *Linear Algebra Appl.* **10** (1975), 285–290.
- [16] S. Cole, M. Eckstein, S. Friedland, K. Życzkowski, Quantum optimal transport, arXiv:2105.06922.
- [17] A. Connes, Compact metric spaces, Fredholm modules, and hyperfiniteness, *Ergodic Theory Dynam. Systems* **9** (1989), 207–220.
- [18] A. Connes, *Noncommutative geometry*, Academic Press, Inc., San Diego, CA, 1994.
- [19] A. Connes, Gravity coupled with matter and the foundation of noncommutative geometry, *Comm. Math. Phys.* **182** (1996), 155–176.
- [20] F. D’Andrea, P. Martinetti, A view on optimal transport from noncommutative geometry, *SIGMA Symmetry Integrability Geom. Methods Appl.* **6** (2010), 057.
- [21] G. De Palma, M. Marvian, C. Rouzé, D. Stilck França, Limitations of variational quantum algorithms: a quantum optimal transport approach, *PRX Quantum* **4** (2023) 010309.
- [22] G. De Palma, M. Marvian, D. Trevisan, S. Lloyd, The quantum Wasserstein distance of order 1, *IEEE Trans. Inf. Theory* **67** (2021) 6627–6643.
- [23] G. De Palma, D. Trevisan, Quantum optimal transport with quantum channels, *Ann. Henri Poincaré* **22** (2021), 3199–3234.
- [24] G. De Palma, C. Rouzé, Quantum Concentration Inequalities, *Ann. Henri Poincaré* **23** (2022), 3391–3429.
- [25] J. de Pillis, Linear transformations which preserve hermitian and positive semi-definite operators, *Pacific J. Math.* **23** (1967), 129–137.
- [26] R. Duvenhage, Quadratic Wasserstein metrics for von Neumann algebras via transport plans, *J. Operator Theory* **88** (2022), 289–308..
- [27] R. Duvenhage, M. Snyman, Balance between quantum Markov semigroups, *Ann. Henri Poincaré* **19** (2018), 1747–1786.
- [28] L. Gao, M. Junge, N. LaRacuenta, Fisher information and logarithmic Sobolev inequality for matrix-valued functions, *Ann. Henri Poincaré* **21** (2020), 3409–3478.
- [29] A. Gilchrist, N. K. Langford, M. A. Nielsen, Distance measures to compare real and ideal quantum processes, *Phys. Rev. A* **71** (2005), 062310.
- [30] F. Golse, C. Mouhot, T. Paul, On the mean field and classical limits of quantum mechanics, *Comm. Math. Phys.* **343** (2016), 165–205.
- [31] F. Golse, T. Paul, The Schrödinger equation in the mean-field and semiclassical regime, *Arch. Ration. Mech. Anal.* **223** (2017), 57–94.
- [32] F. Golse, T. Paul, Wave packets and the quadratic Monge-Kantorovich distance in quantum mechanics, *C. R. Math. Acad. Sci. Paris* **356** (2018), 177–197.
- [33] C. Hirche, C. Rouzé, D. Stilck França, Quantum Differential Privacy: An Information Theory Perspective, arXiv:2202.10717.
- [34] A. S. Holevo, V. Giovannetti, Quantum channels and their entropic characteristics, *Rep. Progr. Phys.* **75** (2012), 046001.
- [35] D. F. Hornshaw, L^2 -Wasserstein distances of tracial W^* -algebras and their disintegration problem, arXiv:1806.01073.

- [36] A. Jamiolkowski, Linear transformations which preserve trace and positive semidefiniteness of operators, *Rep. Mathematical Phys.* **3** (1972), 275–278.
- [37] B. T. Kiani, G. De Palma, M. Marvian, Z.-W. Liu, S. Lloyd, Learning quantum data with the quantum earth mover’s distance, *Quantum Sci. Technol.* **7** (2022), 045002.
- [38] A. Kitaev, Quantum computations; algorithms and error correction, *Russian Math. Surveys* **52** (1997), 1191–1249.
- [39] S. Koudia, A. S. Cacciapuoti, K. Simonov, M. Caleffi, How Deep the Theory of Quantum Communications Goes: Superadditivity, Superactivation and Causal Activation, *IEEE Communications Surveys & Tutorials* **24** (2022), 1926–1956.
- [40] L. Li, K. Bu, D. E. Koh, A. Jaffe, S. Lloyd, Wasserstein Complexity of Quantum Circuits, arXiv:2208.06306.
- [41] L. Narici, E. Beckenstein, *Topological vector spaces*, Second edition, Pure and Applied Mathematics (Boca Raton), 296, CRC Press, Boca Raton, FL, 2011.
- [42] L. Ning, T. T. Georgiou, Metrics between matrix-valued measures via test functions, in *53rd IEEE Conference on Decision and Control*, pages 2642–2647, IEEE, 2014.
- [43] D. Ornstein, An application of ergodic theory to probability theory, *Ann. Probability* **1** (1973), 43–65.
- [44] M. A. Rieffel, Metrics on state spaces, *Doc. Math.* **4** (1999), 559–600.
- [45] C. Rouzé, N. Datta, Concentration of quantum states from quantum functional and transportation cost inequalities, *J. Math. Phys.* **60** (2019), 012202.
- [46] C. Rouzé, D. Stilck França, Learning quantum many-body systems from a few copies, arXiv:2107.03333.
- [47] E. K. Ryu, Y. Chen, W. Li, S. Osher, Vector and matrix optimal mass transport: theory, algorithm, and applications, *SIAM J. Sci. Comput.* **40** (2018), A3675–A3698.
- [48] C. Villani, *Topics in optimal transportation*, Graduate Studies in Mathematics, 58, American Mathematical Society, Providence, RI, 2003.
- [49] T. V. Vu, K. Saito, Thermodynamic unification of optimal transport: thermodynamic uncertainty relation, minimum dissipation, and thermodynamic speed limits, *Phys. Rev. X* **13** (2023), 011013.
- [50] J. Watrous, *The theory of quantum information*, Cambridge University Press, Cambridge, 2018.
- [51] M. Wirth, A noncommutative transport metric and symmetric quantum Markov semigroups as gradient flows of the entropy, arXiv:1808.05419.

DEPARTMENT OF PHYSICS, UNIVERSITY OF PRETORIA, PRETORIA 0002, SOUTH AFRICA

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