Lattice-based Approximate and Robust Majorization

Maxence, Jauberty `firstname.lastname@telecom-paris.fr`[1]

[1]LTCI, Télécom Paris, Institut Polytechnique de Paris, France

***Abstract*—In this paper, we study the majorization lattice within the framework of Local Operations and Classical Communication (LOCC) protocols for bipartite quantum systems. In particular, we address two key problems: the approximation of LOCC transformations and ϵ-robustness under uncertainty. The first problem considers whether, given a desired transformation from a state ρ^{AB} to σ^{AB} that is not theoretically possible, one can identify a state ${\sigma'}^{AB}$ close to σ^{AB} such that ρ^{AB} can be transformed into ${\sigma'}^{AB}$ via LOCC. The second problem investigates whether a transformation from ρ^{AB} to σ^{AB} remains possible when ρ^{A} is known with an uncertainty ϵ, with respect to a given metric.**

Using Nielsen's theorem, these problems are formalized on the majorization lattice, focusing on the majorization of Schmidt coefficients (i.e., the eigenvalues of the reduced state ρ^{A}). We derive a formula for the greatest radius of ϵ-robust majorization between the Schmidt coefficient vectors λ^{ρ} and λ^{σ} in the ℓ^1 metric. Furthermore, we establish a formula for the greatest ϵ for ϵ-robustness approximate majorization. Finally, we demonstrate that an ϵ-robust majorization between Schmidt coefficients is equivalent to an ϵ-robust LOCC transformation between the states ρ^{AB} and σ^{AB}.

***Keywords*—** majorization lattice, approximate majorization, LOCC, robust majorization, trumping majorization

CONTENTS

I. Outline

In section II, we recall important definitions and results such as majorization and its lattice structure, “Local Operation and Classical Communication” (LOCC) protocols and Nielsen’s theorem. In section III, we recall results from [9] and [11] on approximate majorization. In section IV, we study the robustness of majorization relation under perturbation. We begin with some recall and proof of trivial lemmas. We then find a formula for the greatest perturbation allowed (subsection IV-A) to maintain a majorization relation between two vectors. Finally, we extend this result whenever we consider perturbation on density matrices (subsection IV-B). In section V, we use the results of the previous parts to study the robustness of approximate majorization (subsection V-A), the robustness of a separability criterion based on majorization (subsection V-B) and robust and approximate trumping majorization (subsection V-C).

II. Preliminaries

The theory of majorization has been quite popular in diverse topics such as matrix theory, economics combinatorics [10][4] and in entanglement theory [12][15]. It defines a general framework to study the notion of “non-uniformity”. Intuitively, a vector x is majorized by y, or $x \preccurlyeq y$, if x is “more uniform” than y.

Notation :

- Δ_d denotes the set of $d-$ary probability distributions arranged in non-increasing order, i.e. $x \in \mathbb{R}_+^d$ such that $x_1 \geqslant \ldots \geqslant x_d$ and $\|x\|_1 = 1$.
- Let $x \in \Delta_d$, $\mathfrak{m}(x)$ denotes the cardinal of $\mathrm{supp}(x)$, i.e. the number of non-null components of x.
- $u \in \Delta_d$ denotes the uniform distribution, that is $u = \left(\frac{1}{d}, \ldots, \frac{1}{d}\right)$.
- δ denotes the dirac distribution, that is $\delta = (1, \ldots, 0)$.
- $\|x\|_p$ denotes the usual ℓ^p norm on $\mathbb{R}^d$.
- If $\mathcal{H}$ is an Hilbert space, denote $\mathcal{D}(\mathcal{H})$ the set of density operators ρ acting on $\mathcal{H}$, i.e. the hermitian positive semi-definite ρ such that $\mathrm{Tr}\,(\rho) = 1$ see Appendix VIII-A.
- If ρ is a density operator acting on a Hilbert space of dimension d, denote $\lambda(\rho) \in \Delta_d$ the eigenvalues of ρ in non-increasing order.
- For any $a \in \mathbb{R}$, we denote $(a)^+$ the quantity $\max\{a, 0\}$.

Numerous equivalent definitions of majorization have been found, see [10]. Here, we have considered the definition which seem to be the most appropriate in our context, namely the “Lorenz curves” or “cumulative sum” definition.

Definition 1. *Let $x, y \in \Delta_d$, we say that x is majorized by y, denoted as $x \preccurlyeq y$, if*

$$\forall k \in \{1, \ldots, d-1\}, \sum_{i=1}^{k} x_i \leqslant \sum_{i=1}^{k} y_i. \tag{1}$$

We denote $S(y)$ the set of $x \in \Delta_d$ such that $x \preccurlyeq y$.

Remark 1. *Whenever $\preccurlyeq$ is defined on $\mathbb{R}^d$, it requires to also verify $\sum_{i=1}^d x_i = \sum_{i=1}^d y_i$. Since we consider only vectors of Δ_d, this condition is trivially verified.*

The relation of majorization $\preccurlyeq$ defines a partial order on Δ_d. $(\Delta_d, \preccurlyeq)$ admits δ as greatest element and u as least element [10]. The majorization is not a total order, in figure 1, we give an example of vectors x, y such that neither $x \preccurlyeq y$ or $y \preccurlyeq x$ are verified.

Remark 2. *We define the Lorenz curve associated to $x \in \Delta_d$ as the curve obtained by connecting the points $\left(\frac{k}{d}, \sum_{i=1}^k x_i\right)_{0 \leqslant k \leqslant d}$. Hence, $x \preccurlyeq y$ is equivalent to having the Lorenz curve of x under the Lorenz curve of y. See figure 1.*

Remark 3. *It has been shown by Hardy, Littlewood and Polya (see [10]) that $x \preccurlyeq y$ if and only if there exists a doubly stochastic matrix P such that $x = Py$.*

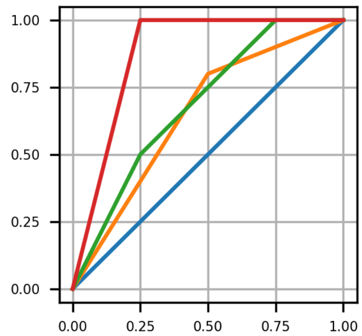


Fig. 1: Example of (non-)majorization between u (blue), $x = (0.4, 0.4, 0.1, 0.1)$ (orange), $y = (0.5, 0.25, 0.25, 0)$ (green) and δ (red). $u \preccurlyeq x \preccurlyeq \delta$ and $u \preccurlyeq y \preccurlyeq \delta$. Moreover, $x \npreccurlyeq y$ and $y \npreccurlyeq x$.

An analogous order has been defined on density matrices considering the “doubly stochastic” characterization. Quite remarkably, this definition can only be reduced to a classical majorization on eigenvalues thanks to Uhlmann theorem (more details are given in Appendix and in [1]).

Definition 2. *Let ρ, σ be two density matrices acting on an Hilbert space $\mathcal{H}$, we say that ρ is majorized by σ, denoted as $\rho \preccurlyeq \sigma$ if $\lambda(\rho) \preccurlyeq \lambda(\sigma)$.*

Denote $S(\sigma)$ the set of density matrices acting on $\mathcal{H}$ such that $\rho \preccurlyeq \sigma$.

A. LOCC Protocol and Nielsen’s Theorem

Among the classes of quantum communication protocols, the *“distant lab”* framework is a well-studied subject [6]. In these protocols, participants are only allowed to use a classical channel to send classical bits and to manipulate independently their reduced state. Using a prior shared entanglement between the involved parties can allow them to indeed share *quantum information*. One of the most famous illustration is the protocol of quantum teleportation *Bennett et al* [3]. Here, we introduce the class of Local Operations and Classical Communication for bipartite systems. Although considering bipartite systems is restrictive, it has been shown that a multipartite system can be simulated with bipartite systems as long as we do not restrict the size of the composite systems [16].

Assume that Alice and Bob share a density operator ρ^{AB}. They do not have an access to this state but only to their reduced operators ρ^A or ρ^B. What kind of information can be shared between Alice and Bob ? The precise resource shared is entanglement and is quantified by Schmidt coefficients [14].

Definition 3 (Schmidt Coefficients)**.** *Let $\mathcal{H}_A$ and $\mathcal{H}_B$ be Hilbert spaces, and $|\psi\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$ a normalized vector. There exists $\lambda_1, \ldots, \lambda_n$, $\{|i_A\rangle\}_i$ and $\{|i_B\rangle\}_i$ respectively sets of orthonormal vectors of $\mathcal{H}_A$ and $\mathcal{H}_B$ such that*

$$\rho^{AB} = |\psi\rangle\langle\psi| = \sum_{1 \leqslant i,j \leqslant n} \sqrt{\lambda_i \lambda_j}\, |i_A\rangle\langle j_A| \otimes |i_B\rangle\langle j_B| \tag{2}$$

where $n = \min\{\dim(\mathcal{H}_A), \dim(\mathcal{H}_B)\}$. $\sqrt{\lambda_1}, \ldots, \sqrt{\lambda_n}$ are called the Schmidt coefficients and are unique up to re-ordering.

Notation 1. *Given ρ a density matrix acting on a space $\mathcal{H}_A \otimes \mathcal{H}_B$ such that $\rho^{AB} = |\psi\rangle\langle\psi|$ for some $|\psi\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$, λ^ρ denote the λ introduced in the Schmidt decomposition non-increasingly ordered.*

One can remark [14] that the reduced operators ρ^A and ρ^B, i.e. the state manipulated respectively by Alice and Bob, share the same eigenvalues λ^ρ.

Definition 4 (LOCC)**.** *The protocol of communication Local Operations and Classical Communication (LOCC) allows Alice and Bob to apply any Local Operations (LO) on their reduced states, such as measurements and unitary transformation, and to communicate through a Classical Channel (CC).*

Remark 4. *In figure 2, we illustrate the LOCC setup. One can think of LOCC as a way of communication in which Alice and Bob can only "act" on their reduced states.*

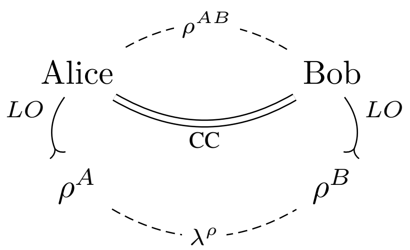


Fig. 2: Scheme of a bipartite LOCC protocol, ρ^{AB} is a pure state and Alice and Bob can only manipulate their reduced states. They are sharing the Schmidt coefficients λ^ρ.

It is natural to ask whether we can transform deterministically a shared state ρ^{AB} into another, say, σ^{AB} through a LOCC protocol. A celebrated result of entanglement theory, namely Nielsen's theorem [12], shows that the evolution of a communication under LOCC protocol is entirely characterized by relations of majorization.

Theorem 1. *Alice and Bob can transform their shared pure state ρ^{AB} into the pure state σ^{AB} using a LOCC protocol if and only if $\lambda^\rho \preccurlyeq \lambda^\sigma$. This is generally written as*

$$\rho^{AB} \xrightarrow{LOCC} \sigma^{AB} \quad \textit{if and only if} \quad \lambda^\rho \preccurlyeq \lambda^\sigma. \tag{3}$$

From Alice's point of view, it is possible to transform her state ρ^A into σ^A if and only if $\lambda\left(\rho^A\right) \preccurlyeq \lambda\left(\sigma^A\right)$, i.e. $\rho^A \preccurlyeq \sigma^A$.

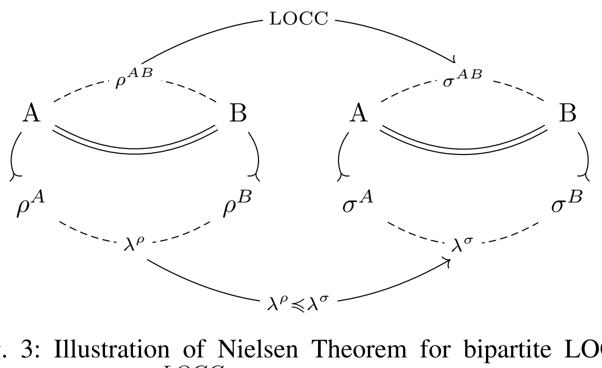


Fig. 3: Illustration of Nielsen Theorem for bipartite LOCC conversion. $\rho^{AB} \xrightarrow{LOCC} \sigma^{AB}$ if and only if $\lambda^\rho \preccurlyeq \lambda^\sigma$. This allows Alice and Bob to go from the system on the left to the system on the right.

B. Majorization Lattice

A lattice is a quadruplet $(\mathcal{L}, \preccurlyeq, \vee, \wedge)$ such that $\mathcal{L}$ is a set partially ordered with respect to $\preccurlyeq$ and for every two elements $x, y \in \mathcal{L}$, there exists a greatest lower bound $x \wedge y$ with respect to $\preccurlyeq$ and a least upper bound $x \vee y$ with respect to $\preccurlyeq$. It is well-known that the majorization order induces a lattice stucture on Δ_d (see [2]).

Theorem 2 (Majorization lattice)**.** *Δ_d ordered by the majorization relation $\preccurlyeq$ can be endowed with a lattice structure.*

Moreover, the lattice $(\Delta_d, \preccurlyeq, \wedge, \vee)$ is a *complete* lattice [2], that is for any non-empty subset S of Δ_d there exists an infimum and a supremum of S in the sense of majorization.

Proposition 1. *Let $S \subset \Delta_d$ non-empty. There exists two unique vectors $\bigwedge S$ and $\bigvee S$ such that*

- 1) $\forall x \in S,\ \bigwedge S \preccurlyeq x$
  2) *If x is such that for any $y \in S$, we have $x \preccurlyeq y$, then $x \preccurlyeq \bigwedge S$.*
- 1) $\forall x \in S,\ x \preccurlyeq \bigvee S$
  2) *If x is such that for any $y \in S$, we have $y \preccurlyeq x$, then $\bigvee S \preccurlyeq x$.*

Proof: The proof of the completeness of the majorization lattice is given in [2]. More details on the construction of such vectors are given in [5]. ■

Remark 5. *Note that $\bigwedge S$ or $\bigvee S$ may not be in S. A trivial example is given whenever $x \not\preccurlyeq y$, then $\bigwedge\{x, y\} = x \wedge y$ and $x \wedge y$ is different than x or y. The same argument holds for $\bigvee\{x, y\}$.*

C. Steepest and Flattest Approximation

Horodecki et al. have introduced the notions of steepest and flattest approximations in [9], these are defined to be the supremum and infimum (in the sense of majorization) of the vectors in Δ_d within a distance of a given vector $x \in \Delta_d$. These are useful whenever one wants to study majorization with approximations [11].

Notation :

- An exponent $^{(D)}$ indicates that the object is defined with respect to a given metric D. For ℓ^p metrics, we write $^{(p)}$. Whenever there is no ambiguity on the D, we may omit the exponent.
- $\mathcal{B}_\epsilon^{(D)}(x)$ denote the ball of radius ϵ and of center x for D intersected with Δ_d, i.e.

$$\mathcal{B}_\epsilon^{(D)}(x) = \{x' \in \Delta_d,\ D(x, x') \leqslant \epsilon\}. \tag{4}$$

Definition 5. *Let $x \in \Delta_d$, D a metric on $\mathbb{R}^d \times \mathbb{R}^d$ and $\epsilon \geqslant 0$, we define the following vectors respectively the steepest and flattest $\epsilon-$approximations [9][11].*

$$\underline{x}_\epsilon^{(D)} = \bigwedge \mathcal{B}_\epsilon^{(D)}(x) \tag{5}$$

$$\overline{x}_\epsilon^{(D)} = \bigvee \mathcal{B}_\epsilon^{(D)}(x). \tag{6}$$

Remark 6. *The term "flattest" and "steepest" are references to the constructions of $\underline{x}_\epsilon^{(1)}$ and $\overline{x}_\epsilon^{(1)}$, see Appendix VIII-B and [9].*

Remark 7. *We can compute $\overline{x}_\epsilon^{(\infty)}$ and $\underline{x}_\epsilon^{(\infty)}$ using the algorithms presented in [11]. In Appendix VIII-B, we recall the form and the computation for ℓ^1 metric given in [9].*

Steepest and flattest approximations have already been studied for ℓ^p-distances [11] [9]. In particular, it has been shown in [11] that $\underline{x}_\epsilon^{(p)}$ and $\overline{x}_\epsilon^{(p)}$ are in $\mathcal{B}_\epsilon^{(p)}(x)$ if and only if $p = 1$ or $p = \infty$. In [11], it has been shown that $\epsilon \mapsto \overline{x}_\epsilon^{(\infty)}$ is additive, i.e. $\overline{\left(\overline{x}_{\epsilon_1}^{(\infty)}\right)}_{\epsilon_2}^{(\infty)} = \overline{x}_{\epsilon_1+\epsilon_2}^{(\infty)}$. We show that this is also verified for the ℓ^1-approximation.

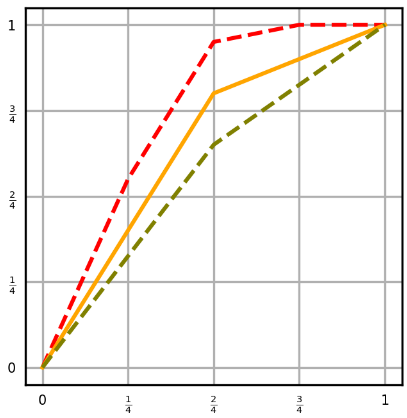


Fig. 4: Lorenz curves of $x = (0.4, 0.4, 0.1, 0.1)$, $\overline{x}_\epsilon^{(1)}$ and $\underline{x}_\epsilon^{(1)}$ for $\epsilon = 0.3$. We have $\overline{x}_\epsilon^{(1)} = (0.55, 0.4, 0.05, 0)$ and $\underline{x}_\epsilon^{(1)} = (0.325, 0.325, 0.175, 0.175)$.

Proposition 2. *Let $x \in \Delta_d$ and $\epsilon_1, \epsilon_2 > 0$, then*

$$\overline{\left(\overline{x}_{\epsilon_1}^{(1)}\right)}_{\epsilon_2}^{(1)} = \overline{x}_{\epsilon_1+\epsilon_2}^{(1)}. \tag{7}$$

Proof: See Appendix VIII-C for a proof. ■

Similarly, we can show that $\epsilon \mapsto \underline{x}_\epsilon$ is also additive for ℓ^1 distance.

Proposition 3. *Let $x \in \Delta_d$ and $\epsilon_1, \epsilon_2 > 0$, then*

$$\underline{\left(\underline{x}_{\epsilon_1}^{(1)}\right)}_{\epsilon_2}^{(1)} = \underline{x}_{\epsilon_1+\epsilon_2}^{(1)}. \tag{8}$$

Proof: See Appendix VIII-C for a proof. ■

III. Approximate Majorization

A classical problem in the study of pure bipartite LOCC protocol arises when Alice and Bob want to perform a transformation they are not allowed to accordingly to Nielsen's theorem, i.e. $\lambda^\rho \not\preccurlyeq \lambda^\sigma$. Alice and Bob can either choose to perform the LOCC protocol but may fail, this has been studied in Vidal's article, but they may also consider to go to a state σ' close to σ whose Schmidt coefficients verify $\lambda^\rho \preccurlyeq \lambda^{\sigma'}$, it has also been studied in [15]. We introduce the notion of $\epsilon-$majorization (namely $(D,\epsilon)-$pre-majorization in [11]).

Definition 6. *Let x, y in Δ_d and $\epsilon > 0$.Given a metric D, we say that x is $\epsilon-$majorized by y if there exists $y' \in \Delta_d$ such that $x \preccurlyeq y'$ and $D(y, y') \leqslant \epsilon$ and we write $x \preccurlyeq_\epsilon y$.*

This relation might be hard to study in general. However it is simply reduced to classical majorization relation (Proposition 4) when we study the ℓ^1 metric [11].

Proposition 4. *Let x, y in Δ_d and $\epsilon > 0$, then $x \underset{\epsilon}{\preccurlyeq} y$ if and only if $x \preccurlyeq \overline{y}_\epsilon^{(1)}$.*

Remark 8. *This result hold for any distance D such that $\overline{y}_\epsilon^{(D)}$. It has been shown that among the ℓ^p distance with $p \geqslant 1$, only the ℓ^1 and ℓ^∞ distances verify this property.*

Horodecki et al [9] have determined an explicit formula for the minimal ϵ to perform an approximate majorization with ℓ^1 metric (see figure 5).

Proposition 5. *Let $x, y \in \Delta_d$ and $\epsilon_{\min}(x, y)$ the minimal ϵ such that $x \preccurlyeq \overline{y}_\epsilon$, then*

$$\epsilon_{\min}(x, y) = 2 \max_{1 \leqslant k \leqslant d} \left\{ \left(\sum_{i=1}^k x_i - y_i \right)^+ \right\} \tag{9}$$

Moreover, $\epsilon_{\min}(x, y)$ is also the minimal ϵ such that $\overline{x}_\epsilon \preccurlyeq y$.

Remark 9. *Horedecki's result is slightly different since it assumes that $x \not\preccurlyeq y$. However, trivially, if $x \preccurlyeq y$ holds, then the minimal $\epsilon_{\min}(x, y) = 0$. Moreover $\max_{1 \leqslant k \leqslant d} \left\{ \left(\sum_{i=1}^k x_i - y_i \right)^+ \right\} = 0$ since $x \preccurlyeq y$.*

It has been shown in [11] that an approximate majorization between two density operators is equivalent to an approximate majorization between their eigenvalues for the ℓ^1 metric.

Proposition 6. *Let ρ, σ be two density operators acting on a space $\mathcal{H}$ and $\epsilon \geqslant 0$. Then, the following are equivalent*

1) *There exists σ' such that $\|\lambda(\sigma - \sigma')\|_1 \leqslant \epsilon$ and $\rho \preccurlyeq \sigma'$.*
2) *$\lambda(\rho) \preccurlyeq_\epsilon \lambda(\sigma)$.*

Hence, if we define $\mathcal{E}_{\min}(\rho, \sigma)$ the minimal $\epsilon \geqslant 0$ such that 1. is verified in Proposition 6, i.e. the minimal ϵ to perform an approximate majorization between ρ and σ, we have

$$\epsilon_{\min}\left(\lambda(\rho, \sigma)\right) = \mathcal{E}_{\min}(\rho, \sigma). \tag{10}$$

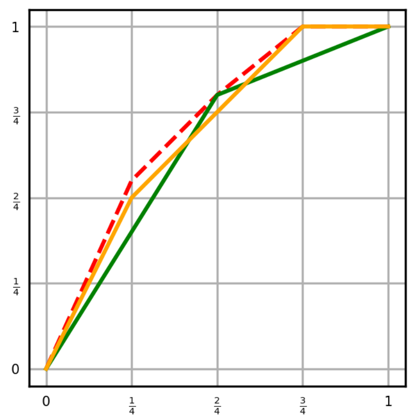


Fig. 5: Lorenz curves of $x = (0.4, 0.4, 0.1, 0.1)$ (in green), $y = (0.5, 0.25, 0.25, 0)$ (in orange) and $\overline{y}_{\epsilon_{\min(x,y)}}$ (in dashed red). Here, $\min(x, y) = 0.1$ and $x \not\preccurlyeq y$ but $x \preccurlyeq \overline{y}_{\epsilon_{\min(x,y)}}$. Intuitively, if we consider $\epsilon < \min(x, y)$, then the dashed curve would slightly below and would violate the majorization condition.

IV. Majorization With Corruption

Assume Alice and Bob are in pure bipartite LOCC setup and share a pure state ρ of squared Schmidt's coefficients λ^ρ and want to perform an LOCC transformation to σ with squared Schmidt's coefficients λ^σ. Nielsen's Theorem [12] tells us it is possible if and only if $\lambda^\rho \preccurlyeq \lambda^\sigma$. However, assume then that Alice and Bob are not exactly in the state ρ but rather in ρ' such that for a given metric D on $\mathbb{R}^d \times \mathbb{R}^d$ we have $D(\lambda^\rho, \lambda^{\rho'}) \leqslant \epsilon$. Is there a greatest radius $\epsilon_{\max}(x, y)$ such that for any $\lambda^{\rho'}$ such that $D(\lambda^\rho, \lambda^{\rho'}) \leqslant \epsilon_{\max}(x, y)$, we have $\rho' \xrightarrow{LOCC} \sigma$. In other words, we are looking the greatest ball $\mathcal{B}_\epsilon^{(D)}(\lambda^\rho)$ included in $S(\lambda^\sigma)$.

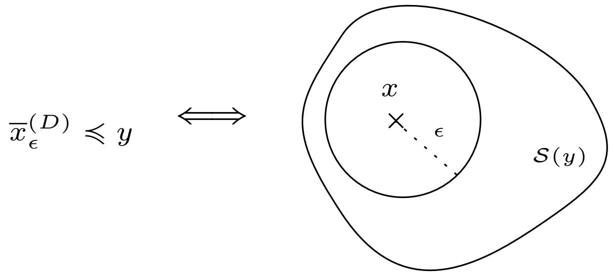


Fig. 6: Illustration of the Proposition 7

.

We consider in the following propositions D to be a metric on $\mathbb{R}^d \times \mathbb{R}^d$. The following trivial proposition gives a relatively simple necessary condition for $\mathcal{B}_\epsilon^{(D)}(x) \subset S(y)$.

Proposition 7. *Let $x, y \in \Delta_d$ and $\epsilon > 0$, then $\mathcal{B}_\epsilon^{(D)}(x) \subset S(y)$ if and only if $\overline{x}_\epsilon^{(D)} \preccurlyeq y$.*

Proof: By definition, $\overline{x}_\epsilon^{(D)}$ is the supremum of $\mathcal{B}_\epsilon^{(D)}(x)$. If $\mathcal{B}_\epsilon^{(D)}(x) \subset S(y)$, then y is greater (in the sense of $\preccurlyeq$) than any element of the ball, hence $\overline{x}_\epsilon^{(D)} \preccurlyeq y$ since $\overline{x}_\epsilon^{(D)}$ is least upper bound. Conversely, assume that $\overline{x}_\epsilon^{(D)} \preccurlyeq y$, then for any x' in the ball, we have $x \preccurlyeq \overline{x}_\epsilon^{(D)} \preccurlyeq y$, i.e. $\mathcal{B}_\epsilon^{(D)}(x) \subset S(y)$. ∎

As a corollary of Proposition 7, it is possible to characterize the interior points of $S(y)$.

Lemma 1. *Let $x, y \in \Delta_d$.x is an interior point of $S(y)$ if and only if*

$$\forall k \in \{1, \dots, \mathfrak{m}(y) - 1\}, \sum_{i=1}^{k} x_i < \sum_{i=1}^{k} y_i. \tag{11}$$

Proof: See Appendix VIII-C. ∎

A. Greatest Radius for the ℓ^1-distance

In this part, we consider D to be the ℓ^1 metric. For the sake of brevity, in the proofs $\overline{x}_\epsilon$ will be denoted as μ^ϵ. We want to establish the greatest ϵ such that $\overline{x}^\epsilon \preccurlyeq y$ (figure 7).

Lemma 2. *Let $x, y \in \Delta_d$ such that x is an interior point of $S(y)$. Let $m \leqslant d$ and $\epsilon_m = \min_{1 \leqslant k < m} \left\{\sum_{i=1}^{k} y_i - x_i\right\}$, then $\mathfrak{m}\left(\overline{x}_{\epsilon_m}\right) \geqslant m$.*

Proof: Let $\epsilon_m = 2 \min \left\{\sum_{i=1}^{k} y_i - x_i\right\}$. $\epsilon \leqslant 2(y_1 - x_1)$, then $x_1 + \frac{\epsilon}{2} \leqslant y_1 \leqslant 1$. We assume y to be different than δ. Hence $x_1 + \frac{\epsilon}{2} < 1$ and there exists l given by the construction of the steepest approximation μ^ϵ of x. Moreover, we have

$$1 < \sum_{i=1}^{l} x_i + \frac{\epsilon}{2}. \tag{12}$$

Assume $l < m$, then

$$1 < \sum_{i=1}^{l} x_i + \frac{\epsilon}{2} \leqslant \sum_{i=1}^{l} x_i + \sum_{i=1}^{l} y_i - x_i = \sum_{i=1}^{l} y_i. \tag{13}$$

Hence, by contradiction, $l \geqslant m$. ∎

Remark 10. *If $\mu^\epsilon \preccurlyeq y$, then μ^ϵ must have at least $\mathfrak{m}(y)$ non-zeros component. Therefore, Lemma 2 says that $\epsilon_{\mathfrak{m}(y)}$ is a potential candidate for a ϵ such that $x \preccurlyeq \mu^\epsilon \preccurlyeq y$.*

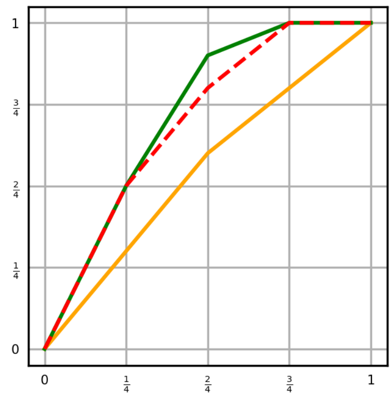


Fig. 7: Lorenz curves of $x = (0.3, 0.3, 0.2, 0.2)$ (orange), $y = (0.5, 0.4, 0.1, 0)$ (green) and $\overline{x}_{\epsilon_{\max(x,y)}}$ (dashed red). Here, $\epsilon_{\max(x,y)} = 0.4$ and $\overline{x}_{\epsilon_{\max(x,y)}} = (0.5, 0.3, 0.2, 0)$. Intuitively, if $\epsilon > \max(x, y)$ then the dashed curve would be above and would violate the majorization condition.

Theorem 3. *Let $x, y \in \Delta_d$ such that x is an interior point of $S(y)$, then for the ℓ^1 distance, we have*

$$\epsilon_{\max}(x, y) = 2 \min_{1 \leqslant k < \mathfrak{m}(y)} \left\{\sum_{i=1}^{k} y_i - x_i\right\}. \tag{14}$$

Proof: Let μ^ϵ such that $\mu^\epsilon \preccurlyeq y$. We know that μ^ϵ has at least more non-null components than y, hence for $1 \leqslant k < \mathfrak{m}(y)$, we have

$$\sum_{i=1}^{k} \mu_i^\epsilon = \sum_{i=1}^{k} x_i + \frac{\epsilon}{2} \leqslant \sum_{i=1}^{k} y_i. \tag{15}$$

Hence, $\epsilon_{max}(x, y) \leqslant \epsilon_{\mathfrak{m}(y)}$.
Consider $\epsilon = \epsilon_{\mathfrak{m}(y)}$, we know that μ^ϵ has at least $\mathfrak{m}(y)$ non-null components. Then, for $k < \mathfrak{m}(y)$, we have

$$\sum_{i=1}^{k} \mu_i^\epsilon = \sum_{i=1}^{k} x_i + \frac{\epsilon}{2} \leqslant \sum_{i=1}^{k} x_i + \sum_{i=1}^{k} y_i - x_i = \sum_{i=1}^{k} y_i. \tag{16}$$

And for $k \geqslant \mathfrak{m}(y)$, $\sum_{i=1}^{k} \mu_i^\epsilon \leqslant \sum_{i=1}^{k} y_i = 1$. Hence $\mu^\epsilon \preccurlyeq y$ and $\epsilon_m \leqslant \epsilon_{max}(x, y)$. ∎

Remark 11. *The proposition holds when x is not an interior point of $S(y)$. If x is not an interior point, trivially $\epsilon_{\max}(x, y) = 0$. Moreover, there exists $k < \mathfrak{m}(y)$ such that $\sum_{i=1}^{k} x_i = \sum_{i=1}^{k} y_i$, hence the minimum will be equal to 0.*

B. Robustness on Matrix Majorization

Whenever we are working with density matrices, a well-known metric is the so-called "Trace distance", we recall its definition

$$\|\rho - \sigma\|_1 = \|\lambda(\rho - \sigma)\|_1. \tag{17}$$

This metric can be considered analogous to the total variation distance (up to a multiplicative constant) as a measure of how much one can discriminate a state upon another [14].

Notation :

- Let ρ be a density matric acting on a space $\mathcal{H}$ an denote $\mathcal{B}_\epsilon(\rho)$ the set of density matrices ρ' acting on $\mathcal{H}$ such that $\|\rho - \rho'\|_1 \leqslant \epsilon$, i.e.

$$\mathcal{B}_\epsilon(\rho) = \left\{\rho' \in \mathcal{D}(\mathcal{H}),\ \|\rho - \rho'\|_1 \leqslant \epsilon\right\}. \tag{18}$$

Alice and Bob are now concerned about the distance between their density matrices. Hence, Alice knows that her state ${\rho'}^A$ is in $\mathcal{B}_\epsilon\left(\rho^A\right)$ and wants to be sure that it verifies ${\rho'}^A \preccurlyeq \sigma^A$.

Definition 7. *Let ρ, σ acting on an Hilbert space $\mathcal{H}$ and $\epsilon \geqslant 0$. We say that ρ is $\epsilon-$robust majorized by y when $\mathcal{B}^1_\epsilon(\rho) \subset S(\sigma)$.*

Definition 8. *Let ρ, σ acting on an Hilbert space $\mathcal{H}$. Define $\mathcal{E}_{\max}(\rho, \sigma)$ the greatest $\epsilon \geqslant 0$ such that $\mathcal{B}^1_\epsilon(\rho) \subset S(\sigma)$.*

We show in Theorem 4 that the greatest ϵ such that $\epsilon-$robust majorization is verified between two density matrices ρ and σ is equal to $\epsilon_{\max}(\lambda(\rho), \lambda(\sigma))$, i.e. the greatest ϵ such that $\epsilon-$robust majorization is verified between the eigenvalues.

Theorem 4. *Let ρ, σ acting on an Hilbert space $\mathcal{H}$. Then, we have*

$$\epsilon_{\max}(\lambda(\rho), \lambda(\sigma)) = \mathcal{E}_{\max}(\rho, \sigma). \tag{19}$$

Proof: Let $\epsilon = \epsilon_{\max}(\lambda(\rho), \lambda(\sigma))$ and $\mathcal{E} = \mathcal{E}_{\max}(\rho, \sigma)$. Consider μ such that $\|\lambda(\rho) - \mu\|_1 \leqslant \mathcal{E}$. Let Λ and M the diagonal matrices naturally defined by $\lambda(\rho)$ and μ. There exists a unitary such that $U\Lambda U^\dagger = \rho$. Hence

$$\|\rho - UMU^\dagger\|_1 = \|\Lambda - M\|_1 = \|\lambda(\rho) - \mu\|_1 \leqslant \mathcal{E}. \tag{20}$$

By definition of $\mathcal{E}$, we have $UMU^\dagger \preccurlyeq \sigma$, hence $\lambda(M) \preccurlyeq \lambda(\sigma)$, i.e. $\mu \preccurlyeq \lambda(\sigma)$. Therefore, $\mathcal{E} \leqslant \epsilon$.

Consider ρ' such that $\|\rho - \rho'\|_1 \leqslant \epsilon$. We can apply Lidskii's theorem to obtain (21)

$$\|\lambda(\rho) - \lambda(\rho')\|_1 \leqslant \|\rho - \rho'\|_1 \leqslant \epsilon. \tag{21}$$

Hence, by definition of ϵ, $\lambda(\rho') \preccurlyeq \lambda(\sigma)$. Therefore, $\rho' \preccurlyeq \sigma$. Thus, $\epsilon \leqslant \mathcal{E}$. ∎

As a trivial corollary, we can show that an $\epsilon-$robust majorization between density matrices is verified if and only if an $\epsilon-$robust majorization is verified between their eigenvalues.

Corollary 1. *Let ρ, σ acting on an Hilbert space $\mathcal{H}$, then ρ is $\epsilon-$robust majorized by σ if and only if $\lambda(\rho)$ is $\epsilon-$robust majorized by $\lambda(\sigma)$*

Proof: $\lambda(\rho)$ is $\epsilon-$robust majorized by $\lambda(\sigma)$ if and only if $\epsilon \leqslant \epsilon_{\max}(\lambda(\rho), \lambda(\sigma))$. Using Proposition 4, we have that $\epsilon \leqslant \mathcal{E}(\rho, \sigma) = \epsilon_{\max}(\lambda(\rho), \lambda(\sigma))$. Hence, $\lambda(\rho)$ is $\epsilon-$robust majorized by $\lambda(\sigma)$ if and only if $\epsilon \leqslant \mathcal{E}(\rho, \sigma)$, which is equivalent to ρ being $\epsilon-$robust majorized by σ. ∎

To enlighten how this result might be considered for LOCC protocol, we restate Nielsen's theorem from the point of view of Alice (or Bob). $\rho \xrightarrow{LOCC} \sigma$ if and only if $\rho^A \preccurlyeq \sigma^A$ (equivalently $\rho^B \preccurlyeq \sigma^B$. From Proposition 4, we have

$$\epsilon_{\max}(\lambda^\rho, \lambda^\sigma) = \mathcal{E}_{\max}(\rho^A, \sigma^A). \tag{22}$$

V. Applications

In this section, we will consider only the ℓ^1 metric and its matrix analogous metric the trace distance.

A. Robustness on Approximate Majorization

Consider that Alice and Bob are no more interested into having exactly σ and accept that there exists $\epsilon_\rho, \epsilon_\sigma$ such that $D(\lambda_\rho, \lambda'_\rho) \leqslant \epsilon_\rho$ and $D(\lambda_\sigma, \lambda'_\sigma) \leqslant \epsilon_\sigma$. Precisely, they want, knowing that ρ might be corrupted, to go to a state *"not-so-far"* from the state σ. Proposition 8 states that this can be characterized using a majorization relation, which can be considered as an application of approximate majorization and robust majorization.

Proposition 8. *Let $x, y \in \Delta_d$, $\epsilon_x \geqslant 0$ and $\epsilon_y \geqslant 0$. There exists $y' \in \mathcal{B}_{\epsilon_y}(y)$ such that $\mathcal{B}_{\epsilon_x}(x) \subset S(y')$ if and only if $\overline{x}_{\epsilon_x} \preccurlyeq \overline{y}_{\epsilon_y}$.*

Proof: Assume $\overline{x}_{\epsilon_x} \preccurlyeq \overline{y}_{\epsilon_y}$. Then, we take $y' = \overline{y}_{\epsilon_y}$ and by the result of the previous part $\mathcal{B}_{\epsilon_x}(x) \subset S(y')$.

Conversely, assume there exists y' as defined above then $\mathcal{B}_{\epsilon_x}(x) \subset S(y') \subset S(\overline{y}_{\epsilon_y})$ by the definition of $\overline{y}_{\epsilon_y}$. Therefore, $\overline{x}_{\epsilon_x} \preccurlyeq \overline{y}_{\epsilon_y}$. ∎

Remark 12. *Note that Proposition 8 is equivalent to : for any $x' \in \mathcal{B}_\epsilon(x)$, there exists $y' \in \mathcal{B}_\epsilon(y)$ such that $x' \preccurlyeq y'$.*

A robust transformation protocol would require to have the greatest ϵ_x and the lowest ϵ_y. The case where $\epsilon = \epsilon_x = \epsilon_y$ and $x \preccurlyeq y$ has already been studied in [9].

Proposition 9. *Let $x, y \in \Delta_d$ such that $x \preccurlyeq y$. Then for any $\epsilon \geqslant 0$, we have*

$$\overline{x}_\epsilon \preccurlyeq \overline{y}_\epsilon. \tag{23}$$

Remark 13. *This cannot provide directly "optimal" ϵ_x or ϵ_y. However, if we allow ϵ to be small, say that Alice and Bob are almost certain that thay have produced vector x, then they can perform a majorization relation to some vector close to y.*

As a trivial corollary of Theorem 3, we find a formula for the greatest radius for an $\epsilon-$approximation of majorization, i.e. $\epsilon_{\max}\left(x, \overline{y}_\epsilon\right)$

Corollary 2. *Let $x, y \in \Delta_d$ and $\epsilon \geqslant 0$ such that $x \preccurlyeq \overline{y}_\epsilon$, then*

$$\epsilon_{\max}\left(x, \overline{y}_\epsilon\right) = 2 \min_{1 \leqslant k < \mathfrak{m}(\overline{y}_\epsilon)} \left\{ \sum_{i=1}^k y_i - x_i \right\} + \epsilon. \tag{24}$$

Proof: Denote $m = \mathfrak{m}\left(\overline{y}_\epsilon\right)$.

$$\begin{aligned} \epsilon_{\max}(x, \overline{y}_\epsilon) &= 2 \min_{1 \leqslant k < m} \left\{ \sum_{i=1}^k \left(\overline{y}_\epsilon\right)_i - x_i \right\} && (25) \\ &= 2 \min_{1 \leqslant k < m} \left\{ \sum_{i=1}^k y_i - x_i + \frac{\epsilon}{2} \right\} && (26) \\ &= 2 \min_{1 \leqslant k < m} \left\{ \sum_{i=1}^k y_i - x_i \right\} + \epsilon. && (27) \end{aligned}$$

(26) comes from the construction of $\overline{y}_\epsilon$, for any $k < m$ $\sum_{i=1}^k \left(\overline{y}_\epsilon\right)_i = \sum_{i=1}^k y_i + \frac{\epsilon}{2}$. ∎

Remark 14. *If $x \preccurlyeq y$, then $2 \min_{1 \leqslant k < \mathfrak{m}(\overline{y}_\epsilon)} \left\{ \sum_{i=1}^k y_i - x_i \right\} \geqslant 0$. For any ϵ, we have $\epsilon \leqslant \epsilon_{\max}\left(x, \overline{y}_\epsilon\right)$. Therefore, we can prove as a corollary the Proposition 9.*

Remark 15. *This result can also be used to find $\epsilon_{\min}(x, y)$. Let x, y such that $x \npreccurlyeq y$, then by definition of $\epsilon_{\min}(x, y)$, $x \preccurlyeq \overline{y}_{\epsilon_{\max}(x,y)}$. Moreover, it would saturate one of the inequality of the majorization, i.e. x would be a boundary point of $S\left(\overline{y}_{\epsilon_{\max}(x,y)}\right)$. As a consequence, $\epsilon_{\max}\left(\overline{y}_{\epsilon_{\min}(x,y)}\right) = 0$. Denote $m = \mathfrak{m}\left(\overline{y}_{\epsilon_{\min}(x,y)}\right)$ Hence, using the formula*

$$2 \min_{1 \leqslant k < m} \left\{ \sum_{i=1}^k y_i - x_i \right\} + \epsilon_{\min}\left(x, y\right) = 0. \tag{28}$$

Hence,

$$\epsilon_{\min}\left(x, y\right) = 2 \max_{1 \leqslant k < m} \left\{ \sum_{i=1}^k x_i - y_i \right\}. \tag{29}$$

B. Robustness of a Separability Criterion

We recall definitions from Appendix VIII-A. If ρ is a density operator acting on a tensor space $\mathcal{H}_1 \otimes \mathcal{H}_2$, it is said to be a product state whenever there exists ρ_1, ρ_2 such that $\rho = \rho_1 \otimes \rho_2$. ρ is separable if it is a convex combination of product states. If ρ is not separable, it is said to be entangled.

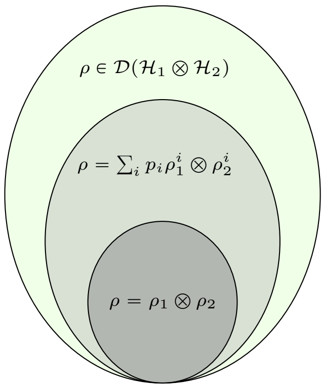


Fig. 8: Inclusions between the sets of all states, separable states and product states. Entangled states are in the clearest part of the drawing.

A well-known problem in entanglement theory is to know whether there is *actually* some entanglement, or the state shared is separable. This problem is considered to be NP-hard in general and has been proven NP-hard for specific cases [8]. In the bipartite case, one can use a separability criterion based on majorization. We will then use Proposition 5 from [9] and Theorem 3 to study the robustness of this criterion.

Proposition 10. *If ρ^{AB} is separable, then $\rho^{AB} \preccurlyeq \rho^{A}$ and $\rho^{AB} \preccurlyeq \rho^{B}$.*

Proof: A proof can be found in [13]. ∎

Remark 16. *Notice that ρ^{AB} and ρ^{A} are not acting on the same Hilbert space. Hence, we have a problem of homogeneity for the eigenvalues. Usually, to consider majorization between vectors of different size, we add $0'$s to the shortest vector.*

Remark 17. *Equivalently, if either $\rho^{AB} \npreccurlyeq \rho^{A}$ or $\rho^{AB} \npreccurlyeq \rho^{B}$, then ρ^{AB} is entangled.*

This criterion may fail for two different reasons. Either entanglement is theoretically detected but practically not detected, or there is no entanglement theoretically but entanglement is detected.

Definition 9. *We say that the criterion is $\epsilon-$corrupted for Alice whenever we have :*

- *False Entanglement Detection : $\rho^{AB} \preccurlyeq \rho^{A}$ and there exists $\rho'^{AB} \in \mathcal{B}_\epsilon\left(\rho^{AB}\right)$ such that $\rho'^{AB} \npreccurlyeq \rho^{A}$*
- *No Entanglement Detection : $\rho^{AB} \npreccurlyeq \rho^{A}$ and there exists $\rho'^{AB} \in \mathcal{B}_\epsilon\left(\rho^{AB}\right)$ such that $\rho'^{AB} \preccurlyeq \rho^{A}$.*

We say that the criterion is $\epsilon-$corrupted whenever it is $\epsilon-$corrupted for Alice or for Bob.

Proposition 11. *Assume that $\rho^{AB} \npreccurlyeq \rho^{A}$. The separability criterion is not $\epsilon-$corrupted for Alice if and only if $\epsilon < \mathcal{E}_{\min}\left(\rho^{AB}, \rho^{A}\right)$.*

Proof: The criterion is $\epsilon-$corrupted if and only if $\lambda\left(\rho^{AB}\right)^{\epsilon} \preccurlyeq \lambda\left(\rho^{A}\right)$. By definition, this is equivalent to $\epsilon \geqslant \mathcal{E}_{\min}\left(\rho^{AB}, \rho^{A}\right)$. ∎

Proposition 12. *Assume that $\rho^{AB} \preccurlyeq \rho^{A}$. The separability criterion is not $\epsilon-$corrupted if and only if $\epsilon \leqslant \mathcal{E}_{\max}\left(\rho^{AB}, \rho^{A}\right)$.*

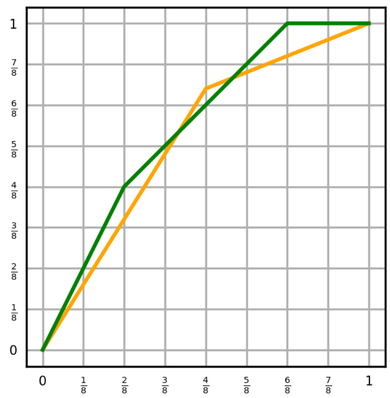
Proof: The criterion is not $\epsilon-$corrupted for Alice if and only $\mathcal{B}_\epsilon\left(\rho^{AB}\right) \subset S\left(\rho^{A}\right)$. By definition, this is equivalent to $\epsilon \leqslant \mathcal{E}_{\max}\left(\rho^{AB}, \rho^{A}\right)$ ∎

C. Robust and Approximate Trumping Majorization

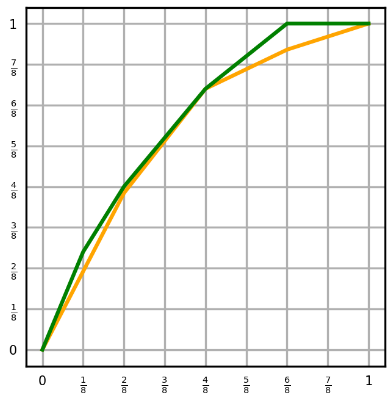
Trumping majorization, or catalytic majorization, have been mainly introduced in the study of LOCC. It has been noticed that some states were inconvertible, i.e. $\lambda^{\rho} \npreccurlyeq \lambda^{\sigma}$, but whenever we consider a well-chosen catalysis, say $c \in \Delta_m$, we have $\lambda^{\rho} \otimes c \preccurlyeq \lambda^{\sigma} \otimes c$ (figure 9). Here, $z \otimes c$ denotes the tensor product of z and c, that is

$$z \otimes c = \left(z_{(i)} c_{(j)}\right)_{i,j} \tag{30}$$

where $\left(z_{(i)}\right)_i$ and $\left(c_{(j)}\right)_j$ are permutations of z and c such that $z \otimes c$ is arranged in non-increasing order.



(a)



(b)

Fig. 9: We consider $x = (0.4, 0.4, 0.1, 0.1)$ (orange) and $y = (0.5, 0.25, 0.25, 0)$ (green), we recall in (a) that $x \npreccurlyeq y$ with their Lorenz curves. In (b), we have plotted the Lorenz curves of $x \otimes c$ and of $y \otimes c$ with $c = (0.6, 0.4)$, we can see that $x \otimes c \preccurlyeq y \otimes c$.

Definition 10. *Let $x, y \in \Delta_d$, x is said to be trumped by y, denoted as $x \preccurlyeq_T y$ if there exists $m \in \mathbb{N}$ and $c \in \Delta_m$ such that $x \otimes c \preccurlyeq y \otimes c$.*

We denote $\mathcal{T}(y)$ the set of vectors trumped by y.

$\preccurlyeq_T$ is known to be a partial order on Δ_d. More details on this can be found in [7]. Whenever $d \geqslant 4$, there exists vectors $x, y \in \Delta_d$ such that $x \npreccurlyeq y$ and $x \preccurlyeq_T y$. See figure 9b, we can show that $x \otimes c \preccurlyeq y \otimes c$ for $c = (0.6, 0.4)$ and $x \npreccurlyeq y$.

Remarkably, both robustness and approximation of trumping majorization can be characterized using steepest and flattest approximations.

Proposition 13. *Let $x, y \in \Delta_d$ and $\epsilon \geqslant 0$, $\mathcal{B}_\epsilon(x) \subset \mathcal{T}(y)$ if and only if $\overline{x}^{\epsilon} \preccurlyeq_T y$.*

Proof: If $\mathcal{B}_\epsilon(x) \subset \mathcal{T}(y)$, then $\overline{x}^\epsilon \preccurlyeq_T y$ because $\overline{x}^\epsilon \in \mathcal{B}_\epsilon(x)$.
Conversely, assume that $\overline{x}^\epsilon \preccurlyeq_T y$. Then, for any $x' \in \mathcal{B}_\epsilon(x)$

$$x' \preccurlyeq \overline{x}^\epsilon \preccurlyeq_T y. \tag{31}$$

Since $x' \preccurlyeq \overline{x}^\epsilon$ implies $x' \preccurlyeq_T \overline{x}^\epsilon$, we have $x' \preccurlyeq_T y$. ∎

Proposition 14. *Let $x, y \in \Delta_d$ and $\epsilon \geqslant 0$, the following are equivalent*

1) *There exists $y' \in \mathcal{B}_\epsilon(y)$ such that $x \preccurlyeq_T y'$*
2) $x \preccurlyeq_T \overline{y}^\epsilon$.

Proof: Assume that there exists $y' \in \mathcal{B}_\epsilon(y)$ such that $x \preccurlyeq_T y'$.

$$x \preccurlyeq_T y' \preccurlyeq \overline{y}^\epsilon. \tag{32}$$

Hence, $x \preccurlyeq_T \overline{y}^\epsilon$.
Conversely, assume that $x \preccurlyeq_T \overline{y}^\epsilon$. Since $\overline{y}^\epsilon \in \mathcal{B}_\epsilon(y)$, we can choose $y' = \overline{y}^\epsilon$. ∎

M.Klimesh and S.Daftuar have already given in [7] a sufficient condition for a point to be in the interior of $\mathcal{T}(y)$. Using the previous proposition, we give a necessary condition whenever $\mathfrak{m}(y) = d$, i.e. y has only non-null coordinates.

Proposition 15. *Let $x, y \in \Delta_d$ such that $\mathfrak{m}(y) = d$, x is an interior point of $\mathcal{T}(y)$ if and only if $x \preccurlyeq_T y$ and $x_1 < y_1$ and $x_d > y_d$.*

Proof: The sufficiency has already been proven in [7]. Let x be an interior point. Let $\epsilon > 0$ small enough such that $\mathfrak{m}(\overline{x}^\epsilon) = d$ and $\mathcal{B}_\epsilon(x) \subset \mathcal{T}(y)$. Hence, we have trumping majorization with a catalyst, say $c \in \Delta_m$,

$$\overline{x}^\epsilon \otimes c \preccurlyeq y \otimes c. \tag{33}$$

We have $(\overline{x}^\epsilon \otimes c)_1 = \left(x_1 + \frac{\epsilon}{2}\right) c_1$, and $(y \otimes c)_1 = y_1 c_1$. Therefore

$$\left(x_1 + \frac{\epsilon}{2}\right) c_1 \leqslant y_1 c_1 \tag{34}$$

Hence $x_1 + \frac{\epsilon}{2} \leqslant y_1$, therefore $x_1 < y_1$.
Similarly, $(\overline{x}^\epsilon \otimes c)_{md} = \left(x_d - \frac{\epsilon}{2}\right) c_m$ and $(y \otimes c)_{md} = y_d c_m$. Hence, by majorization

$$\left(x_d - \frac{\epsilon}{2}\right) c_m \geqslant y_d c_m. \tag{35}$$

Thus, $x_d > y_d$. ∎

Proposition 16. *Let $x, y \in \Delta_d$ with $\mathfrak{m}(y) = d$, denote $\epsilon_{\max}(x,y)^T$ the greatest ϵ such that $\overline{x}^\epsilon \preccurlyeq_T y$, then*

$$\epsilon_{\max}(x,y) \leqslant \epsilon_{\max}(x,y)^T \leqslant 2\min\{y_1 - x_1, x_d - y_d\}. \tag{36}$$

Proof: Denote $\epsilon_n = \epsilon_{\max}(x,y)^T - 2^{-n}$. For n large enough, $\overline{x}^{\epsilon_n}$ is an interior point. There exists $\alpha > 0$ such that $\epsilon_n + \alpha < \epsilon_{\max}(x,y)^T$.

$$\overline{\overline{x}^{\epsilon_n}}^\alpha = \overline{x}^{\epsilon_n + \alpha} \preccurlyeq \overline{x}^{\epsilon_{\max}(x,y)^T} \preccurlyeq_T y. \tag{37}$$

Hence, for n large enough,

$$(\overline{x}^{\epsilon_n})_1 < y_1 \tag{38}$$
$$(\overline{x}^{\epsilon_n})_d > y_d \tag{39}$$

We assumed that $\mathfrak{m}(y) = d$, then by construction of $(\overline{x}^{\epsilon_n})$, we have

$$x_1 + \frac{\epsilon_n}{2} < y_1 \tag{40}$$
$$x_d - \frac{\epsilon_n}{2} < y_d. \tag{41}$$

Hence, as n goes to infinity, we have $\epsilon_{\max}(x,y)^T \leqslant 2\min\{y_1 - x_1, x_d - y_d\}$.

The left inequality is trivial whenever x is not an interior point of $S(y)$, $\epsilon_{\max}(x,y) = 0$ in this case. If x is an interior point of $S(y)$, then $\overline{x}^{\epsilon_{\max(x,y)}} \preccurlyeq y$. Hence $\overline{x}^{\epsilon_{\max(x,y)}} \preccurlyeq_T y$ and by definition

$$\epsilon_{\max(x,y)} \leqslant \epsilon^T_{\max(x,y)}. \tag{42}$$

∎

Corollary 3. *Let $x, y \in \Delta_d$ such that $x \not\preccurlyeq y$ and $x \preccurlyeq_T y$, then for $\epsilon \geqslant \epsilon_{\min}(x,y)$*

$$\epsilon + 2 \min_{1 \leqslant k < \mathfrak{m}(\overline{y}^\epsilon)} \left\{ \sum_{i=1}^{k} y_k - x_k \right\} \leqslant \epsilon_{\max}\left(x, \overline{y}_\epsilon\right)^T \tag{43}$$

Proof: Denote $\epsilon = \epsilon_{\min}(x,y)$. By definition, $x \preccurlyeq \overline{y}_\epsilon$. Hence, we can use the previous Proposition :

$$\epsilon_{\max}\left(x, \overline{y}_\epsilon\right) \leqslant \epsilon_{\max}\left(x, \overline{y}_\epsilon\right)^T. \tag{44}$$

$\epsilon_{\max}\left(x, \overline{y}_\epsilon\right)$ can be explicitly computed using Corollary 2

$$\epsilon_{\max}\left(x, \overline{y}_\epsilon\right) = \epsilon + 2 \min_{1 \leqslant k < \mathfrak{m}(\overline{y}_\epsilon)} \left\{ \sum_{i=1}^{k} y_i - x_i \right\}. \tag{45}$$

∎

Remark 18. *We recall that $\epsilon_{\max}\left(x, \overline{y}_\epsilon\right)^T$ denote the supremum of all ϵ' such that $\overline{x}_{\epsilon'} \preccurlyeq_T \overline{y}_\epsilon$. Hence, this Corollary gives a lower bound on the robustness of an approximate trumping majorization whenever $\epsilon = \epsilon_{\min}(x,y)$.*

Assume that Alice and Bob know that $\lambda^\rho \preccurlyeq_T \lambda^\sigma$ with $\lambda^\rho \otimes c \preccurlyeq \lambda^\sigma \otimes c$ and want to be sure that they will still verify the trumping relation with $c \in \Delta_m$. Formally, given $x, y \in \Delta_d$ and $c \in \Delta_m$ verifying $x \otimes c \preccurlyeq y \otimes c$ we want ϵ such

$$\forall x' \in \mathcal{B}_\epsilon(x),\ x' \otimes c \preccurlyeq y \otimes c. \tag{46}$$

And denote $\mathcal{T}(y,c)$ the set of vectors x in Δ_d such that $x \otimes c \preccurlyeq y \otimes c$.

Proposition 17. *Let $x, y \in \Delta_d$, $\mathcal{B}_\epsilon(x) \subset T(y,c)$ if and only if $\overline{x}_\epsilon \otimes c \preccurlyeq y \otimes c$.*

We show that a sufficient condition to have $\mathcal{B}_\epsilon(x) \subset T(y,c)$ is to have $\mathcal{B}_\epsilon(x \otimes c) \subset S(y \otimes c)$ which is far easier to study. First, we establish a bound on the minimal ϵ such that $x \otimes c \preccurlyeq \overline{y \otimes c}_\epsilon$ given their distance.

Lemma 3. *Let $x, y \in \Delta_d$ such that $\|x - y\|_1 \leqslant \epsilon$ and $c \in \Delta_m$ then*

$$\epsilon_{\min}(x \otimes c, y \otimes c) \leqslant \epsilon \tag{47}$$

Proof: Consider $k < md$ such that

$$\sum_{i=1}^{k} (x \otimes c)_i - (y \otimes c)_i \geqslant 0. \tag{48}$$

We can write the following sum in the following way

$$\sum_{i=1}^{k} x_i = \sum_{i=1}^{d} x_i \sum_{j=1}^{r_i} c_j. \tag{49}$$

In other words, we consider the k greatest components of $(y \otimes c)$ and order them by x_i and r_i is the number of $x_i c_j$ in these k greatest components. We have

$$\sum_{i=1}^{k} (y \otimes c)_i \geqslant \sum_{i=1}^{d} y_i \sum_{j=1}^{r_i} c_j \tag{50}$$

because, by definition, the sum on the left is the sum of the k greatest components of $(y \otimes c)$. Hence

$$\left| \sum_{i=1}^{k} (x \otimes c)_i - (y \otimes c)_i \right| \leqslant \left| \sum_{i=1}^{d} (x_i - y_i) \sum_{j=1}^{r_i} c_j \right| \quad (51)$$

$$\leqslant \sum_{i=1}^{d} |x_i - y_i| \sum_{j=1}^{r_i} c_j \quad (52)$$

$$\leqslant \sum_{i=1}^{d} |x_i - y_i| \quad (53)$$

$$= \|x - y\|_1 \quad (54)$$

$$\leqslant \epsilon. \quad (55)$$

Therefore, for any $k \in \{1, ..., md\}$, we have

$$\sum_{i=1}^{k} (x \otimes c)_i - (y \otimes c)_i \leqslant \epsilon. \quad (56)$$

Thus, $\epsilon_{\min} (y \otimes c, x \otimes c) \leqslant \epsilon$. ■

Proposition 18. *Let $x \in \Delta_d$, $c \in \Delta_m$ and $\epsilon \geqslant 0$, then*

$$\overline{x}_\epsilon \otimes c \preccurlyeq \overline{x \otimes c}_\epsilon \quad (57)$$

$$\underline{x \otimes c}_\epsilon \preccurlyeq \underline{x}_\epsilon \otimes c. \quad (58)$$

Proof: $\|x - \overline{x}_\epsilon\|_1 \leqslant \epsilon$, hence $\epsilon_{\min} (\overline{x}_\epsilon \otimes c, x \otimes c) \leqslant \epsilon$. Then

$$\overline{x}_\epsilon \otimes c \preccurlyeq \overline{x \otimes c}_\epsilon. \quad (59)$$

Similarly using Proposition 5, $\|x - \underline{x}_\epsilon\|_1 \leqslant \epsilon$ then $\epsilon_{\min} (x \otimes c, \underline{x}_\epsilon \otimes c) \leqslant \epsilon$. Then

$$\underline{x \otimes c}_\epsilon \preccurlyeq \underline{x}_\epsilon \otimes c. \quad (60)$$

■

Remark 19. *Hence, if $\overline{x \otimes c}_\epsilon \preccurlyeq y \otimes c$, i.e. $\mathcal{B}_\epsilon(x \otimes c) \subset S(y \otimes c)$, then*

$$\overline{x}_\epsilon \otimes c \preccurlyeq \overline{x \otimes c}_\epsilon \preccurlyeq y \otimes c. \quad (61)$$

Thus, $\mathcal{B}_\epsilon(x) \subset \mathcal{T}(y, c)$.

Similarly, we can introduce $\epsilon_{\max}(x, y, c)^T$ the greatest ϵ such that for any $x' \in \mathcal{B}_\epsilon(x)$ we have $x' \otimes c \preccurlyeq y \otimes c$. As a corollary, we can establish a lower bound on this quantity.

Corollary 4. *Let $x, y \in \Delta_d$, $c \in \Delta_m$ and $\epsilon \geqslant 0$, then*

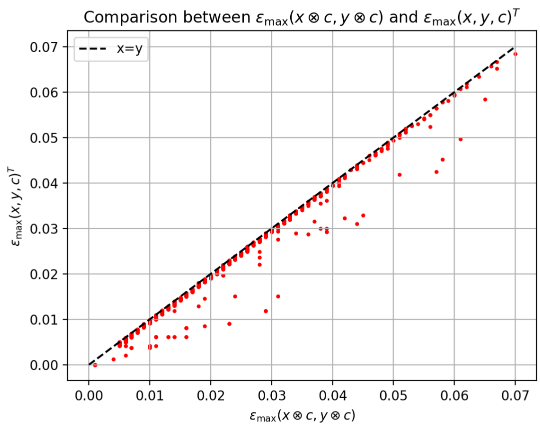
$$\epsilon_{\max} (x \otimes c, y \otimes c) \leqslant \epsilon_{\max}(x, y, c)^T. \quad (62)$$

Results on figure 10 may be pondered by some observations on inequality (62). It seems that for catalysis c close to uniform distribution, the bound is tight. However, it may be degraded whenever the catalysis gets far from the uniform.

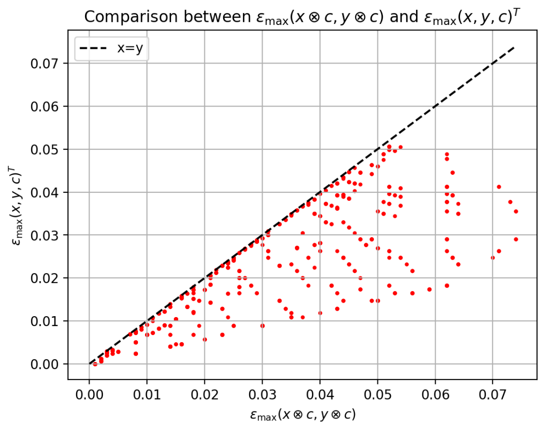
VI. Conclusion

We have considered the problem of robustness of majorization upon perturbation of vectors through the framework of the majorization lattice. We have derived for ℓ^1 metric a greatest radius for the corrupted vectors to be within in order to maintain a majorization relation. This has been used to study the robustness of a LOCC protocol in theory, but may be applied in different contexts. For example, in V-B, we have established a necessary and sufficient condition for a separability criterion to be robust. These results can be used to derive bounds on approximate and robust trumping majorization which is considered to be a complex order to study.

In [11], many results on lattice-based approximation with ℓ^∞ metric have been studied. It would be interesting to study the robustness for this metric that can be considered more "restrictive" than the ℓ^1 metric.



(a) Here $c = (0.6, 0.4)$.



(b) Here $c = (0.7, 0.3)$

Fig. 10: For a given c, we plot for couples $x, y \in \Delta_d$ such that $x \otimes c \preccurlyeq y \otimes c$ the comparison between an estimation of $\epsilon_{\max}(x, y, c)^T$ and $\epsilon_{\max}(x \otimes c, y \otimes c)$.

VII. Acknowledgements

We would like to thank Olivier Rioul and Julien Béguinot for suggesting this topic and helping us through this rich theory of majorization.

VIII. Appendix

A. Quantum Formalism

While classical information theory relies on random variables and probability distributions, it is well-known that quantum theory cannot be described entirely by this formalism. Thus, accordingly to the first postulate of quantum mechanics quantum states are described by unit vectors of an Hilbert space. In quantum information theory however, states are mainly described by statistical mixing of these quantum mechanics states. That is, considering $|\psi_1\rangle, ..., |\psi_n\rangle$, the coupling $(|\psi_1\rangle, p_1), ..., (|\psi_n\rangle, p_n)$ where p is a probability distribution. These can be entirely described by density operators.

Definition 11 (Density Operators)**.** *Let $\mathcal{H} = \mathbb{C}^d$ be an Hilbert space. A density operator ρ on $\mathcal{H}$ is a positive semi-definite matrix that verifies Tr $(\rho) = 1$.*

Denote $\mathcal{D}(\mathcal{H})$ the set of density operators on $\mathcal{H}$.

We can explicitely define a quantum state in information theory, using density operators.

Definition 12. *Let $\mathcal{H}_A$ be an Hilbert space associated with a quantum system A, a quantum state, or state, is a density operator ρ_A on $\mathcal{H}_A$.*

The following definitions and properties enlighten how density operators are used in quantum information theory.

Definition 13 (Pure / Mixed States). *Let $\rho \in \mathcal{D}(\mathcal{H})$ a quantum state. ρ is said to be pure if its rank is equal to one. Otherwise, it is said to be a mixed state.*

The notion of purity can be interpreted as a notion of certainty. There might be randomness on measurement, according to Born's rule, however if the state is pure, we are sure of which state the quantum system is in.

Proposition 19. *Let $\mathcal{H}$ be an Hilbert space, a quantum state $\rho \in \mathcal{D}(\mathcal{H})$ is pure if and only if there exists $|\psi\rangle \in \mathcal{H}$ such that*

$$\rho = |\psi\rangle\langle\psi| \tag{63}$$

Corollary 5. *Let $\mathcal{H}$ be an Hilbert space, let $\rho \in \mathcal{D}(\mathcal{H})$ a mixed quantum state, then there exists $p_1, ..., p_r$ a probability distribution and $\rho_1, ..., \rho_r$ pure quantum states in $\mathcal{D}(\mathcal{H})$ such that*

$$\rho = \sum_{i=1}^{r} p_i \rho_i \tag{64}$$

Then ρ can also be expressed in a vector notation by spectral theorem

$$\rho = \sum_{i=1}^{r} p_i |\psi_i\rangle\langle\psi_i| \tag{65}$$

with $|\psi_1\rangle, ..., |\psi_r\rangle$ orthogonal states.

Proposition 20. *Let $\mathcal{H}$ be an Hilbert space, let $\rho \in \mathcal{D}(\mathcal{H})$, then*

$$Tr\left(\rho^2\right) \leqslant 1 \tag{66}$$

with equality if and only if ρ is pure.

In quantum information, and generally in information theory, it is common to consider multipartite system. For example, the coupling of two random variables (X, Y) will usually be studied for bipartite communication through channel. Here, we define such multipartite system for quantum information.

Definition 14 (Hilbert Space of multipartite system). *If we consider $A_1, ..., A_n$ quantum systems with associated Hilbert spaces $\mathcal{H}_1, ..., \mathcal{H}_n$, the Hilbert space associated to the multipartite system $A_1...A_n$ is $\mathcal{H}_1 \otimes ... \otimes \mathcal{H}_n$, where $\otimes$ denotes the tensor operation.*

It is relatively easy to see that if $\rho_1 \in \mathcal{D}(\mathcal{H}_1)$ and $\rho_2 \in \mathcal{D}(\mathcal{H}_2)$, then $\rho_1 \otimes \rho_2 \in \mathcal{D}(\mathcal{H}_1 \otimes \mathcal{H}_2)$. However, the converse is not generally true : if $\rho \in \mathcal{D}(\mathcal{H}_1 \otimes \mathcal{H}_2)$, ρ may be not equal to $\rho_1 \otimes \rho_2$ for any $\rho_1 \in \mathcal{D}(\mathcal{H}_1)$ and $\rho_2 \in \mathcal{D}(\mathcal{H}_2)$. We introduce definitions for multipartite states.

Definition 15 (Product State). *Let $\mathcal{H}_1, ..., \mathcal{H}_n$ be Hilbert spaces, let $\rho \in \mathcal{D}(\mathcal{H}_1 \otimes ... \otimes \mathcal{H}_n)$ be a density operator on product space. ρ is said to be a product space if*

$$\exists \rho_1 \in \mathcal{H}_1, ..., \exists \rho_n \in \mathcal{H}_n,\ \rho = \rho_1 \otimes ... \otimes \rho_n \tag{67}$$

Definition 16 (Separable / Entangled States). *Let $\rho \in \mathcal{D}(\mathcal{H}_1 \otimes ... \otimes \mathcal{H}_n)$, ρ is said to be separable if it is convex combination of product states. Otherwise, the state is said to be entangled.*

Consider that AB is a bipartite state of Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B$. We would like to have a linear map such that given a state of $\mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$, we get the state of system A. This motivates the definition of partial trace

Definition 17 (Partial Trace). *Let $\mathcal{H}_A$,$\mathcal{H}_B$ be Hilbert spaces, the partial trace over $\mathcal{H}_B$ Tr_B is defined as the unique linear map such that*

$$Tr_B\left(|a_1\rangle\langle a_2| \otimes |b_1\rangle\langle b_2|\right) = |a_1\rangle\langle a_2|\, Tr\left(|b_1\rangle\langle b_2|\right) \tag{68}$$

$\forall\, |a_1\rangle, |a_2\rangle \in \mathcal{H}_A, \forall\, |b_1\rangle, |b_2\rangle \in \mathcal{H}_B$

This can be seen as the operation of marginalization for classical random variables, from (X, Y) we get X. Here, from ρ^{AB}, we get ρ^A the reduced operator on A.

Definition 18 (Reduced operator). *Let $\mathcal{H}_A, \mathcal{H}_B$ be two Hilbert spaces, let $\rho^{AB} \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$, we define the reduced operator ρ^A on A*

$$\rho^A = Tr_B\left(\rho^{AB}\right) \tag{69}$$

In quantum information theory, we can define channel as superoperators mapping density operators. To be well-defined and for operational purpose such maps must verify some axioms.

Definition 19. *Let $\mathcal{H}$ and $\mathcal{H}'$ be two Hilbert spaces. A map Φ is said to be a channel if it verifies*

1) *Convex-linearity : for any $(p_1, ..., p_n)$ and $\rho_1, ..., \rho_n$ such that $\|p\|_1 = 1$, we have*

$$\Phi\left(\sum_{i=1}^{n} p_i \rho_i\right) = \sum_{i=1}^{n} p_i \Phi(\rho_i). \tag{70}$$

2) *Complete Positivity : for any Hilbert space $\mathcal{H}_R$, for any ρ_R in $\mathcal{D}(\mathcal{H}_R)$ and for any $\rho \in \mathcal{D}(\mathcal{H})$, $\rho_R \otimes \Phi(\rho)$ is semi-definite positive.*

Remark 20. *One would consider that only positivity, i.e. $\Phi(\rho) \geqslant 0$ for any ρ, should be verified. However, tensorization is a key operation in quantum information. In particular, if we consider bipartite states $\rho^A \otimes \rho^B$ and if we want to act only on ρ^B, our operation may lead to a matrix that cannot be considered as a state. A well-known example of positive but not completely positive map is the transposition (see [14] Box 8.2).*

An extension of bistochastic matrix for superoperators are the unital channels. These can be used to defined an analogous majorization order on density operators.

Definition 20. *A channel Φ is said to be unital if $\Phi\left(\frac{1}{d} I_d\right) = \frac{1}{d} I_d$.*

Remark 21. *We recall that a matrix P is said to be bistochastic whenever all its coefficients are non-negative, $u^T P = u^T$ and $Pu = u$. In other words, P as an operator on Δ_d admits as one of its fixed point u. Similarly, Φ admits as one of its fixed point the uniform density matrix.*

Theorem 5. *Let $\rho, \sigma \in \mathcal{D}(\mathcal{H})$, $\rho \preccurlyeq \sigma$ if and only if there exists a unital channel Φ such that $\rho = \Phi(\sigma)$.*

Proof: A proof of this theorem can be found in [1]. ■

B. Computation of Steepest and Flattest Approximation

Let $\epsilon > 0$ and $x \in \Delta_d$. We present here the construction $\overline{x}_\epsilon$. If $\|x - \delta\|_1 \leqslant \epsilon$, then $\overline{x}_\epsilon = \delta$. Otherwise, we can define an integer $l \in \{1, ..., d\}$ such that

$$\sum_{i=1}^{l-1} x_i + \frac{\epsilon}{2} \leqslant 1 \tag{71}$$

and

$$\sum_{i=1}^{l} x_i + \frac{\epsilon}{2} > 1. \tag{72}$$

$\overline{x}_\epsilon$ is then defined by

$$\left(\overline{x}_\epsilon\right)_k = \begin{cases} x_1 + \frac{\epsilon}{2} & \text{if } k = 1 \\ x_k & \text{if } 1 < k < l \\ 1 - \sum_{i=1}^{l-1} \left(\overline{x}_\epsilon\right)_i & \text{if } k = l \\ 0 & \text{otherwise.} \end{cases} \tag{73}$$

We then present the construction of $\underline{x}_\epsilon$. If $\|x - u\|_1 \leqslant \epsilon$, then $\underline{x}_\epsilon = u$. Otherwise, consider the following functions for $p, q \in [0, 1]$.

$$\gamma_1(x, p) = \sum_{i=1}^{d} (x_i - p)^+ \tag{74}$$

$$\gamma_2(x, q) = \sum_{i=1}^{d} (q - x_i)^+ . \tag{75}$$

We then compute p^* and q^* such that $\gamma_1(x, p^*) = \frac{\epsilon}{2}$ and $\gamma_2(x, q^*) = \frac{\epsilon}{2}$. These can be explicitly computed. Consider l to be the integer such that

$$\gamma_1(x, x_l) \leqslant \frac{\epsilon}{2} < \gamma_1(x, x_{l+1}) \tag{76}$$

then, $x_{l+1} < p^* \leqslant x_l$ (γ_1 is decreasing in p). We can then simplify the sum

$$\gamma_1(x, p^*) = \sum_{i=1}^{l} x_i - p^* \tag{77}$$

$$= \sum_{i=1}^{l} x_i - lp^* \tag{78}$$

$$= \frac{\epsilon}{2}. \tag{79}$$

Hence,

$$p^* = \frac{1}{l}\left(\sum_{i=1}^{l} x_i - \frac{\epsilon}{2}\right). \tag{80}$$

Similarly, denote r the integer such that

$$\gamma_2(x, x_{r+1}) \leqslant \frac{\epsilon}{2} < \gamma_2(x, x_r). \tag{81}$$

Then, $x_r \leqslant q^* < x_{r-1}$ because γ_2 is increasing in q. As a consequence, we can simplify γ_2

$$\gamma_2(x, q^*) = \sum_{i=r}^{d} q^* - x_i \tag{82}$$

$$= (d - r + 1)q^* - \sum_{i=r}^{d} x_i. \tag{83}$$

Hence, we have

$$q^* = \frac{1}{d - r + 1}\left(\sum_{i=r}^{d} x_i + \frac{\epsilon}{2}\right). \tag{84}$$

The flattest is then explicitly defined using p^* and q^*.

$$(\underline{x}_\epsilon)_k = \begin{cases} p^* & \text{if } k \leqslant l \\ x_k & \text{if } l < k < r \\ q^* & \text{otherwise.} \end{cases} \tag{85}$$

C. Proofs

Proof of Proposition 2.

Proof: We are going to show successively that $\overline{\left(\overline{x}_{\epsilon_1}\right)}_{\epsilon_2} \in \mathcal{B}_{\epsilon_1+\epsilon_2}(x)$ and $\overline{x}_{\epsilon_1+\epsilon_2} \in \mathcal{B}_{\epsilon_2}\left(\overline{x}_{\epsilon_1}\right)$.

The first one comes directly from triangle inequality

$$\|x - \overline{\left(\overline{x}_{\epsilon_1}\right)}_{\epsilon_2}\|_1 = \|x - \overline{x}_{\epsilon_1} + \overline{x}_{\epsilon_1} - \overline{\left(\overline{x}_{\epsilon_1}\right)}_{\epsilon_2}\|_1 \tag{86}$$

$$\leqslant \epsilon_1 + \epsilon_2 \tag{87}$$

Hence, $\overline{\left(\overline{x}_{\epsilon_1}\right)}_{\epsilon_2} \in \mathcal{B}_{\epsilon_1+\epsilon_2}(x)$, thus $\overline{\left(\overline{x}_{\epsilon_1}\right)}_{\epsilon_2} \preccurlyeq \overline{x}_{\epsilon_1+\epsilon_2}$. The second one is quite handy. Denote respectively $m = \mathfrak{m}\left(\overline{x}_{\epsilon_1}\right)$ and $l = \mathfrak{m}\left(\overline{x}_{\epsilon_1+\epsilon_2}\right)$. We also write μ^ϵ for $\overline{x}_\epsilon$.

$$\begin{aligned} \sum_{i=1}^{d} \left|\mu_i^{\epsilon_1} - \mu_i^{\epsilon_1+\epsilon_2}\right| &= \left|\mu_1^{\epsilon_1} - \mu_1^{\epsilon_1+\epsilon_2}\right| + \sum_{i=2}^{d} \left|\mu_i^{\epsilon_1} - \mu_i^{\epsilon_1+\epsilon_2}\right| \\ &= \frac{\epsilon_2}{2} + \sum_{i=l}^{d} \left|\mu_i^{\epsilon_1} - \mu_i^{\epsilon_1+\epsilon_2}\right| \end{aligned}$$

If $m = l$, then

$$\begin{aligned} &\frac{\epsilon_2}{2} + \sum_{i=l}^{d} \left|\mu_i^{\epsilon_1} - \mu_i^{\epsilon_1+\epsilon_2}\right| \\ &= \frac{\epsilon_2}{2} + \left|1 - \sum_{i=1}^{l-1} \mu_i^{\epsilon_1} - \left(1 - \sum_{i=1}^{l-1} \mu_i^{\epsilon_1+\epsilon_2}\right)\right| \\ &= \frac{\epsilon_2}{2} + \frac{\epsilon_2}{2}. \end{aligned}$$

If $l < m$ (by construction $l \leqslant m$), then we have

$$\begin{aligned} &\frac{\epsilon_2}{2} + \sum_{i=l}^{d} \left|\mu_i^{\epsilon_1} - \mu_i^{\epsilon_1+\epsilon_2}\right| \\ &= \frac{\epsilon_2}{2} + \left|\mu_l^{\epsilon_1} - \mu_l^{\epsilon_1+\epsilon_2}\right| + \left|\mu_m^{\epsilon_1} - \mu_m^{\epsilon_1+\epsilon_2}\right| + \sum_{i=l}^{m-1} x_i \\ &= \frac{\epsilon_2}{2} + \left|1 - \sum_{i=1}^{l-1} \mu_i^{\epsilon_1} - x_l\right| + \left|1 - \sum_{i=1}^{m-1} \mu_i^{\epsilon_1+\epsilon_2}\right| + \sum_{i=l}^{m-1} x_i \\ &= \frac{\epsilon_2}{2} + x_l + \sum_{i=1}^{l-1} \mu_i^{\epsilon_1} - \sum_{i=1}^{m-1} \mu_i^{\epsilon_1+\epsilon_2} + \sum_{i=l}^{m-1} x_i \\ &= \frac{\epsilon_2}{2} + \frac{\epsilon_2 + \epsilon_1}{2} - \frac{\epsilon_1}{2} - \sum_{i=l}^{m-1} x_i + \sum_{i=l}^{m-1} x_i \\ &= \epsilon_2. \end{aligned}$$

Hence $\overline{x}_{\epsilon_1+\epsilon_2} \in \mathcal{B}_{\epsilon_2}\left(\overline{x}_{\epsilon_1}\right)$, thus $\overline{x}_{\epsilon_1+\epsilon_2} \preccurlyeq \overline{\overline{x}_{\epsilon_1}}_{\epsilon_2}$. Therefore, we have equality between the two vectors. ■

Proof of Proposition 3.

Proof: See [9] to have more details on the construction of $\underline{x}_\epsilon$. We can show that the construction of $\underline{x}_\epsilon$ depends on the construction of $p, q \in [0, 1]$ such that

$$\gamma_1(p, x) = \sum_{i=1}^{d} (x_i - p)^+ = \frac{\epsilon}{2}$$

$$\gamma_2(q, x) = \sum_{i=1}^{d} (q - x_i)^+ = \frac{\epsilon}{2}.$$

When p, q are found, $\left(\underline{x}_\epsilon\right)_i = p$ if $x_i \geqslant p$, $\left(\underline{x}_\epsilon\right)_i = q$ if $x_i \leqslant q$ and otherwise $\left(\underline{x}_\epsilon\right)_i = x_i$. In particular, we denote l_I^ϵ the greatest integer such that $x_{l_I^\epsilon} \geqslant p$.

Consider p, q defined for $\underline{x}_{\epsilon_1}$, p', q' for $\underline{\left(\underline{x}^{(1)}_{\epsilon_1}\right)^{(1)}}_{\epsilon_2}$ and p'', q'' for $\underline{x}_{\epsilon_1+\epsilon_2}$. We have $\gamma_1(x, p) + \gamma_1\left(\underline{x}_{\epsilon_1}, p'\right) = \frac{\epsilon_1+\epsilon_2}{2}$, then

$$\begin{aligned}
\frac{\epsilon_1 + \epsilon_2}{2} &= \sum_{i=1}^{d} (x_i - p)^+ + \left(\underline{x}_{\epsilon_1} - p'\right)^+ \\
&= \sum_{i=1}^{l_I^{\epsilon_1}} x_i - p + p - p' + \sum_{i=l_I^{\epsilon_1}+1}^{d} \left(\underline{x}_{\epsilon_1} - p'\right)^+ \\
&= \sum_{i=1}^{d} \left(x - p'\right)^+
\end{aligned}$$

By unicity of p'', we know that $p' = p''$. Similarly, we can show that $q' = q''$

$$\begin{aligned}
\gamma_2(x, p) + \gamma_2\left(\underline{x}_{\epsilon_1}, p'\right) &= \sum_{i=1}^{d} (q - x_i)^+ + \left(q' - \underline{x}_{\epsilon_1}\right)^+ \\
&= \sum_{i=r^{\epsilon_2}}^{d} q - x_i + \left(q' - q\right)^+ \\
&+ \sum_{i=1}^{r^{\epsilon_2}-1} \left(q' - \underline{x}_{\epsilon_1}\right)^+ \\
&= \sum_{i=r^{\epsilon_2}}^{d} q' - x_i + \sum_{i=1}^{r^{\epsilon_2}-1} \left(q' - x\right)^+ \\
&= \sum_{i=1}^{d} \left(q' - x\right)^+ \\
&= \frac{\epsilon_1 + \epsilon_2}{2}.
\end{aligned}$$

Hence, $q' = q''$.

Since $p \geqslant p' = p''$ and $q \leqslant q' = q''$. It follows that both vectors are equal since

$$\left\{i \leqslant d\,;\, x_i \geqslant p''\right\} = \left\{i \leqslant d\,;\, \left(\underline{x}_{\epsilon_1}\right)_i \geqslant p''\right\}. \tag{88}$$

and

$$\left\{j \leqslant d\,;\, x_j \leqslant p''\right\} = \left\{j \leqslant d\,;\, \left(\underline{x}_{\epsilon_1}\right)_j \leqslant p''\right\}. \tag{89}$$

■

Proof of Lemma 1

Proof: Denote μ^ϵ the steepest $\epsilon-$approximation of x. We assume there exists ϵ is such that $\mu^\epsilon \preccurlyeq y$, i.e. x is an interior point of $S(y)$. We can consider ϵ as small as possible, such that μ^ϵ also has only non-null coordinates. Then, μ^ϵ is defined as $\mu_1^\epsilon = x_1 + \frac{\epsilon}{2}$, $\mu^\epsilon_{\mathfrak{m}(x)} = x_{\mathfrak{m}(x)} - \frac{\epsilon}{2}$ and $\mu_i^\epsilon = x_i$ otherwise. The following relation if verified

$$x \preccurlyeq \mu^\epsilon \preccurlyeq y \quad \text{and} \quad x \preccurlyeq y \tag{90}$$

Assume there exists $k < \mathfrak{m}(y) - 1 < \mathfrak{m}(x)$ such that $\sum_{i=1}^k x_i = \sum_{i=1}^k y_i$. By the above majorization relation, we get

$$\sum_{i=1}^{k} x_i = \sum_{i=1}^{k} \mu_i^\epsilon = \sum_{i=1}^{k} y_i. \tag{91}$$

However, $\sum_{i=1}^k \mu_i^\epsilon = \sum_{i=1}^k x_i + \frac{\epsilon}{2}$. Whence, we have $\epsilon = 0$, which contradicts our initial hypothesis.

Conversely, assume that $\sum_i^k x_i < \sum_i^k y_i$ for all $k < \mathfrak{m}(y)$. Then, denote $\epsilon \leqslant \min_{1 \leqslant k < \mathfrak{m}(y)}\{\sum_i^k y_i - x_i\}$ such that $x_{\mathfrak{m}(x)} - \frac{\epsilon}{2} > 0$, then μ^ϵ has $\mathfrak{m}(x)$ non-null coefficients and

$$\sum_{i=1}^{k} \mu_i^\epsilon = \sum_{i=1}^{k} x_i + \frac{\epsilon}{2} \leqslant \sum_{i=1}^{k} x_i + \sum_{i=1}^{k} y_i - x_i = \sum_{i=1}^{k} y_i. \tag{92}$$

Hence $\mu^\epsilon \preccurlyeq y$, i.e. x is an interior point of $S(y)$. ■

References

[1] A.Uhlmann. “Sätze über Dichtematrizen”. In: *Wiss. Z. Karl-Marx-Univ. Leipzig, Math.-Nat. R.* 20 (1971), pp. 633–637.

[2] R.B. Bapat. “Majorization and singular values. III”. In: *Linear Algebra and its Applications* 145 (1991), pp. 59–70. ISSN: 0024-3795. DOI: https://doi.org/10.1016/0024-3795(91)90287-7. URL: https://www.sciencedirect.com/science/article/pii/0024379591902877.

[3] Charles H. Bennett et al. “Teleporting an unknown quantum state via dual classical and Einstein-Podolsky-Rosen channels”. In: *Phys. Rev. Lett.* 70 (13 Mar. 1993), pp. 1895–1899. DOI: 10.1103/PhysRevLett.70.1895. URL: https://link.aps.org/doi/10.1103/PhysRevLett.70.1895.

[4] R. Bhatia. *Matrix analysis*. Graduate texts in mathematics. Springer, 1997. ISBN: 9783540948469. URL: https://books.google.fr/books?id=f0ioPwAACAAJ.

[5] G M Bosyk et al. “Optimal common resource in majorization-based resource theories”. In: *New Journal of Physics* 21.8 (Aug. 2019), p. 083028. DOI: 10.1088/1367-2630/ab3734. URL: https://dx.doi.org/10.1088/1367-2630/ab3734.

[6] Eric Chitambar et al. “Everything You Always Wanted to Know About LOCC (But Were Afraid to Ask)”. In: *Communications in Mathematical Physics* 328.1 (Mar. 2014), pp. 303–326. ISSN: 1432-0916. DOI: 10.1007/s00220-014-1953-9. URL: http://dx.doi.org/10.1007/s00220-014-1953-9.

[7] Sumit Daftuar and Matthew Klimesh. “Mathematical structure of entanglement catalysis”. In: *Phys. Rev. A* 64 (4 Sept. 2001), p. 042314. DOI: 10.1103/PhysRevA.64.042314. URL: https://link.aps.org/doi/10.1103/PhysRevA.64.042314.

[8] Sevag Gharibian. *Strong NP-Hardness of the Quantum Separability Problem*. 2009. arXiv: 0810.4507 [quant-ph].

[9] Michał Horodecki, Jonathan Oppenheim, and Carlo Sparaciari. “Extremal distributions under approximate majorization”. In: *Journal of Physics A: Mathematical and Theoretical* 51.30 (June 2018), p. 305301. ISSN: 1751-8121. DOI: 10.1088/1751-8121/aac87c. URL: http://dx.doi.org/10.1088/1751-8121/aac87c.

[10] Albert W. Marshall, Ingram Olkin, and Barry C. Arnold. *Inequalities: Theory of Majorization and its Applications*. Vol. 143. 2011.

[11] C Massri et al. “Extremal elements of a sublattice of the majorization lattice and approximate majorization”. In: *Journal of Physics A: Mathematical and Theoretical* 53.21 (May 2020), p. 215305.

[12] M. A. Nielsen. "Conditions for a Class of Entanglement Transformations". In: *Physical Review Letters* 83.2 (July 1999), pp. 436–439. ISSN: 1079-7114. DOI: 10.1103/physrevlett.83.436. URL: http://dx.doi.org/10.1103/PhysRevLett.83.436.

[13] M. A. Nielsen and J. Kempe. "Separable States Are More Disordered Globally than Locally". In: *Physical Review Letters* 86.22 (May 2001), pp. 5184–5187. ISSN: 1079-7114. DOI: 10.1103/physrevlett.86.5184. URL: http://dx.doi.org/10.1103/PhysRevLett.86.5184.

[14] Michael A. Nielsen and Isaac L. Chuang. *Quantum Computation and Quantum Information*. Cambridge University Press, 2010.

[15] Guifré Vidal. "Entanglement of Pure States for a Single Copy". In: *Physical Review Letters* 83.5 (Aug. 1999), pp. 1046–1049. ISSN: 1079-7114. DOI: 10.1103/physrevlett.83.1046. URL: http://dx.doi.org/10.1103/PhysRevLett.83.1046.

[16] Hayata Yamasaki et al. "Multipartite entanglement outperforming bipartite entanglement under limited quantum system sizes". In: *Phys. Rev. A* 98 (5 Nov. 2018), p. 052313. DOI: 10.1103/PhysRevA.98.052313. URL: https://link.aps.org/doi/10.1103/PhysRevA.98.052313.