

Lecture Notes in Physics 902

Ved Prakash Gupta  
Prabha Mandayam  
V.S. Sunder

# The Functional Analysis of Quantum Information Theory

A Collection of Notes Based on  
Lectures by Gilles Pisier, K. R.  
Parthasarathy, Vern Paulsen and  
Andreas Winter

 Springer

# Lecture Notes in Physics

Volume 902

## *Founding Editors*

W. Beiglböck  
J. Ehlers  
K. Hepp  
H. Weidenmüller

## *Editorial Board*

M. Bartelmann, Heidelberg, Germany  
B.-G. Englert, Singapore, Singapore  
P. Hänggi, Augsburg, Germany  
M. Hjorth-Jensen, Oslo, Norway  
R.A.L. Jones, Sheffield, United Kingdom  
M. Lewenstein, Barcelona, Spain  
H. von Löhneysen, Karlsruhe, Germany  
J.-M. Raimond, Paris, France  
A. Rubio, Donostia, San Sebastian, Spain  
S. Theisen, Potsdam, Germany  
D. Vollhardt, Augsburg, Germany  
J.D. Wells, Ann Arbor, USA  
G.P. Zank, Huntsville, USA

# The Lecture Notes in Physics

The series Lecture Notes in Physics (LNP), founded in 1969, reports new developments in physics research and teaching—quickly and informally, but with a high quality and the explicit aim to summarize and communicate current knowledge in an accessible way. Books published in this series are conceived as bridging material between advanced graduate textbooks and the forefront of research and to serve three purposes:

- to be a compact and modern up-to-date source of reference on a well-defined topic
- to serve as an accessible introduction to the field to postgraduate students and nonspecialist researchers from related areas
- to be a source of advanced teaching material for specialized seminars, courses and schools

Both monographs and multi-author volumes will be considered for publication. Edited volumes should, however, consist of a very limited number of contributions only. Proceedings will not be considered for LNP.

Volumes published in LNP are disseminated both in print and in electronic formats, the electronic archive being available at [springerlink.com](http://springerlink.com). The series content is indexed, abstracted and referenced by many abstracting and information services, bibliographic networks, subscription agencies, library networks, and consortia.

Proposals should be sent to a member of the Editorial Board, or directly to the managing editor at Springer:

Christian Caron  
Springer Heidelberg  
Physics Editorial Department I  
Tiergartenstrasse 17  
69121 Heidelberg/Germany  
[christian.caron@springer.com](mailto:christian.caron@springer.com)

More information about this series at  
<http://www.springer.com/series/5304>

Ved Prakash Gupta • Prabha Mandayam •  
V.S. Sunder

# The Functional Analysis of Quantum Information Theory

A Collection of Notes Based on Lectures  
by Gilles Pisier, K.R. Parthasarathy,  
Vern Paulsen and Andreas Winter

 Springer

Ved Prakash Gupta  
School of Physical Sciences  
Jawaharlal Nehru University  
New Delhi, India

Prabha Mandayam  
Department of Physics  
Indian Institute of Technology Madras  
Chennai, India

V.S. Sunder  
The Institute of Mathematical Sciences  
Chennai, India

ISSN 0075-8450

Lecture Notes in Physics

ISBN 978-3-319-16717-6

DOI 10.1007/978-3-319-16718-3

ISSN 1616-6361 (electronic)

ISBN 978-3-319-16718-3 (eBook)

Library of Congress Control Number: 2015940545

Springer Cham Heidelberg New York Dordrecht London

© Springer International Publishing Switzerland 2015

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made.

Printed on acid-free paper

Springer International Publishing AG Switzerland is part of Springer Science+Business Media  
([www.springer.com](http://www.springer.com))

# Preface

These notes grew out of a 2-week workshop on *The Functional Analysis of Quantum Information Theory* that was held at the Institute of Mathematical Sciences, Chennai during 26/12/2011 to 06/01/2012. This was initially the brain-child of Ed Effros, who was to have been one of the four principal speakers; but as things unfolded, he had to pull out at the eleventh hour due to unforeseen and unavoidable circumstances. But he mentored us through the teething stages with suggestions on possible and desirable substitutes. After a few hiccups, we arrived at a perfectly respectable cast of characters, largely owing to Prof. K.R. Parthasarathy agreeing to fill the breach to the extent that his somewhat frail health would permit. While everybody else had been asked to give five 90-min lectures each, he agreed gamely to give three 60-min lectures, which were each as delightful as one has expected his lectures to be.

The three other speakers were Gilles Pisier, Vern Paulsen, and Andreas Winter. Given the impeccable clarity in their lectures, it was not surprising that the workshop had a substantial audience (a mixed bag of mathematicians, physicists, and theoretical computer scientists) for the entire 2 weeks, and several people wanted to know if the proceedings of the workshop could be published. Given the wide scope of the problems discussed here, we hope these notes will prove useful to students and research scholars across all three disciplines.

The quite non-trivial and ambitious task of trying to put together a readable account of the lectures was taken upon by the three of us, with periodic assistance from Madhushree Basu and Issan Patri. When we finally finished a first draft some 28 months later and solicited permission from the speakers for us to publicize this account, they responded quite promptly and positively, with Vern even offering to slightly edit and expand “his” part and Gilles uncomplainingly agreeing to edit the last few pages of “his” part.

As even a casual reading of these notes will show, different parts have been written by different people, and some material has been repeated (the Kraus decomposition for instance). Further the schism in the notational conventions of physicists and mathematicians (e.g., in which of its two arguments the inner product is (conjugate) linear) reflects which convention which speaker followed!

This schism was partly because we did not want to introduce new errors by trying to adopt a convention other than in the notes taken from the lectures. Thus, this must be viewed not as a textbook in a polished and complete form, but rather as a transcription of notes derived from lectures, reflecting the spontaneity of live classroom interactions and conveying the enthusiasm and intuition for the subject. It goes without saying that any shortcomings in these notes are to be attributed to the scribes, while anything positive should be credited to the speakers.

New Delhi, India  
Chennai, India  
Chennai, India

Ved Prakash Gupta  
Prabha Mandayam  
V.S. Sunder

# Acknowledgements

Firstly, we express our gratitude to so many people in the institute (IMSc) who helped in so many ways to make the workshop as successful as it was, not least our genial Director Balu for generously funding the whole enterprise. We also thank our systems people who cheerfully videographed all the lectures, these videos being made available as a You-tube playlist at <http://www.imsc.res.in/~sunder/faqit.html>; and the reader is strongly urged to view these videos in parallel with reading these notes, to really get the full benefit of them.

PM was a post-doctoral fellow with the quantum information group at IMSc during the writing of this manuscript. She would like to thank IMSc and the quantum information group in particular for the opportunity to participate in the workshop and be a part of this effort.

VSS also wishes to explicitly thank Dr. S. Ramasami, the former Secretary of the Department of Science and Technology, for having gladly, promptly, and unconditionally permitted the use of the generous grant (the J.C. Bose Fellowship) he receives from them in a wonderful example of the “reasonable accommodation” that the UN Convention on the Rights of Persons with Disability speaks about, which has permitted him to travel when needed and be a functional mathematician in spite of health constraints.

We will be remiss if we did not record our appreciation of the extremely pleasant and civil manner in which we were requested by the Editors of Springer LNP to permit them to publish these notes. This association/correspondence has been an altogether refreshing and heartening experience.





# Contents

<b>1 Operator Spaces</b> .....	1
1.1 Operator Spaces .....	1
1.1.1 Completely Bounded and Completely Positive Maps .....	2
1.1.2 Operator Systems .....	2
1.1.3 Fundamental Factorisation of CB Maps .....	3
1.2 More on CB and CP Maps .....	18
1.3 Ruan’s Theorem and Its Applications .....	27
1.3.1 Ruan’s Theorem .....	27
1.3.2 Some Applications and Some Basic Facts .....	29
1.3.3 <i>min</i> and <i>max</i> Operator Space Structures on a Banach Space .....	30
1.4 Tensor Products of Operator Spaces .....	31
1.4.1 Injective Tensor Product .....	31
1.4.2 Projective Tensor Product .....	32
1.4.3 General Remarks .....	33
1.4.4 A Passing Remark on the Haagerup Tensor Product .....	34
1.5 Tensor Products of $C^*$ -Algebras .....	34
1.5.1 <i>min</i> and <i>max</i> Tensor Products of $C^*$ -Algebras .....	34
1.5.2 Kirchberg’s Theorem .....	35
References .....	37
<b>2 Entanglement in Bipartite Quantum States</b> .....	39
2.1 Quantum States, Observables and Probabilities .....	39
2.2 Entanglement .....	42
2.2.1 Schmidt Decomposition .....	43
2.2.2 Unitary Bases, EPR States and Dense Coding .....	46
2.3 Schmidt Rank of Bipartite Entangled States .....	48
2.3.1 Subspaces of Minimal Schmidt Rank .....	48

2.4	Schmidt Number of Mixed States .....	52
2.4.1	Test for Schmidt Number $k$ Using $k$ -Positive Maps .....	53
2.4.2	Schmidt Number of Generalized Werner States .....	56
	References .....	62
<b>3</b>	<b>Operator Systems</b> .....	<b>63</b>
3.1	Theorems of Choi .....	63
3.1.1	Douglas Factorization .....	64
3.1.2	Choi-Kraus Representation and Choi Rank .....	64
3.2	Quantum Error Correction .....	67
3.2.1	Applications of Choi's Theorems to Error Correction .....	67
3.2.2	Shor's Code: An Example .....	70
3.3	Matrix Ordered Systems and Operator Systems .....	72
3.3.1	Duals of Matrix Ordered Spaces .....	74
3.3.2	Choi-Effros Theorem .....	75
3.4	Tensor Products of Operator Systems .....	76
3.4.1	Minimal Tensor Product of Operator Systems .....	78
3.4.2	Maximal Tensor Product of Operator Systems .....	78
3.5	Graph Operator Systems .....	84
3.5.1	Dual of a Graph Operator System .....	85
3.6	Three More Operator System Tensor Products .....	86
3.6.1	The Commuting Tensor Product $\otimes_c$ .....	86
3.6.2	The Tensor Products $\otimes_{el}$ and $\otimes_{er}$ .....	87
3.6.3	Lattice of Operator System Tensor Products .....	87
3.7	Some Characterizations of Operator System Tensor Products .....	88
3.7.1	Exact Operator Systems .....	88
3.7.2	Weak Expectation Property (WEP) .....	88
3.7.3	Operator System Local Lifting Property (OSLLP) .....	89
3.7.4	Double Commutant Expectation Property (DCEP) .....	89
3.8	Operator System Tensor Products and the Conjectures of Kirchberg and Tsirelson .....	90
3.8.1	Special Operator Sub-systems of the Free Group $C^*$ -Algebras .....	90
3.8.2	Kirchberg's Conjecture .....	90
3.8.3	Quotient of an Operator System .....	91
	References .....	92
<b>4</b>	<b>Quantum Information Theory</b> .....	<b>93</b>
4.1	Zero-Error Communication Via Quantum Channels .....	93
4.1.1	Conditions for Zero-Error Quantum Communication .....	95
4.1.2	Zero-Error Capacity and Lovasz $\vartheta$ Function .....	98
4.2	Strong Subadditivity and Its Equality Case .....	105
4.2.1	Monotonicity of Relative Entropy: Petz Theorem .....	107
4.2.2	Structure of States that Saturate Strong Subadditivity .....	109
4.2.3	An Operator Algebraic Proof of the Koashi-Imoto Theorem .....	111

- 4.3 Norms on Quantum States and Channels ..... 114
- 4.4 Matrix-Valued Random Variables ..... 119
  - 4.4.1 Matrix Tail Bounds ..... 121
  - 4.4.2 Destroying Correlations ..... 124
  - 4.4.3 State Merging ..... 129
- References ..... 133
  
- Further Reading** ..... 135
  
- Index** ..... 137

# Chapter 1

## Operator Spaces

**Abstract** Starting from the definitions of operator spaces and operator systems (the natural ambience to discuss the notions of complete boundedness and complete positivity, respectively) this chapter, which is based on lectures by Gilles Pisier, quickly gets to the heart of the matter by discussing several closely related theorems—Stinespring’s theorem on dilation of completely positive maps, Arveson’s extension theorems (for completely bounded as well as completely positive maps), related results by Haagerup and Paulsen, Nachbin’s Hahn-Banach theorem—as more or less different perspectives of the same result; we also see the power/use of the clever matrix trick relating positivity of self-adjoint  $2 \times 2$  block matrices and operator norm relations between its entries. After discussing the classical row (R) and column (C) Hilbert spaces, Schur multipliers etc., we proceed with Ruan’s axioms for an “abstract” operator space, applications of the formalism to highlight analogy with classical Banach space theory (e.g.:  $R^* \cong C$ ), min and max operator space structures, tensor products of  $C^*$ -algebras, nuclear pairs (whose tensor products have a unique “ $C^*$ -cross-norm”), and the chapter concludes with Kirchberg’s theorem on  $(C^*(\mathbb{F}_n), B(H))$  being a nuclear pair.

### 1.1 Operator Spaces

**Definition 1.1.1** A closed subspace  $E \subset B(H)$  for some Hilbert space  $H$  is called an operator space.

The requirement of “closed”-ness is imposed because we want to think of operator spaces as “quantised (or non-commutative) Banach spaces”. This assumption ensures that operator spaces are indeed Banach spaces with respect to the CB-norm (see Definition 1.1.2 and the subsequent remarks). Conversely every Banach space can be seen to admit such an embedding.

*(Reason:* If  $E$  is a Banach space, then the unit ball of  $E^*$ , equipped with the weak\*-topology is a compact Hausdorff space (by Alaoglu’s theorem), call it  $X$ ; and the Hahn-Banach theorem shows that  $E$  embeds isometrically into  $C(X)$ . Finally, we may use an isometric representation of the  $C^*$ -algebra  $C(X)$  on some Hilbert space  $H$ ; and we would finally have our isometric embedding of  $E$  into  $B(H)$ .)

Note that, for each Hilbert space  $H$ , there is a natural identification between  $M_n(B(H))$  and  $B(H^{(n)})$ , where  $H^{(n)} := \underbrace{H \oplus \cdots \oplus H}_{n\text{-copies}}$ . For each operator space  $E \subset B(H)$ , and  $n \geq 1$ , let

$$M_n(E) = \{[a_{ij}] : a_{ij} \in E, 1 \leq i, j \leq n\} \subset M_n(B(H)).$$

Then, each  $M_n(E)$  inherits a norm, given by

$$\|a\|_n := \|a\|_{B(H^{(n)})} = \sup \left\{ \left( \sum_{i=1}^n \left\| \sum_{j=1}^n a_{ij}(h_j) \right\|^2 \right)^{1/2} : h_j \in H, \sum_{j=1}^n \|h_j\|^2 \leq 1 \right\}$$

for all  $a = [a_{ij}] \in M_n(E)$ . In particular, each operator space  $E \subset B(H)$  comes equipped with a sequence of norms  $(\|\cdot\|_n \text{ on } M_n(E), n \geq 1)$ .

### 1.1.1 Completely Bounded and Completely Positive Maps

For a linear map  $u : E \rightarrow F$  between two operator spaces, for each  $n \geq 1$ , let  $u_n : M_n(E) \rightarrow M_n(F)$  be defined by  $u_n([a_{i,j}]) = [u(a_{i,j})]$  for all  $[a_{i,j}] \in M_n(E)$ .

**Definition 1.1.2** A linear map  $u : E \rightarrow F$  is called **completely bounded (CB)** if  $\|u\|_{cb} := \sup_{n \geq 1} \|u_n\| < \infty$ . Write  $CB(E, F) = \{u : E \xrightarrow{CB} F\}$  and equip it with the CB norm  $\|\cdot\|_{cb}$ .

Note that  $\|u\| = \|u\|_1 \leq \|u\|_{cb}$ ; so, any Cauchy sequence in  $CB(E, F)$  is also a Cauchy sequence in  $B(E, F)$  and hence one deduces that  $CB(E, F)$  is a Banach space (as  $B(E, F)$  is a Banach space). Complete boundedness has many properties similar to boundedness, for example:  $\|vu\|_{cb} \leq \|v\|_{cb}\|u\|_{cb}$ . We shall see later that  $CB(E, F)$  also inherits an appropriate operator space structure.

### 1.1.2 Operator Systems

**Definition 1.1.3** A (closed) subspace  $S \subset B(H)$  for some Hilbert space  $H$  is called an **operator system**, if

- $1 \in S$
- $x \in S \Rightarrow x^* \in S$ .

Note that if  $S \subset B(H)$  is an operator system, then  $S_+ := S \cap B(H)_+$  linearly spans  $S$  since if  $a \in S_h = \{x \in S : x = x^*\}$ , then  $a = \|a\|1 - (\|a\|1 - a)$ . Also, if  $S$  is an operator system, so also is  $M_n(S) (\subset B(H^{(n)}))$ .

**Definition 1.1.4** If  $S$  is an operator system, a linear mapping  $u : S \rightarrow B(K)$  is said to be:

- **positive** if  $x \in S_+ \Rightarrow u(x) \in B(K)_+$ .
- **$n$ -positive** if  $u_n : M_n(S) \rightarrow M_n(B(K))$  is positive.
- **completely positive (CP)** if it is  $n$ -positive for all  $n \in \mathbb{N}$ .
- **completely contractive** if  $\|u_n\| \leq 1 \forall n \in \mathbb{N}$

### 1.1.3 Fundamental Factorisation of CB Maps

Our goal, in this section, is to prove the following fundamental factorisation theorem:

**Theorem 1.1.5 (Fundamental Factorisation)** *Given an operator space  $E$  contained in a unital  $C^*$ -algebra  $A$ , the following conditions on a linear map  $u : E \rightarrow B(K)$  are equivalent:*

1.  $u \in CB(E, B(K))$  and  $\|u\|_{cb} \leq 1$ ;
2. there is a Hilbert space  $\hat{H}$ , a unital  $*$ -representation  $\pi : A \rightarrow B(\hat{H})$  and linear maps  $V, W : K \rightarrow \hat{H}$  with  $\|V\|, \|W\| \leq 1$  such that  $u(a) = V^* \pi(a) W$  for all  $a \in E$ .

Moreover if  $E$  is an operator system then  $u$  is CP if and only if it admits such a factorisation with  $V = W$ .

There are many names and a long history associated with this fact. Initially, Stinespring established his version of this in 1955 as a factorisation of CP maps defined on  $C^*$ -algebras, making it appear as an operator analogue of the classic GNS construction associated to a state on a  $C^*$ -algebra. Then Arveson proved his extension theorem in 1969—to the effect that a CP map on an operator system can be extended to a CP map on a  $C^*$ -algebra containing it (thereby generalizing Stinespring’s theorem so as to be valid for CP maps on operator systems). The final full generalization to CB maps on operator spaces, which may be attributed to several authors—Wittstock, Haagerup, Paulsen—came around 1981. We shall give a proof of the theorem which exhibits it as an extension theorem. But first, some examples:

*Example 1.1.6* The “transpose” mapping  $M_n \xrightarrow{T} M_n$ , given by  $T(a) = a^t$ , is CB with  $\|T\| = 1$  but  $\|T\|_{cb} = n$ .

*Proof* Let  $e_{i,j}$  be the matrix units of  $M_n$ . Consider the matrix  $a \in M_n(M_n)$  with  $(i, j)$ th entry given by  $e_{j,i}$ . It is not hard to see that  $a$  is a permutation matrix, hence unitary and  $\|a\| = 1$ .

Also,  $\frac{T_n(a)}{n}$  is seen to be the projection in  $M_n(M_n)$  onto the one-dimensional subspace spanned by the vector  $v = e_1 + e_{n+2} + e_{2n+3} + \cdots + e_{n^2}$  in  $\ell_2^{n^2}$ , where

$\{e_i\}$  is the standard orthonormal basis of  $\ell_2^{n^2}$ . Hence  $\|T_n(a)\| = n$ . In particular,

$$\|T\|_{cb} \geq \|T_n\| \geq \frac{\|T_n(a)\|}{\|a\|} = n.$$

Conversely, consider the subalgebra  $\Delta_n \subset M_n$  of  $n \times n$  diagonal matrices. Let  $E : M_n \rightarrow \Delta_n$  be the *conditional expectation* defined by  $E(a) = E([a_{i,j}]) = \text{diag}(a_{1,1}, \dots, a_{n,n})$ —which is easily verified to satisfy  $E(dad') = dE(a)d'$  for all  $d, d' \in \Delta_n$  and  $a \in M_n$ . Consider the permutation matrix

$$u = \begin{pmatrix} 0 & 0 & \cdots & 0 & 1 \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{pmatrix}.$$

A little calculation shows that  $E(au^k) = \text{diag}(a_{1,k+1}, a_{2,k+2}, \dots, a_{n,n})$ , and hence  $E(au^k)u^{-k}$  has  $a_{i,k+i}$  at  $(i, k+i)^{\text{th}}$  position for each  $1 \leq k \leq n, 1 \leq i \leq n$ , (indices are modulo  $n$ ) as its only non-zero entries. Thus  $a = \sum_{k=1}^n E(au^k)u^{-k}$  and  $a^t = \sum_{k=1}^n u^k E(au^k)$ .

Define  $u_l, u_r : M_n \rightarrow M_n$  by  $u_l(a) = ua$  and  $u_r(a) = au$ , respectively. Then  $u_l^k$  and  $u_r^k$  are isometries for all  $k$ . As  $(u_l)_n$  (resp.,  $(u_r)_n$ ) is seen to be the map of  $M_n(M_n)$  given by left- (resp., right-) multiplication by the block-diagonal unitary matrix with all diagonal blocks being given by  $u$ , we conclude that  $u_l, u_r$  are CB maps. (They are in fact *complete isometries*.)

Moreover, a conditional expectation is a CB map (since  $E : M_n \rightarrow \Delta_n$  being a conditional expectation implies that, for all  $k \in \mathbb{N}$ ,  $E_k : M_k(M_n) \rightarrow M_k(\Delta_n)$  is also a conditional expectation whence a projection and  $\|E_k\| \leq 1$  for all  $k$ ).

Hence the equation  $T = \sum_{k=1}^n u_r^k \circ E \circ u_l^k$  expresses  $T$  as the sum of  $n$  *complete contractions* thereby showing that also  $\|T\|_{cb} \leq n$ , whereby  $\|T\|_{cb} = n$ .  $\square$

*Example 1.1.7* It follows from the previous example that the transpose mapping, when thought of as a self-map  $T : K(\ell^2) \rightarrow K(\ell^2)$  of the space of compact operators on  $\ell^2$ , is isometric but not CB [since all finite rank matrices are in  $K(\ell^2)$ ].

**Theorem 1.1.8 (Stinespring)** *Given a unital  $C^*$ -algebra  $A$ , a Hilbert space  $K$ , and a unital CP map  $u : A \rightarrow B(K)$ , there exist a Hilbert space  $H$ , a  $*$ -representation  $\pi : A \rightarrow B(H)$  and an isometric embedding  $V : K \hookrightarrow H$  such that  $u(a) = V^*\pi(a)V$  for all  $a$  in  $A$ . Equivalently, if we identify  $K$  via  $V$  as a subspace of  $H$ , then  $u(a) = P_K\pi(a)|_K$ .*

*Proof* Let  $\alpha = \sum_{i \in I} a_i \otimes k_i, \beta = \sum_{j \in J} b_j \otimes k_j \in A \otimes K$ . The fact that  $u$  is CP implies that the equation  $\langle \alpha, \beta \rangle = \sum_{i \in I, j \in J} \langle u(b_i^* a_j) k_j, k_i \rangle_K$  defines a positive semi-definite sesquilinear form on  $A \otimes K$ . (*Reason:*  $u_n$  is positive and if



$I = \{1, \dots, n\}$ , then

$$\begin{pmatrix} a_1^* & 0 & \cdots & 0 \\ a_2^* & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ a_n^* & 0 & \cdots & 0 \end{pmatrix} \begin{pmatrix} a_1 & a_2 & \cdots & a_n \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix} \geq 0 \Rightarrow \langle \alpha, \alpha \rangle \geq 0.$$

Define  $H$  to be the result of “separation and completion of this semi-inner-product space”; i.e.,  $H$  is the completion of the quotient of  $A \otimes K$  by the radical  $N$  of the form  $\langle \cdot, \cdot \rangle$ . And it is fairly painless to verify that the equations

$$\pi(a)(b \otimes k + N) = ab \otimes k + N \text{ and } V(k) = (1 \otimes k) + N$$

define a  $*$ -representation  $\pi : A \rightarrow B(H)$  and an isometry  $V : K \rightarrow H$  which achieve the desired results.  $\square$

Before we get to a CB version of Arveson’s extension theorem, which may be viewed as a non-commutative Hahn-Banach theorem, we shall pause to discuss a parallel precursor, i.e., a (commutative) Banach space version due to Nachbin, namely:

**Proposition 1.1.9 (Nachbin’s Hahn-Banach Theorem)** *Given an inclusion  $E \subset B$  of Banach spaces, any  $u \in B(E, L^\infty(\mu))$ <sup>1</sup> admits an extension  $\tilde{u} \in B(B, L^\infty(\mu))$  such that  $\|\tilde{u}\| = \|u\|$ .*

The proof of Nachbin’s theorem relies on the following observation.

**Observation 1.1.10** For Banach spaces  $E$  and  $F$ , we have isometric identifications:

$$\begin{aligned} B(E, F^*) &:= \{\text{all bounded linear maps from } E \text{ to } F^*\} \\ &\simeq \text{Bil}(E \times F) := \{\text{all bounded bilinear forms on } E \times F\} \\ &\simeq (E \hat{\otimes} F)^*, \end{aligned}$$

where  $\hat{\otimes}$  is the projective tensor product, i.e., the completion of the linear span of the elementary tensor products with respect to the norm  $\|t\|_\wedge := \inf\{\sum_i \|a_i\| \cdot \|b_i\| : t = \sum_i a_i \otimes b_i\}$ .

---

<sup>1</sup>When we write symbols such as  $L^1(\mu)$  or  $L^\infty(\mu)$ , it will be tacitly assumed that we are talking about (equivalence classes of almost everywhere agreeing) complex-valued functions on some underlying measure space  $(\Omega, \mu)$ ; when we wish to allow vector-valued functions, we will write  $L^1(\mu; X)$ , etc.

The first isomorphism follows from the identification  $B(E, F^*) \ni T \mapsto \psi_T \in \text{Bil}(E \times F)$  given by  $\psi_T(a, b) = T(a)(b)$  for  $a \in E, b \in F$ , since

$$\begin{aligned} \|T\| &= \sup\{\|T(a)\| : \|a\| \leq 1\} \\ &= \sup\{|T(a)(b)| : \|a\|, \|b\| \leq 1\} \\ &= \|\psi_T\|, \end{aligned}$$

while the second isomorphism follows from the identification  $\text{Bil}(E \times F) \ni S \mapsto \phi_S \in (E \hat{\otimes} F)^*$  given by  $\phi_S(a \otimes b) = S(a, b)$ . On one hand,  $\|\phi_S\| = \sup\{|\phi_S(t)| : \|t\|_{E \hat{\otimes} F} \leq 1\}$ . On the other hand, for each representation  $t = \sum_i a_i \otimes b_i \in E \otimes F$ ,

$$\begin{aligned} |\phi_S(t)| &= \left| \sum_i S(a_i, b_i) \right| \\ &\leq \sum_i |S(a_i, b_i)| \\ &\leq \|S\| \sum_i \|a_i\| \cdot \|b_i\|. \end{aligned}$$

Taking infimum over all such representations of  $t$ , we obtain  $|\phi_S(t)| \leq \|S\| \|t\|_{E \hat{\otimes} F}$  implying  $\|\phi_S\| \leq \|S\|$ . Conversely, since  $\|a \otimes b\|_{E \hat{\otimes} F} = \|a\| \cdot \|b\|$  (as is easily verified),

$$\begin{aligned} \|\phi_S\| &\geq \sup\{|\phi_S(a \otimes b)| : \|a\| \cdot \|b\| \leq 1\} \\ &= \sup\{|S(a, b)| : \|a\| \cdot \|b\| \leq 1\} \\ &= \|S\|. \end{aligned}$$

*Proof of Theorem 1.1.9* Deduce from the above observation that

$$\begin{aligned} B(E, L^\infty(\mu)) &\cong (E \hat{\otimes} L^1(\mu))^* \\ &\cong L^1(\mu, E)^* \\ &\subseteq L^1(\mu, B)^* \\ &\cong B(B, L^\infty(\mu)), \end{aligned}$$

where the first and last isomorphisms follow from the Observation 1.1.10, the inclusion is an isometric imbedding as a consequence of the classical Hahn-Banach theorem, and the second isomorphism is justified in two steps as follows:

*Step I:* If  $X \hat{\otimes} Y$  denotes the projective tensor product of Banach spaces  $X$  and  $Y$ , if  $D_X$  (resp.  $D_Y$ ) is a dense subspace of  $X$  (resp.,  $Y$ ), and if  $\phi$  is any map

from the algebraic tensor product  $D_X \otimes D_Y$  to any Banach space  $Z$  such that  $\|\phi(x \otimes y)\|_Z = \|x\| \|y\| \quad \forall x \in D_X, y \in D_Y$ , then  $\phi$  extends uniquely to an element  $\tilde{\phi} \in B((X \hat{\otimes} Y), Z)$ .

*Reason:* If  $\sum_i x_i \otimes y_i = \sum_j u_j \otimes v_j$ , then, observe that

$$\begin{aligned} \|\phi(\sum_i x_i \otimes y_i)\| &= \|\sum_j \phi(u_j \otimes v_j)\| \\ &\leq \sum_j \|\phi(u_j \otimes v_j)\| \\ &= \sum_j \|u_j\| \|v_j\|. \end{aligned}$$

Taking the infimum over all such choices of  $u_j, v_j$ , we find that

$$\|\phi(\sum_i x_i \otimes y_i)\|_Z \leq \|\sum_i x_i \otimes y_i\|_{\wedge},$$

and the desired conclusion is seen to easily follow..

*Step II:* The assignment

$$L^1(\mu) \otimes Y \ni f \otimes y \mapsto f(\cdot) y \in L^1(\mu; Y)$$

extends to an isometric isomorphism  $\tilde{\phi}$  of  $L^1(\mu) \hat{\otimes} Y$  onto  $L^1(\mu; Y)$  for any Banach space  $Y$ .

*Reason:* Let  $X = L^1(\Omega, \mu)$ , and let  $D_X$  denote the subspace of measurable functions with finite range, and let  $D_Y = Y$ . Note that any element of  $D_X$  admits an essentially unique expression of the form  $\sum_i c_i 1_{E_i}$ , for some Borel partition  $\Omega = \bigsqcup_i E_i$ . It follows that any element of  $D_X \otimes Y$  has an essentially unique expression of the form  $\sum_i 1_{E_i} \otimes y_i$ , for some Borel partition  $\Omega = \bigsqcup_i E_i$  and some  $y_i \in Y$ . It follows from Step I that there exists a contractive operator  $\tilde{\phi} \in B(L^1(\mu) \hat{\otimes} Y, L^1(\mu; Y))$  such that  $\tilde{\phi}(1_{\Delta} \otimes y) = 1_{\Delta}(\cdot)y$ . Observe that if  $\sum_i 1_{E_i} \otimes y_i \in D_X \otimes Y$ , then

$$\begin{aligned} \|\tilde{\phi}(\sum_i 1_{E_i} \otimes y_i)\| &= \int \|\sum_i 1_{E_i}(\cdot) y_i\| \\ &= \sum_i \mu(E_i) \|y_i\| \\ &= \sum_i \|1_{E_i}\| \|y_i\| \\ &\geq \|\sum_i 1_{E_i} \otimes y_i\|_{L^1(\mu) \hat{\otimes} Y}, \end{aligned}$$

thereby showing that  $\tilde{\phi}$  is isometric on the dense subspace  $D_X \otimes Y$  and hence completing the proof.

**Theorem 1.1.11 (Arveson's Extension Theorem)** *Given an inclusion  $E \subset A$  of an operator space into a unital  $C^*$ -algebra, and any  $CB$  map  $u : E \rightarrow B(K)$  for some Hilbert space  $K$ , there exists  $\tilde{u} : A \xrightarrow{CB} B(K)$  extending  $u$  such that  $\|\tilde{u}\|_{cb} = \|u\|_{cb}$ .*

*Proof* Our strategy will involve proving that:

1. If, for  $t \in \overline{K} \otimes E \otimes H$ , we define

$$\begin{aligned} \gamma(t) &:= \inf \left\{ \left( \sum \|k_i\|^2 \right)^{1/2} \|[a_{i,j}]\|_{M_n(E)} \left( \sum \|h_j\|^2 \right)^{1/2} : t \right. \\ &= \left. \sum_{i,j=1}^n \bar{k}_i \otimes a_{i,j} \otimes h_j \right\}, \end{aligned}$$

then  $\gamma$  is a norm;

2. there is a natural isometric identification  $CB(E, B(H, K)) \cong (\overline{K} \otimes E \otimes H)^*$ , with  $(\overline{K} \otimes E \otimes H)^*$  being given the norm dual to  $\gamma$ ; and
3. if  $B$  is an "intermediate" operator space, meaning  $E \subset B \subset A$ , then  $(\overline{K} \otimes E \otimes H, \gamma) \subset (\overline{K} \otimes B \otimes H, \gamma)$  is an isometric embedding; i.e.,

$$\|\cdot\|_{\overline{K} \otimes E \otimes H} = (\|\cdot\|_{\overline{K} \otimes B \otimes H})|_{\overline{K} \otimes E \otimes H}.$$

It is obvious that saying  $\gamma(t) \geq 0$  amounts to showing that  $t$  admits an expression of the form  $t = \sum_{i,j=1}^n \bar{k}_i \otimes a_{i,j} \otimes h_j$ . In fact, there is such an expression with the matrix  $[a_{i,j}]$  being diagonal: indeed, if  $t = \sum_{j=1}^n \bar{k}_j \otimes e_j \otimes h_j$ , we can set  $a_{ij} = \delta_{ij} e_i$ . For  $t = \sum_{i,j=1}^n \bar{k}_i \otimes a_{i,j} \otimes h_j$ , let us use the suggestive notation  $t = (\bar{k}) \otimes a \otimes (h)$ .

We may assume that the infimum defining  $\gamma$  may be taken only over collections  $(\bar{k}) \otimes (a) \otimes (h)$  which satisfy

$$\|(a)\|_{M_n(E)} = 1 \text{ and } \|(\bar{k})\| = \sqrt{\sum_i \|\bar{k}_i\|^2} = \|(h)\|.$$

(Reason: This may be achieved by "spreading scalars over the tensor factors". More precisely, if  $C^2 = \|(\bar{k})\| \cdot \|a\| \cdot \|(h)\|$  and if we define

$$(\bar{k}') = \frac{C}{\|(\bar{k})\|} (\bar{k}), (a') = \frac{1}{\|(a)\|} (a), (h') = \frac{C}{\|(h)\|} (h),$$

then  $(\bar{k}') \otimes (a') \otimes (h') = (\bar{k}) \otimes (a) \otimes (h)$  and  $\|(a')\|_{M_n(E)} = 1$  and  $\|(\bar{k}')\| = \|(h')\|$ .)

For any other  $t' = \begin{pmatrix} \bar{k}' \\ \end{pmatrix} \otimes a' \otimes (h')$ , we find that  $t + t' = \begin{pmatrix} \bar{k} & \bar{k}' \\ \end{pmatrix} \otimes \begin{pmatrix} a & 0 \\ 0 & a' \end{pmatrix} \otimes \begin{pmatrix} h \\ h' \end{pmatrix}$ .

Thus,

$$\gamma(t + t') \leq \left( \sum \|k_i\|^2 + \sum \|k'_p\|^2 \right)^{1/2} \left( \sum \|h_j\|^2 + \sum \|h'_q\|^2 \right)^{1/2}.$$

(Note the crucial use of the fact that the operator norm of a direct sum of operators is the maximum of the norms of the summands—this is one of the identifying features of these matrix norms, as will be shown when we get to *Ruan's theorem*.)

For now, let us make the useful observation that:

$\gamma(t) < 1$  iff there exists a decomposition  $t = \begin{pmatrix} \bar{k} \\ \end{pmatrix} \otimes (a) \otimes (h)$  with  $\|a\|_{M_n(E)} = 1$  and  $\|\begin{pmatrix} \bar{k} \\ \end{pmatrix}\| = \|(h)\|^2 < 1$ . (†)

Now conclude from remark (†) that the above decompositions may be chosen to satisfy  $\|a\|_{M_n(E)} = 1$  and  $\|\begin{pmatrix} \bar{k} \\ \end{pmatrix}\| = \|(h)\| \sim \sqrt{\gamma(t)}$  and  $\|(a')\|_{M_n(E)} = 1$  and  $\|\begin{pmatrix} \bar{k}' \\ \end{pmatrix}\| = \|(h')\| \sim \sqrt{\gamma(t')}$ ; here  $\sim$  denotes approximate equality. (We are being slightly sloppy here in the interest of sparing ourselves the agony of performing calisthenics with epsilons.)

Thus,  $\gamma(t + t') \leq (\gamma(t) + \gamma(t'))^{1/2}(\gamma(t) + \gamma(t'))^{1/2} = \gamma(t) + \gamma(t')$ . Hence the triangle inequality holds for  $\gamma$ . It remains only to prove that  $\gamma(t) = 0 \Rightarrow t = 0$ .

So, suppose  $\gamma(t) = 0$ . Fix an arbitrary  $(\phi, \chi, \psi) \in (\bar{K})^* \times E^* \times H^*$  and  $\epsilon > 0$ . By assumption, there exists a decomposition  $t = \sum_{i,j=1}^n \bar{k}_i \otimes a_{i,j} \otimes h_j$ , with  $\|[a_{ij}]\|_{B(H^{(n)})} = 1$  and  $\|\begin{pmatrix} \bar{k} \\ \end{pmatrix}\|_{\bar{K}^{(n)}} = \|(h)\|_{H^{(n)}} < \epsilon$ . Hence,

$$\begin{aligned} |(\phi \otimes \chi \otimes \psi)(t)| &= \left| \sum \phi(\bar{k}_j) \chi(a_{ij}) \psi(h_j) \right| \\ &= \langle \chi_n(a) \begin{pmatrix} \psi(h_1) \\ \vdots \\ \psi(h_n) \end{pmatrix}, \begin{pmatrix} \phi(\bar{k}_1) \\ \vdots \\ \phi(\bar{k}_n) \end{pmatrix} \rangle \\ &\leq \|\chi_n(a)\| \left( \sum_{j=1}^n |\psi(h_j)|^2 \right)^{1/2} \left( \sum_{i=1}^n |\phi(\bar{k}_i)|^2 \right)^{1/2} \\ &\leq \|\chi_n\| \|a\| \|\phi\| \|\begin{pmatrix} \bar{k} \\ \end{pmatrix}\| \|\psi\| \|(h)\| \\ &\leq \|\chi\|_n \|\phi\| \|\psi\| \epsilon^2. \end{aligned}$$

Hence,

$$(\phi \otimes \chi \otimes \psi)(t) = 0 \quad \forall \phi \in (\bar{K})^*, \chi \in E^*, \psi \in H^*.$$

But this is seen to imply that  $t = 0$ , thereby proving (1).

As for (2), first note that  $B(H, K)$  has a natural operator space structure, by viewing it as the (2,1)-corner of  $B(H \oplus K)$  or equivalently by identifying

$M_n(B(H, K))$  with  $B(H^{(n)}, K^{(n)})$ . Consider the mapping

$$CB(E, B(H, K)) \ni u \mapsto \Phi_u \in (\bar{K} \otimes E \otimes H)^*$$

defined by  $\Phi_u(t) = \sum_{i,j} \langle u(a_{i,j})h_j, k_i \rangle$ , if  $t = \sum_{i,j} \bar{k}_i \otimes a_{i,j} \otimes h_j$ . Note that

$$\begin{aligned} \|\Phi_u\|_{\gamma^*} \leq 1 &\Leftrightarrow \sup_{\gamma(t) < 1} |\Phi_u(t)| \leq 1 \\ &\Leftrightarrow \sup \left\{ \left| \sum_{i,j} \langle u(a_{i,j})h_j, k_i \rangle \right| : \sum_j \|h_j\|^2 \right. \\ &\quad \left. = \sum_i \|k_i\|^2 < 1, \|[a_{i,j}]\|_{M_n(E)} = 1 \right\} \leq 1 \\ &\Leftrightarrow \sup \{ \|[u(a_{i,j})]\|_{M_n(B(H,K))} : [a_{i,j}] \in M_n(E), \|[a_{i,j}]\|_{M_n(E)} \\ &\quad = 1, n \geq 1 \} \leq 1 \\ &\Leftrightarrow \|u\|_{cb} \leq 1 \end{aligned}$$

Hence  $\|\Phi_u\|_{\gamma} = \|u\|_{cb}$  and the map  $u \mapsto \Phi_u$  is isometric.

We only need to verify now that  $\Phi$  is surjective. Suppose  $\Psi \in (\bar{K} \otimes E \otimes H)^*$ . Fix  $a \in E$ . If  $h \in H, k \in K$  and  $k \mapsto \bar{k}$  is an antiunitary map, then consider the clearly sesquilinear form defined on  $H \times K$  by  $[h, k] = \Psi(\bar{k} \otimes a \otimes h)$ . By definition of  $\gamma$ , we have  $\|[h, k]\| \leq \|\Psi\| \cdot \gamma(\bar{k} \otimes a \otimes h) \leq \|\Psi\| \cdot \|a\|_E \|h\| \|k\|$ , and hence  $[\cdot, \cdot]$  is a bounded sesquilinear form and there exists  $u \in B(H, K)$  such that  $\Psi(\bar{k} \otimes a \otimes h) = \langle uh, k \rangle$ . It is a routine application of the definition of  $\gamma$  to verify that  $u \in CB(H, K)$ , thus completing the proof of (2).

As for (3), it is clear that if  $t \in \bar{K} \otimes E \otimes H$ , it follows that  $\|t\|_B \leq \|t\|_E$ —where we write  $\|\cdot\|_C = \|\cdot\|_{(\hat{K} \otimes C \otimes H, \gamma)}$ —since the infimum over a larger collection is smaller. In order to prove the reverse inequality, we shall assume that  $\|t\|_B < 1$ , and show that  $\|t\|_E < 1$ . The assumption  $\|t\|_B < 1$  implies that there exists a decomposition

$$t = \sum_{i=1}^m \sum_{j=1}^n \bar{k}_i \otimes a_{i,j} \otimes h_j = (\bar{k}_1, \dots, \bar{k}_m) \otimes [a_{i,j}]_{m \times n} \otimes \begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix}$$

with  $a_{i,j} \in B$  such that

$$\left( \sum_{i=1}^m \|\bar{k}_i\|^2 \right)^{\frac{1}{2}} \cdot \|[a_{i,j}]\|_{M_{m \times n}(B)} \cdot \left( \sum_{j=1}^n \|h_j\|^2 \right)^{\frac{1}{2}} < 1.$$

It follows from the linear algebraic Observation 1.1.12 (discussed at the end of this proof) that there exist  $r \leq m$ , a linearly independent set  $\{\bar{k}'_p : 1 \leq p \leq r\}$  of vectors, and a matrix  $C \in M_{m \times r}$  (resp.,  $s \leq n$ , a linearly independent set  $\{h_q : 1 \leq q \leq s\}$  of vectors, and a matrix  $D \in M_{n \times s}$ ) such that  $\|C\| \leq 1$ ,  $\bar{k}_i = \sum_{p=1}^r \bar{c}_{ip} \bar{k}'_p$ , and  $\sum_{i=1}^m \|\bar{k}_i\|^2 = \sum_{p=1}^r \|\bar{k}'_p\|^2$  (resp.,  $\|D\| \leq 1$ ,  $h_j = \sum_{q=1}^s \bar{d}_{jq} h'_q$ , and  $\sum_{j=1}^n \|h_j\|^2 = \sum_{q=1}^s \|h'_q\|^2$ ).

Then, note that

$$\begin{aligned} t &= \sum_{i=1}^m \sum_{j=1}^n \bar{k}_i \otimes a_{ij} \otimes h_j \\ &= \sum_{i=1}^m \sum_{j=1}^n \left( \sum_{p=1}^r \bar{c}_{ip} \bar{k}'_p \right) \otimes a_{ij} \otimes \left( \sum_{q=1}^s \bar{d}_{jq} h'_q \right) \\ &= \sum_{p=1}^r \sum_{q=1}^s \bar{k}'_p \otimes a'_{pq} \otimes h'_q, \end{aligned}$$

where  $a'_{pq} = \sum_{i=1}^m \sum_{j=1}^n \bar{c}_{ip} a_{ij} \bar{d}_{jq}$ , i.e.,  $A' =: [a'_{pq}] = C^* A D$  and hence  $\|A'\| \leq \|A\| = 1$ .

Observe that since  $\{\bar{k}'_p : 1 \leq p \leq r\}$  and  $\{h'_q : 1 \leq q \leq s\}$  are linearly independent sets, then  $a'_{pq}$  for all  $p, q$  are forced to be in  $E$ , as  $t \in \bar{K} \otimes E \otimes H!$  (*Reason:* For each  $p, q$  take  $f_p \in (\bar{K})^*$  such that  $f_i(\bar{k}'_i) = \delta_{i,p}$  and  $g_q \in H^*$  such that  $g_q(h'_q) = \delta_{q,q}$ . Then  $f_p \otimes Id_E \otimes g_q : \bar{K} \otimes E \otimes H \rightarrow E$  and maps  $t$  to  $a'_{pq}$ .)

Thus we find—since  $\sum_{i=1}^m \|\bar{k}_i\|^2 = \sum_{j=1}^n \|\bar{k}'_j\|^2$ ,  $\sum_{j=1}^n \|h_j\|^2 = \sum_{j=1}^n \|h'_j\|^2$ ,  $A' \in M_{r \times s}(E)$  and  $\|A'\| = \|C^* A D\| \leq \|A\|$ —that

$$\begin{aligned} \|t\|_E &\leq \left( \sum_{p=1}^r \|\bar{k}'_p\|^2 \right)^{\frac{1}{2}} \cdot \|A'\| \cdot \left( \sum_{q=1}^s \|h'_q\|^2 \right)^{\frac{1}{2}} \\ &\leq \left( \sum_{i=1}^m \|\bar{k}_i\|^2 \right)^{\frac{1}{2}} \cdot \|A\| \cdot \left( \sum_{j=1}^n \|h_j\|^2 \right)^{\frac{1}{2}} \\ &< 1, \end{aligned}$$

thus establishing that  $\|t\|_B < 1 \Rightarrow \|t\|_E \leq 1$ ; hence, indeed  $\|t\|_B = \|t\|_E$ .

To complete the proof of the theorem, note that if  $u \in CB(E, B(K))$ , and if  $\Phi_u \in (\bar{K} \otimes E \otimes K)^*$  corresponds to  $u$  as in the proof of (2), then it is a consequence of (3) and the classical Hahn-Banach theorem that there exists a  $\tilde{\Phi} \in (\bar{K} \otimes A \otimes K)^*$  which extends, and has the same norm, as  $\Phi_u$ . Again, by (2), there exists a unique  $\tilde{u} \in CB(A, B(K))$  such that  $\tilde{\Phi} = \tilde{\Phi}_{\tilde{u}}$ . It follows from the definitions that  $\tilde{u}$  extends, and has the same CB-norm as,  $u$ .  $\square$

Now for the “linear algebraic” observation used in the above proof:

**Observation 1.1.12** *If  $h_1, \dots, h_n$  are elements of a Hilbert space  $H$ , then there exist  $r \leq n$ , vectors  $\{h'_j : 1 \leq j \leq r\}$  in  $H$  and a rectangular matrix  $C = [c_{ij}] \in M_{n \times r}$  such that*

1.  $CC^*$  is an orthogonal projection, and in particular,  $\|C\|_{M_{n \times r}} \leq 1$ ;
2.  $h_i = \sum_{j=1}^r c_{ij} h'_j$  for  $1 \leq i \leq n$ ;
3.  $\sum_{i=1}^n \|h_i\|^2 = \sum_{j=1}^r \|h'_j\|^2$ ; and
4.  $\{h'_j : 1 \leq j \leq r\}$  is a linearly independent basis for the linear span of  $\{h_i : 1 \leq i \leq n\}$ .

*Proof* Consider the linear operator  $T : \ell_n^2 \rightarrow H$  defined by  $Te_i = h_i$  for each  $1 \leq i \leq n$ , where of course  $\{e_i : 1 \leq i \leq n\}$  is the standard orthonormal basis for  $\ell_n^2$ . Let  $\{f_j : 1 \leq j \leq r\}$  denote any orthonormal basis for  $M = \ker^\perp(T)$  and let  $P$  denote the orthogonal projection of  $\ell_n^2$  onto  $M$ . Set  $h'_j = Tf_j$  for  $1 \leq j \leq r$ . Define  $c_{ij} = \langle f_j, e_i \rangle$  for  $1 \leq i \leq n, 1 \leq j \leq r$  and note that the matrix  $C = [c_{ij}] \in M_{n \times r}$  satisfies

$$\begin{aligned} (CC^*)_{i'i'} &= \sum_{j=1}^r c_{ij} \bar{c}_{i'j} \\ &= \sum_{j=1}^r \langle f_j, e_i \rangle \langle e_{i'}, f_j \rangle \\ &= \left\langle \sum_{j=1}^r \langle e_{i'}, f_j \rangle f_j, \sum_{j=1}^r \langle e_i, f_j \rangle f_j \right\rangle \\ &= \langle Pe_{i'}, Pe_i \rangle \\ &= \langle Pe_{i'}, e_i \rangle \end{aligned}$$

and so,  $CC^*$  denotes the projection  $P$ . Now observe that

$$h_i = Te_i = TPe_i = \sum_j \langle e_i, f_j \rangle Tf_j = \sum_j c_{ij} h'_j.$$

Finally observe that if  $\{f_j : r < j \leq n\}$  is an orthonormal basis for  $\ker(T)$ , then,

$$\sum_{j=1}^r \|h'_j\|^2 = \sum_{j=1}^r \|Tf_j\|^2 = \|TP\|_{HS}^2 = \|T\|_{HS}^2 = \sum_{i=1}^n \|Te_i\|^2 = \sum_{i=1}^n \|h_i\|^2,$$

thereby completing the proof of the observation.  $\square$

Now we discuss completely positive (CP) maps. We begin with a key lemma relating positivity to norm bounds:



**Lemma 1.1.13** For Hilbert spaces  $H$  and  $K$ , let  $a \in B(H)^+$ ,  $b \in B(K)^+$  and  $x \in B(K, H)$ . Then,

1.  $\begin{pmatrix} 1 & x \\ x^* & 1 \end{pmatrix} \in B(H \oplus K)_+ \Leftrightarrow \|x\| \leq 1$ ; and,
2. more generally  $\begin{pmatrix} a & x \\ x^* & b \end{pmatrix} \in B(H \oplus K)_+ \Leftrightarrow |\langle xk, h \rangle| \leq \sqrt{\langle ah, h \rangle \langle bk, k \rangle} \forall h \in H, k \in K$ .

*Proof* First note that (2)  $\Rightarrow \|x\| \leq \sqrt{\|a\| \cdot \|b\|} \Rightarrow$  (1)

As for (2),

$$\begin{aligned} \begin{pmatrix} a & x \\ x^* & b \end{pmatrix} \geq 0 &\Leftrightarrow \left\langle \begin{pmatrix} a & x \\ x^* & b \end{pmatrix} \begin{pmatrix} h \\ k \end{pmatrix}, \begin{pmatrix} h \\ k \end{pmatrix} \right\rangle \geq 0 \forall h \in H, k \in K \\ &\Leftrightarrow \langle ah, h \rangle + \langle bk, k \rangle + 2\operatorname{Re}\langle xk, h \rangle \geq 0 \forall h \in H, k \in K \\ &\Leftrightarrow |\langle xk, h \rangle| \leq \frac{\langle ah, h \rangle + \langle bk, k \rangle}{2} \forall h \in H, k \in K \\ &\Leftrightarrow \forall t > 0, |\langle xk, h \rangle| \leq \frac{1}{2}(t\langle ah, h \rangle + \frac{1}{t}\langle bk, k \rangle) \forall h \in H, k \in K. \end{aligned}$$

(on replacing  $h$  by  $\sqrt{t}h$  and  $k$  by  $\frac{k}{\sqrt{t}}$ )

Hence,

$$|\langle xk, h \rangle| \leq \inf_{t>0} \frac{1}{2} \left( t\langle ah, h \rangle + \frac{1}{t}\langle bk, k \rangle \right) = \sqrt{\langle ah, h \rangle \langle bk, k \rangle}$$

(using the fact that  $\frac{a+b}{2} \geq \sqrt{ab}$  and that  $\frac{a+b}{2} = \sqrt{ab} \Leftrightarrow a = b$ ). □

**Proposition 1.1.14** Let  $A$  be a unital  $C^*$ -algebra,  $S \subset A$  be an operator system and  $K$  be any Hilbert space. Then

1.  $\forall u \in CP(S, B(K)), \|u\| = \|u\|_{cb} = \|u(1)\|$ .
2. If  $u : S \rightarrow B(K)$  is linear with  $u(1) = 1$ , then  $\|u\|_{cb} \leq 1 \Leftrightarrow u$  is CP.

*Proof* (1)  $u$  is CP  $\Rightarrow u(S_+) \subset B(K)_+ \Rightarrow u(S_h) \subset B(K)_h \Rightarrow u(x^*) = u(x)^* \forall x \in S$  (by Cartesian decomposition). Hence by Lemma 1.1.13(1), we see that

$$\|x\| \leq 1 \Rightarrow \begin{pmatrix} 1 & x \\ x^* & 1 \end{pmatrix} \geq 0 \Rightarrow \begin{pmatrix} u(1) & u(x) \\ u(x^*) & u(1) \end{pmatrix} \geq 0.$$

Then, by Lemma 1.1.13(2), we find that  $\|u(x)\| \leq \|u(1)\|$  and hence  $\|u\| = \|u(1)\|$ . Similarly,  $x \in M_n(S)$  with  $\|x\| \leq 1$  implies  $\begin{pmatrix} 1 & x \\ x^* & 1 \end{pmatrix} \geq 0$  and hence

$$\|u_n(x)\| \leq \|u_n(1)\| = \|\text{diag}_n(u(1), \dots, u(1))\| = \|u(1)\|.$$

Hence,  $\|u\|_{cb} = \sup_n \|u_n\| \leq \|u(1)\|$ , i.e.,  $\|u\|_{cb} = \|u(1)\|$ , as desired.

(2) The proof of  $(\Leftarrow)$  is an immediate consequence of part (1) of this proposition and the assumed unitality of  $u$ .

We shall prove  $(\Rightarrow)$  using the fact that if  $\phi \in C^*$  for a commutative unital  $C^*$ -algebra  $C$  (i.e., we may assume  $C = C(\Omega)$  for some compact  $\Omega$ ), and if  $\phi(1) = 1$ , then  $\|\phi\| \leq 1 \Leftrightarrow \phi \geq 0$ . (It is a fact, which we shall not go into here, that fact is a special case of (2)!)

We first prove positivity of  $u$ , i.e., we need to verify that  $\langle u(x)h, h \rangle \geq 0$ ,  $\forall x \in S_+, h \in K$ . We may assume, without loss of generality that  $\|h\| = 1$ . Consider the commutative unital  $C^*$ -subalgebra  $A_0 = C^*(\{1, x\})$  of  $A$ . The linear functional  $\phi_0$  defined on  $S \cap A_0$  by  $\phi_0(a) := \langle u(a)h, h \rangle$  is seen to be bounded with  $\|\phi_0\| = \phi_0(1) = 1$ . Let  $\phi \in A_0^*$  be a Hahn-Banach extension of  $\phi_0$ . Since  $\|\phi\| = 1 = \phi(1)$ , it follows from the fact cited in the previous paragraph that  $\phi \geq 0$ . Hence  $\langle u(x)h, h \rangle = \phi(x) \geq 0$ , and the arbitrariness of  $h$  yields the positivity of  $u(x)$ . Thus, indeed  $u$  is a positive map.

To prove positivity of  $u_n$ ,  $n > 1$ , we need to verify that  $\langle u_n(x)h, h \rangle \geq 0$ ,  $\forall x \in M_n(S)_+, h \in K^{(n)}$ . First deduce from the positivity of  $u$  that  $u(a^*) = u(a)^* \forall a \in S$ , from which it follows that  $u_n(S_h) \subset M_n(B(K))_h$ . Also, note that  $u(1) = 1 \Rightarrow u_n(1) = 1$ . The assumption that  $\|u_n\| = 1$  now permits us, exactly as in the case  $n = 1$ , (by now considering the commutative  $C^*$ -subalgebra  $A_n := C^*(\{1, x\})$  of  $M_n(A)$ ) that  $u_n$  is also positive. Thus, indeed  $u$  is CP.  $\square$

**Corollary 1.1.15 (Arveson's Extension Theorem—CP Version)** *If  $S$  is an operator system contained in a  $C^*$ -algebra  $A$ , then any CP map  $u : S \rightarrow B(K)$  extends to a CP map  $\tilde{u} : A \rightarrow B(K)$ .*

*Proof* Normalise and assume  $\|u(1)\| \leq 1$ , whence it follows from the positivity of  $u$  that  $0 \leq u(1) \leq 1$ . Pick any state on  $A$ , i.e., a positive (=positivity-preserving element) of  $A^*$  such that  $\phi(1) = 1$ . (The existence of an abundance of states is one of the very useful consequences of the classical Hahn-Banach Theorem.)

Now, consider the map  $U : S \oplus S (\subset A \oplus A) \rightarrow B(K)$  defined by

$$U\left(\begin{pmatrix} x & 0 \\ 0 & y \end{pmatrix}\right) = u(x) + \phi(y)(1 - u(1))$$

and note that  $U$  is unital and completely positive, in view of Lemma 1.1.14(2); and hence [by Proposition 1.1.14(1)]  $\|U\|_{cb} = 1$ . Appeal to Theorem 1.1.11 to find an extension  $\tilde{U} \in CB(A \oplus A, B(K))$  with  $\|\tilde{U}\|_{cb} = 1$ . As  $\tilde{U}$  inherits the property of being unital from  $U$ , it follows by an application of Proposition 1.1.14(2) that  $\tilde{U}$  is

CP. Finally, if we define  $\tilde{u}(x) = \tilde{U}\left(\begin{pmatrix} x & 0 \\ 0 & 0 \end{pmatrix}\right)$  for  $x \in A$ , it is clear that  $\tilde{u}$  is a CP extension to  $A$  of  $u$ .  $\square$

**Lemma 1.1.16** *Let  $A$  be a unital  $C^*$ -algebra. Then,  $\phi \in A_+^* \Rightarrow \phi \in CP(A, \mathbb{C})$ .*

*Proof* We need to show that if  $[a_{i,j}] \in M_n(A)_+$ , then we have  $\sum_{i,j} \phi(a_{i,j})h_j\bar{h}_i \geq 0 \forall h \in \mathbb{C}^n$ . But that follows since  $\phi$  is positive and  $\sum_{i,j} h_j\bar{h}_i a_{i,j} = \begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix}^* [a_{i,j}] \begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix} \in A_+$ .  $\square$

We now discuss yet another useful  $2 \times 2$  matrix trick; this one also serves as a conduit from operator spaces to operator systems.

**Theorem 1.1.17** *Suppose  $E \subset B(H)$  is a subspace and  $w : E \rightarrow B(K)$  is a linear map. Define  $S = \left\{ \begin{pmatrix} \lambda 1 & x \\ y^* & \mu 1 \end{pmatrix} : x, y \in E, \lambda, \mu \in \mathbb{C} \right\} \subset M_2(B(H))$  and  $W : S \rightarrow M_2(B(K))$  by  $W\left(\begin{pmatrix} \lambda 1 & x \\ y^* & \mu 1 \end{pmatrix}\right) = \begin{pmatrix} \lambda 1 & w(x) \\ w(y)^* & \mu 1 \end{pmatrix}$ . Then  $E$  is an operator space and  $\|w\|_{cb} \leq 1 \Leftrightarrow S$  is an operator system and  $W$  is CP.*

*(Here the operator space/system structure on  $E/S$  is the natural one induced from the other.)*

*Proof* ( $\Leftarrow$ ) First note that if  $S$  is an operator system and  $W$  is linear and necessarily unital, then  $E$  (identified as the (1,2)-corner of  $S$ ) is an operator space, while it follows from two applications of Lemma 1.1.13(1) that if  $x \in E$ , then  $\|x\|_E \leq 1 \Rightarrow X = \begin{pmatrix} 1 & x \\ x^* & 1 \end{pmatrix} \in S_+$  and so  $W$  positive  $\Rightarrow W(X) = \begin{pmatrix} 1 & w(x) \\ w(x)^* & 1 \end{pmatrix} \in M_2(B(K))_+ \Rightarrow \|w(x)\|_{B(K)} \leq 1$ .

Now suppose  $x(n) \in M_n(E)$  and  $\|x(n)\|_{M_n(E)} \leq 1$ . Define  $X(n) \in M_n(S)$  by  $X(n)_{ij} = \begin{pmatrix} \delta_{ij} & x(n)_{ij} \\ (x(n)^*)_{ij} & \delta_{ij} \end{pmatrix}$ , where the Kronecker symbol  $\delta_{ij}$  denotes the  $ij$  entry of the identity matrix. It is not hard to see that there exists a permutation matrix  $P \in M_{2n}$ —independent of  $x(n)$ —such that  $PX(n)P^* = \begin{pmatrix} 1 & x(n) \\ (x(n)^*) & 1 \end{pmatrix}$ .

Similarly, define  $y(n) = w_n(x(n)) \in M_n(B(K))$  and  $Y(n) \in M_n(S)$  by  $Y(n)_{ij} = \begin{pmatrix} \delta_{ij} & y(n)_{ij} \\ (y(n)^*)_{ij} & \delta_{ij} \end{pmatrix}$  and deduce that with  $P$  as above, we have  $PY(n)P^* = \begin{pmatrix} 1 & y(n) \\ (y(n)^*) & 1 \end{pmatrix}$ .

It follows from the definitions that

$$(y(n)^*)_{ij} = (w_n(x(n))^*)_{ij}$$

$$\begin{aligned}
&= (w(x(n)_{ji}))^* \\
&= (w_n(x(n)^*))_{ij}
\end{aligned}$$

and hence,  $W_n(X(n)) = Y(n)$ . Now,

$$\begin{aligned}
\|x(n)\| \leq 1 &\Rightarrow PX(n)P^* \geq 0 \\
&\Rightarrow X(n) \geq 0 \\
&\Rightarrow Y(n) = W_n(X(n)) \geq 0 \\
&\Rightarrow PY(n)P^* \geq 0 \\
&\Rightarrow \|y_n\| \leq 1
\end{aligned}$$

so  $\|w_n\| \leq 1$ .

( $\Rightarrow$ ) It is clear that  $S$  is naturally an operator system if  $E$  is an operator space. We first prove positivity of  $W$ . (That of the  $W_n$ 's is proved similarly.)

We need to show that  $A = \begin{pmatrix} \lambda & b \\ c^* & \mu \end{pmatrix} \in S_+ \Rightarrow W(A) \geq 0$ . Since  $A \geq 0$ , we have  $b = c \in E$  and  $\lambda, \mu \geq 0$ . Also,  $A + \epsilon 1 \geq 0$  and since, by the definition of  $W$ , it is seen that  $W(A) = \lim_{\epsilon \downarrow 0} W(A + \epsilon 1)$ , we may assume without loss of generality that  $\lambda, \mu > 0$ .

Then,  $A = \begin{pmatrix} \lambda 1 & b \\ b^* & \mu 1 \end{pmatrix} = \begin{pmatrix} \sqrt{\lambda} & 0 \\ 0 & \sqrt{\mu} \end{pmatrix} \begin{pmatrix} 1 & x \\ x^* & 1 \end{pmatrix} \begin{pmatrix} \sqrt{\lambda} & 0 \\ 0 & \sqrt{\mu} \end{pmatrix}$ , where  $x = b/\sqrt{\lambda\mu}$ .

Hence

$$\begin{aligned}
\begin{pmatrix} \lambda 1 & b \\ b^* & \mu 1 \end{pmatrix} \geq 0 &\Leftrightarrow \begin{pmatrix} 1 & x \\ x^* & 1 \end{pmatrix} \geq 0 \\
&\Leftrightarrow \|x\| \leq 1 \\
&\Rightarrow \|w(x)\| \leq 1 \\
&\Rightarrow \begin{pmatrix} 1 & w(x) \\ w(x^*) & 1 \end{pmatrix} \geq 0 \\
&\Rightarrow W(A) = \begin{pmatrix} \sqrt{\lambda} & 0 \\ 0 & \sqrt{\mu} \end{pmatrix} \begin{pmatrix} 1 & w(x) \\ w(x^*) & 1 \end{pmatrix} \begin{pmatrix} \sqrt{\lambda} & 0 \\ 0 & \sqrt{\mu} \end{pmatrix} \geq 0.
\end{aligned}$$

Arguing with the permutation as in the proof of the reverse implication, and proceeding as above (in the  $n = 1$  case), we find that  $\|w_n\| \leq 1 \Rightarrow W_n$  is positive.  $\square$

Now we are ready **to prove the Fundamental Factorisation Theorem 1.1.5**.

*Proof* Normalise and assume  $\|u\|_{cb} \leq 1$ . Let the operator system  $S$  and  $U : S \rightarrow M_2(B(K)) = B(K^{(2)})$  be the (unital) CP map associated to the operator space

$E$  and the completely contractive map  $u : E \rightarrow B(K)$  (like the  $w \leftrightarrow W$ ) as in Theorem 1.1.17. As  $E$  is a subset of the  $C^*$ -algebra  $A$  by assumption, we find that  $S \subset M_2(A)$

Then by the CP version (Corollary 1.1.15) of Arveson's extension theorem, we can extend  $U$  to a CP map  $\tilde{U} : M_2(A) \rightarrow B(K^{(2)})$ .

As  $M_2(A)$  is a unital  $C^*$ -algebra and  $\tilde{U}$  is a unital CP map, we may conclude from Stinespring's theorem that there exists a representation  $\sigma : M_2(A) \rightarrow B(H)$  and an isometry  $T : K \oplus K \rightarrow H$  such that  $\tilde{U}(\tilde{x}) = T^* \sigma(\tilde{x}) T$  for all  $\tilde{x} \in M_2(A)$ . Hence,

$$\begin{aligned} \begin{pmatrix} 0 & u(x) \\ 0 & 0 \end{pmatrix} &= U \begin{pmatrix} 0 & x \\ 0 & 0 \end{pmatrix} = \tilde{U} \begin{pmatrix} 0 & x \\ 0 & 0 \end{pmatrix} = T^* \sigma \left( \begin{pmatrix} 0 & x \\ 0 & 0 \end{pmatrix} \right) T \\ &= T^* \sigma \left( \begin{pmatrix} x & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \right) T = T^* \pi(x) T', \end{aligned}$$

where  $T' = \sigma \left( \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \right) T$  and  $\pi(x) = \sigma \left( \begin{pmatrix} x & 0 \\ 0 & 0 \end{pmatrix} \right)$ . Now, consider the projection  $P := \sigma \left( \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \right)$  and define  $\hat{H} = P(H)$ ,  $Vk = PT(k, 0)$ ,  $Wk = \sigma \left( \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \right) T(k, 0)$ , and note that  $P$  commutes with  $\sigma \left( \begin{pmatrix} x & 0 \\ 0 & 0 \end{pmatrix} \right)$  for each  $x \in A$ , so that the equation  $\pi(x) = \sigma \left( \begin{pmatrix} x & 0 \\ 0 & 0 \end{pmatrix} \right)_{\hat{H}}$  does define a (unital) representation of  $A$  on  $\hat{H}$ ; and we finally see that  $V, W : K \rightarrow \hat{H}$  and indeed

$$u(x) = V^* \pi(x) W \text{ for all } x \in E, \tag{1.1}$$

with  $\|V\|, \|W\| \leq 1$ , as desired.

In the converse direction, if  $u$  admits the factorisation (1.1), then it is seen that also

$$u_n([a_{ij}]) = \begin{pmatrix} V^* \\ \vdots \\ V^* \end{pmatrix} \pi_n([a_{ij}]) (W, \dots, W)$$

and as  $\pi_n$  is as much of a representation (and hence contractive) as  $\pi$ , it is seen that  $\|u\|_{cb} \leq 1$ .

The non-trivial implication in the final assertion of the theorem is a consequence of Arveson's extension Theorem 1.1.15 and Stinespring's theorem.  $\square$

*Remark 1.1.18* For  $V = W$ ,  $u(x) = V^* \pi(x) V \Rightarrow u$  is CP (since a representation is obviously CP).

**Corollary 1.1.19** Any  $u \in CB(S, B(K))$  can be decomposed as  $u = u_1 - u_2 + i(u_3 - u_4)$ , with  $u_j \in CP(S, B(K))$ ,  $j = 1, 2, 3, 4$ .

*Proof* The proof is a consequence of the fundamental factorisation, the polarisation identity and Remark 1.1.18.  $\square$

## 1.2 More on CB and CP Maps

**Theorem 1.2.1 (Kraus Decomposition)** A linear map  $u : M_n \rightarrow M_m$  is CP if and only if there exists a

family  $\{V_p : 1 \leq p \leq N\} \subset M_{n \times m}$  with  $N \leq nm$   
such that  $u(a) = \sum_{p=1}^N V_p^* a V_p$  for all  $a \in M_n$ .

*Proof* Suppose  $u$  is CP. Then, by the fundamental factorization Theorem 1.1.5, there is a Hilbert space  $\hat{H}$ , a  $*$ -representation  $\pi : M_n \rightarrow B(\hat{H})$  and a map  $V : \ell_2^m \rightarrow \hat{H}$  such that  $u(a) = V^* \pi(a) V$ . It is a basic fact that for any representation of  $M_n$ , as above, there exists a Hilbert space  $H$  such that  $\hat{H} = (\cong) \ell_2^n \otimes H$  and  $\pi(a) = a \otimes 1$ .

Further, it is also true that there exists a subspace  $H_1 \subset H$  with  $\dim H_1 \leq mn$  such that  $V(\ell_2^m) \subset \ell_2^n \otimes H_1$ .

(Reason: If  $\{e_j : 1 \leq j \leq m\}$  and  $\{e_i : 1 \leq i \leq n\}$  are orthonormal bases for  $\ell_2^m$  and  $\ell_2^n$ , respectively, we see that there must exist operators  $T_i : \ell_2^m \rightarrow H$  such that

$$V e_j = \sum_{i=1}^n e_i \otimes T_i e_j \quad \forall 1 \leq j \leq m.$$

Clearly, then, if  $H_1 = \text{span}\{T_i e_j : 1 \leq j \leq m, 1 \leq i \leq n\}$ , then  $\dim H_1 \leq nm$  and  $V(\ell_2^m) \subset \ell_2^n \otimes H_1$ .)

Therefore, it follows that if  $\{e_p, 1 \leq p \leq N\}$  is an orthonormal basis for  $H_1$ , then  $N \leq nm$  and there exist  $V_p : \ell_2^m \rightarrow \ell_2^n, 1 \leq p \leq N$  such that  $V(\xi) = \sum_p V_p(\xi) \otimes e_p$  for all  $\xi \in \ell_2^m$ .

Finally, it is seen that for all  $a \in M_n, \xi, \eta \in \ell_2^m$ , we have

$$\begin{aligned} \langle u(a)\xi, \eta \rangle &= \langle (a \otimes 1) \left( \sum_p V_p(\xi) \otimes e_p \right), \sum_q V_q(\eta) \otimes e_q \rangle \\ &= \sum_p \langle a V_p(\xi), V_p(\eta) \rangle \\ &= \left\langle \sum_p V_p^* a V_p(\xi), \eta \right\rangle \end{aligned}$$

Conversely, it is not hard to see that any map admitting a decomposition of the given form is necessarily CP.  $\square$

The above result may be regarded as one of the first links between Operator Space Theory and Quantum Information Theory.

**Definition 1.2.2** If a linear map  $u : M_n \rightarrow M_m$  is CP and preserves the trace, then it is called a **quantum channel**.

*Remark 1.2.3* In the set up of Theorem 1.2.1,  $u$  preserves the trace if and only if  $\sum_p V_p V_p^* = I$ ; while it is identity-preserving if and only if  $\sum_p V_p^* V_p = I$ .

**Theorem 1.2.4 (Choi)** *The following conditions on a linear map  $u : M_n \rightarrow B(K)$  are equivalent:*

1.  $u$  is CP.
2.  $u$  is  $n$ -positive.
3.  $[u(e_{ij})] \in M_n(B(K))_+$ , where  $\{e_{ij} : 1 \leq i, j \leq n\}$  is the canonical system of matrix units for  $M_n$ .

*Proof* (3)  $\implies$  (1).  $[u(e_{ij})] \geq 0 \implies [u(e_{ij})] = X^*X$  for some  $X = [x_{ij}] \in M_n(B(K))$ . So,  $u(e_{ij}) = \sum_k x_{ki}^* x_{kj}$ . In particular,

$$u(a) = \sum_{ij} a_{ij} u(e_{ij}) = \sum_{i,j,k} x_{ki}^* a_{ij} x_{kj} = \sum_k \left( \sum_{ij} x_{ki}^* a_{ij} x_{kj} \right)$$

for all  $a = [a_{ij}] \in M_n$ .

Define  $u_k(a)$  to be the element of  $M_n(B(K))$  with  $(i, j)$ th entry equal to  $\sum_{ij} x_{ki}^* a_{ij} x_{kj}$ . Then, clearly,

$$u_k(a) = \begin{pmatrix} x_{k1}^* & x_{k2}^* & \cdots & x_{kn}^* \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \dots & & & \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} x_{k1} \\ x_{k2} \\ \dots \\ x_{kn} \end{pmatrix}$$

Hence  $u$  has the form  $u(a) = \sum_{k=1}^n V_k^* a V_k$ , where  $V_k : K \rightarrow \ell_2^n \otimes K$  is given by  $V_k(\xi) = \sum_j e_j \otimes x_{kj}(\xi)$  with  $\{e_i\}$  denoting the standard o.n.b. of  $\ell_2^n$ . The desired implication follows now from Theorem 1.2.1.

The implication (1)  $\implies$  (2) is trivial while (2)  $\implies$  (3) is a consequence of the fact that  $[e_{ij}] \in (M_n)_+$ .  $\square$

**Lemma 1.2.5 (Roger Smith)** *Fix  $N \geq 1$  (resp., a compact Hausdorff space  $\Omega$ ) and an operator space  $E$ . Then every bounded linear map  $u : E \rightarrow M_N$  (resp.,  $u : E \rightarrow C(\Omega, M_N)$ ) is CB with  $\|u\|_{cb} = \|u_N : M_n(E) \rightarrow M_N(M_N)\|$  (resp.,  $\|u\|_{cb} = \|u_N : M_n(E) \rightarrow M_N(C(\Omega, M_N))\|$ ). In particular, every bounded linear map  $u : E \rightarrow C(\Omega)$  is CB with  $\|u\|_{cb} = \|u\|$ .*

*Proof* Suppose first that  $u \in B(E, M_n(\mathbb{C}))$ . To prove that  $\|u\|_{cb} = \|u_N\|$ , it clearly suffices to verify that  $\|u_n\| \leq \|u_N\| \forall n \geq N$ . We need to verify that if  $a = [a_{ij}] \in M_n(E)$  and  $x_1, \dots, x_n, y_1, \dots, y_n \in \mathbb{C}^N$  satisfy  $\|a\|_{M_n(E)} \leq 1$  and  $\sum_{i=1}^n \|x_i\|^2 =$

$\sum_{i=1}^n \|y_i\|^2 = 1$ , then

$$\left| \sum_{i,j=1}^n \langle u(a_{ij})y_j, x_i \rangle \right| \leq \|u_N\|.$$

For this, appeal first to Observation 1.1.12 to find  $\alpha, \beta \in M_{n \times n}(\mathbb{C})$  and  $x'_1, \dots, x'_n, y'_1, \dots, y'_n \in \mathbb{C}^n$  such that

1.  $\sum_{i=1}^n \|x'_i\|^2 = \sum_{j=1}^n \|y'_j\|^2 = 1$ ;
2.  $x_i = \sum_{l=1}^n \alpha_{il}x'_l$  and  $y_j = \sum_{k=1}^n \beta_{jk}y'_k$  for  $1 \leq i, j \leq n$ , and
3.  $\|\alpha\|_{M_{n \times n}(\mathbb{C})}, \|\beta\|_{M_{n \times n}(\mathbb{C})} \leq 1$

Deduce, then, that

$$\begin{aligned} \left| \sum_{i,j=1}^n \langle u(a_{ij})y_j, x_i \rangle \right| &= \left| \sum_{i,j=1}^n \sum_{k,l=1}^n \langle u(a_{ij})\beta_{jk}y'_k, \alpha_{il}x'_l \rangle \right| \\ &= \left| \sum_{k,l=1}^n \langle u(\alpha^* a \beta)_{lk} y'_k, x'_l \rangle \right| \\ &\leq \|u_N(\alpha^* a \beta)\| \\ &\leq \|u_N\| \cdot \|\alpha^* a \beta\| \\ &\leq \|u_N\|, \end{aligned}$$

as desired.

Next, suppose  $u \in B(E, C(\Omega; M_N(\mathbb{C})))$ . Let us introduce the notation  $\Omega \ni \omega \mapsto \phi_\omega \in E^*$  where  $\phi_\omega(x) = u(x)(\omega)$ , and note that

$$\begin{aligned} \|\phi_\omega\| &= \sup\{|\phi_\omega(x)| : \|x\| \leq 1\} \\ &= \sup\{|u(x)(\omega)| : \|x\| \leq 1\} \\ &\leq \sup\{\|u(x)\| : \|x\| \leq 1\} \\ &= \|u\|. \end{aligned}$$

Now conclude that if  $[a_{ij}] \in M_n(E)$  and  $u(a_{ij}) = f_{ij}$ , then

$$\begin{aligned} \|u_n([a_{ij}])\| &= \sup_{\omega \in \Omega} \|[f_{ij}(\omega)]\| \\ &= \sup_{\omega \in \Omega} \|\phi_\omega([a_{ij}])\| \\ &= \sup_{\omega \in \Omega} \|(\phi_\omega)_n([a_{ij}])\| \end{aligned}$$



$$\begin{aligned}
&\leq \sup_{\omega \in \Omega} \|(\phi_\omega)_n\| \cdot \|[a_{ij}]\|_{M_n(E)} \\
&= \sup_{\omega \in \Omega} \|\phi_\omega\| \cdot \|[a_{ij}]\|_{M_n(E)} \\
&\leq \|u\| \cdot \|[a_{ij}]\|_{M_n(E)},
\end{aligned}$$

where we have used the fact that  $\|\phi_\omega\|_{cb} = \|\phi_\omega\|$  as  $\phi_\omega \in B(E, \mathbb{C})$ , and we hence find that  $\|u\|_{cb} \leq \|u\|$ .  $\square$

**Proposition 1.2.6** *Let  $E \subset B(H)$  and  $F \subset B(K)$  be operator spaces and  $u \in B(E, F)$ . Then  $\|u\|_{cb} = \|u\|$  in the following cases:*

1.  $\text{Rank}(u) \leq 1$ .
2.  $C^*(F)$  is commutative.
3.  $E = F = R$  or  $E = F = C$ .

*Proof* 1. If  $0 \neq f \in u(E)$ , then, clearly there exists  $\phi \in E^*$  such that  $u(e) = \phi(e)f \forall e \in E$ . Deduce from the case  $N = 1$  of Roger Smith's Lemma 1.2.5 that  $\|\phi\|_{cb} = \|\phi\|$ , and hence if  $a = [a_{ij}] \in M_n(E)$ , then  $\|u_n(a)\| = \|\phi(a_{ij})f\| \leq \|\phi_n(a)\| \cdot \|f\| \leq \|\phi\| \cdot \|a\| \cdot \|f\| = \|u\| \cdot \|a\|$  and hence indeed  $\|u\|_{cb} = \|u\|$ .

2. This is an immediate consequence of Roger Smith's Lemma 1.2.5.

3. Let us discuss the case of  $C := \overline{\text{span}}\{e_{i1} : i \geq 1\} \subset B(\ell_2)$  which clearly admits the identification  $C \cong \ell^2 \cong B(\mathbb{C}, \ell^2)$  thus:

$$C(\xi) = \sum_{i=1}^{\infty} \lambda_i e_{i1} \leftrightarrow \xi = \sum_{i=1}^{\infty} \lambda_i e_i \leftrightarrow \rho_\xi(\alpha) = \alpha \xi.$$

Observe now that if  $u \in B(C)$  corresponds as above to  $u' \in B(\ell^2)$  and to  $u'' \in B(B(\mathbb{C}, \ell^2))$ , then

$$u(C(\xi)) = C(u'\xi) \text{ and } u''(\rho_\xi) = \rho_{u'\xi},$$

and hence it follows that  $\|u\|_{cb} = \|u''\|_{cb} = \|u'\| = \|u\|$ , where we have used the obvious fact that

$$\|u''_n([\rho_{\xi_{ij}}])\| = \|[\rho_{u'\xi_{ij}}]\| = \left\| \begin{bmatrix} u'' & 0 & \cdots & 0 \\ 0 & u'' & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & u'' \end{bmatrix} [\rho_{\xi_{ij}}] \right\| \leq \|u''\| \cdot \|[\rho_{\xi_{ij}}]\|.$$

The case of  $R$  is proved similarly.

Note that  $R$  and  $C$  also have Hilbert space structures with o.n.b.'s  $\{e_{1j}\}$  and  $\{e_{i1}\}$  respectively.  $\square$

**Proposition 1.2.7** *For all  $u \in CB(R, C)$  (resp.,  $u \in CB(C, R)$ ),  $\|u\|_{cb} = \|u\|_{HS}$ .*

*Proof* First consider  $u \in CB(R, C)$ . (The case of  $u \in CB(C, R)$  is treated analogously.) Notice that if

$$x_n = \begin{pmatrix} e_{11} & 0 & \cdots & 0 \\ e_{12} & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ e_{1n} & 0 & \cdots & 0 \end{pmatrix} \in M_n(R),$$

then  $\|x_n\| = 1$ , and hence  $\|u_n(x_n)\|_{M_n} \leq \|u\|_{cb}$ .

Hence,

$$\|u\|_{cb} \geq \|u_n(x_n)\| = \left\| \begin{pmatrix} u(e_{11}) & 0 & \cdots & 0 \\ u(e_{12}) & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ u(e_{1n}) & 0 & \cdots & 0 \end{pmatrix} \right\| = \left\| \sum_{j=1}^n u(e_{1j}^*)u(e_{1j}) \right\|^{\frac{1}{2}}.$$

So, if  $u(e_{1j}) = \sum_{i=1}^{\infty} u_{ij}e_{i1}$ , then

$$\begin{aligned} \left\| \sum_{j=1}^n u(e_{1j}^*)u(e_{1j}) \right\|^{\frac{1}{2}} &= \left\| \sum_{j=1}^n \sum_{i,k=1}^{\infty} (u_{ij}e_{i1})^* (u_{kj}e_{k1}) \right\|^{\frac{1}{2}} \\ &= \left\| \sum_{j=1}^n \sum_{i,k=1}^{\infty} \overline{u_{ij}}u_{kj}e_{i1}e_{k1} \right\|^{\frac{1}{2}} \\ &= \left\| \sum_{j=1}^n \sum_{i,k=1}^{\infty} \overline{u_{ij}}u_{kj}\delta_{ik}e_{11} \right\|^{\frac{1}{2}} \\ &= \left( \sum_{j=1}^n \sum_{k=1}^{\infty} |u_{kj}|^2 \right)^{\frac{1}{2}}. \end{aligned}$$

Letting  $n \rightarrow \infty$ , we see that  $\|u\|_{HS} \leq \|u\|_{cb}$ .

For the reverse inequality, since the linear span of the  $e_{1j}$ 's is dense in  $R$ , deduce that  $\{\sum_{j=1}^m x_j \otimes e_{1j} : m \in \mathbb{N}, x_1, \dots, x_m \in M_n(\mathbb{C})\}$  is dense in  $M_n(\mathbb{C}) \otimes R = M_n(R)$ , so it suffices to verify that

$$\left\| \sum_{j=1}^m x_j \otimes u(e_{1j}) \right\|_{M_n(C)} \leq \|u\|_{HS} \cdot \left\| \sum_{j=1}^m x_j \otimes e_{1j} \right\|_{M_n(R)}$$

As  $R$  and  $C$  are both isometric to Hilbert space,  $u$  is in particular a bounded operator between Hilbert spaces, and its matrix with respect to the natural orthonormal

bases for domain and range is  $((u_{ij}))$  as above. And we see that

$$\begin{aligned}
& \left\| \sum_{j=1}^m x_j \otimes u(e_{1j}) \right\|_{M_n(\mathbb{C})} \\
&= \left\| \sum_{j=1}^m \left( x_j \otimes \sum_{i=1}^{\infty} u_{ij} e_{i1} \right) \right\|_{M_n(\mathbb{C})} \\
&= \left\| \sum_{i=1}^{\infty} \left( \sum_{j=1}^m u_{ij} x_j \right) \otimes e_{i1} \right\|_{M_n(\mathbb{C})} \\
&= \left\| \left( \sum_{i=1}^{\infty} \left( \sum_{j=1}^m u_{ij} x_j \otimes e_{i1} \right) \right)^* \left( \sum_{k=1}^{\infty} \left( \sum_{j=1}^m u_{kj} x_j \right) \otimes e_{k1} \right) \right\|_{M_n(\mathbb{C})}^{\frac{1}{2}} \\
&= \left\| \left( \sum_{i=1}^{\infty} \left( \sum_{j=1}^m \overline{u_{ij}} x_j^* \right) \otimes e_{i1} \right) \left( \sum_{k=1}^{\infty} \left( \sum_{j=1}^m u_{kj} x_j \right) \otimes e_{k1} \right) \right\|_{M_n(\mathbb{C})}^{\frac{1}{2}} \\
&= \left\| \sum_{i=1}^{\infty} \left( \sum_{j=1}^m \overline{u_{ij}} x_j^* \right) \left( \sum_{j=1}^m u_{ij} x_j \right) \right\|_{M_n(\mathbb{C})}^{\frac{1}{2}} \\
&\leq \left( \sum_{i=1}^{\infty} \left\| \sum_{j=1}^m u_{ij} x_j \right\|^2 \right)^{\frac{1}{2}}.
\end{aligned}$$

On the other hand,

$$\begin{aligned}
\left\| \sum_{j=1}^m u_{ij} x_j \right\| &= \left\| \begin{bmatrix} x_1 & \cdots & x_m \\ 0 & \cdots & 0 \\ \vdots & \vdots & \vdots \\ 0 & \cdots & 0 \end{bmatrix} \begin{bmatrix} u_{i1} & 0 & \cdots & 0 \\ u_{i2} & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ u_{i2} & 0 & \cdots & 0 \end{bmatrix} \right\| \\
&\leq \left\| \begin{bmatrix} x_1 & \cdots & x_m \\ 0 & \cdots & 0 \\ \vdots & \vdots & \vdots \\ 0 & \cdots & 0 \end{bmatrix} \right\| \cdot \left\| \begin{bmatrix} u_{i1} & 0 & \cdots & 0 \\ u_{i2} & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ u_{i2} & 0 & \cdots & 0 \end{bmatrix} \right\| \\
&= \left\| \sum_{j=1}^m x_j x_j^* \right\|^{\frac{1}{2}} \left( \sum_{j=1}^m |u_{ij}|^2 \right)^{\frac{1}{2}},
\end{aligned}$$

and hence we see that

$$\begin{aligned}
\left\| \sum_{j=1}^m x_j \otimes u(e_{1j}) \right\|_{M_n(C)} &\leq \left( \sum_{i=1}^{\infty} \left\| \sum_{j=1}^m x_j x_j^* \right\| \sum_{j=1}^m |u_{ij}|^2 \right)^{\frac{1}{2}} \\
&\leq \left( \left\| \sum_{j=1}^m x_j x_j^* \right\| \sum_{i,j=1}^{\infty} |u_{ij}|^2 \right)^{\frac{1}{2}} \\
&= \left\| \sum_{j=1}^m x_j \otimes e_{1j} \right\|_{M_n(R)} \|u\|_{HS}.
\end{aligned}$$

□

**Definition 1.2.8** Let  $E$  and  $F$  be operator spaces. A linear map  $u : E \rightarrow F$  is said to be a

1. **complete isomorphism** if  $u$  is a linear isomorphism such that  $u$  and  $u^{-1}$  are both CB.
2. **complete contraction** if  $u$  is CB with  $\|u\|_{cb} \leq 1$ .
3. **complete isometry** if  $u_n$  is an isometry for all  $n \geq 1$ .

**Corollary 1.2.9**  $R$  and  $C$  are not isomorphic to each other as operator spaces.

*Proof* Suppose  $u : R \rightarrow C$  and  $u^{-1} : C \rightarrow R$  are both CB maps. Then  $u$  and  $u^{-1}$  are Hilbert Schmidt and hence compact on an infinite dimensional Hilbert space, a contradiction. □

It is a fact that even for  $n$ -dimensional row and column spaces  $R_n$  and  $C_n$ , we have

$$\inf\{\|u\|_{cb}\|u^{-1}\|_{cb} : u \in CB(R_n, C_n) \text{ invertible}\} = n.$$

This is “worst possible” in view of the fact that for any two  $n$ -dimensional operator spaces  $E$  and  $F$ ,

$$\inf\{\|u\|_{cb}\|u^{-1}\|_{cb} : u \in CB(F, F) \text{ invertible}\} \leq n.$$

**Proposition 1.2.10** Let  $A$  be a commutative  $C^*$ -algebra and  $u \in B(A, B(K))$ . If  $u$  is positive, then  $u$  is completely positive.

*Proof* We first indicate the proof for finite-dimensional  $A$ .

Thus, suppose  $\dim A = n$ . Then

$$A \simeq \ell_n^\infty = \left\{ \begin{pmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{pmatrix} : \lambda_i \in \mathbb{C} \right\}.$$

Now  $u : \ell_n^\infty \rightarrow B(K)$  is positive if and only if  $a_j := u(e_j) \geq 0, \forall j$ . Finally, we get

$$u(x) = u\left(\sum_j x_j e_j\right) \sim V^* x V \quad \forall x \in \ell_n^\infty,$$

where  $V = \begin{bmatrix} a_1^{\frac{1}{2}} \\ \vdots \\ a_n^{\frac{1}{2}} \end{bmatrix}$  and hence  $u$  is CP.

We now give a sketch of the proof in the general case. It suffices to show that if  $X$  is a compact Hausdorff space, and  $\phi : C(X) \rightarrow B$  is a positive map into a unital  $C^*$ -algebra  $B$ , then  $\phi$  is  $n$ -positive for an arbitrarily fixed  $n$ . The first thing to observe is that there exist natural identifications of  $M_n(C(X)) \cong C(X; M_n(\mathbb{C}))$  with  $C^*$ -completions of the algebraic tensor product  $C(X) \otimes M_n(\mathbb{C})$  whereby the elementary tensor  $f \otimes T$  corresponding to  $f(\cdot)T$ ; and further  $\phi_n = \phi \otimes \mathbb{I}_{M_n(\mathbb{C})}$  which maps elementary tensors  $f \otimes T$ , with  $f, T \geq 0$  to the element  $\phi(f) \otimes T$  of  $B \otimes M_n(\mathbb{C})$  which is positive, being the product of two commuting positive elements. Finally, an easy partition of identity argument shows that if  $P(\cdot) \in C(X; M_n(\mathbb{C}))$  is positive, then  $P$  is approximable by elements of the form  $\sum_i f_i \otimes T_i$ , with  $f_i \in C(X)_+, T_i \in M_n(\mathbb{C})_+$ , so we may conclude that indeed  $\phi_n(P) \geq 0$ .  $\square$

The corresponding fact for CB maps is false: i.e., if  $u : \ell_n^\infty \rightarrow B(K)$  and  $u(e_j) = a_j$  (with  $\{e_j : 1 \leq j \leq n\}$  denoting the standard basis of  $\ell_n^\infty$ ), then  $\|u\| (= \sup\{\|\sum_k z_k a_j\| : |z_j| \leq 1, z_j \in \mathbb{C}\}) \neq \|u\|_{cb}$ —it turns out that

$$\|u\|_{cb} = \inf\left\{\left\|\sum_j x_j x_j^*\right\|^{1/2} \left\|\sum_j y_j^* y_j\right\|^{1/2} : a_j = x_j y_j, x_j, y_j \in B(K)\right\}.$$

**Proposition 1.2.11 (Schur Multipliers)**

1. Let  $\varphi_{ij} \in \mathbb{C}, 1 \leq i, j \leq n$  and consider  $u_\varphi : M_n \rightarrow M_n$  given by  $[a_{ij}] \mapsto [a_{ij}\varphi_{ij}]$ . Then,

$$\|u_\varphi\|_{cb} = \inf_i \left\{ \sup_j \|x_i\|_H \sup_j \|y_j\|_H : \varphi_{ij} = \langle x_i, y_j \rangle_H, x_i, y_j \in H \right\}$$

where the infimum runs over all such possible Hilbert spaces  $H$  and vectors  $x_i, y_j$ . Also,  $u_\varphi$  is CP if and only if  $[\varphi_{ij}] \geq 0$ .

2. Let  $G$  be a discrete group and  $C_\lambda^*(G)$  denote the norm-closure in  $B(\ell_2(G))$  of  $\{\lambda_g : g \in G\}$ , its “reduced group  $C^*$ -algebra” where  $\lambda : G \rightarrow B(\ell_2(G))$  is the left regular representation of  $G$  given by  $\lambda_g(\xi_h) = \xi_{gh}$ , where  $\{\xi_g : g \in G\}$  is the canonical orthonormal basis of  $\ell_2(G)$ . Let  $\varphi : G \rightarrow \mathbb{C}$  be any function. Consider  $T_\varphi : \text{span}\{\lambda_g : g \in G\} \rightarrow B(\ell_2(G))$  given by  $T_\varphi(\sum_g c_g \lambda_g) = \sum_g c_g \varphi(g) \lambda_g$ .

Then,

$$\begin{aligned} \|T_\varphi\|_{cb} &= \inf\{\|x\|_\infty\|y\|_\infty : x, y \in \ell^\infty(G, H) \text{ such that } \langle x(t), y(s) \rangle_H \\ &= \varphi(st^{-1}), \forall s, t \in G\} \end{aligned}$$

*Proof* (1) Suppose  $\phi_{ij} = \langle x_i, y_j \rangle_H, x_i, y_j \in H$ . If  $(a_{ij}) \in (M_n)$ , we can—by polar decomposition and diagonalisation of positive operators—find vectors  $u_i, v_j$  in some Hilbert space  $K$  such that  $a_{ij} = \langle u_i, v_j \rangle_K$  and  $\|(a_{ij})\| = (\max_i \|u_i\|)(\max_j \|v_j\|)$ . Then it follows that

$$\phi_{ij}a_{ij} = \langle x_i \otimes u_i, y_j \otimes v_j \rangle_{H \otimes K}$$

and hence,

$$\|(\phi_{ij}a_{ij})\| = (\max_i \|x_i\|)(\max_j \|y_j\|)(\max_i \|u_i\|)(\max_j \|v_j\|)$$

and we see that

$$\|u_\varphi\|_{cb} \leq \inf\{\sup_i \|x_i\|_H \sup_j \|y_j\|_H : \phi_{ij} = \langle x_i, y_j \rangle_H, x_i, y_j \in H\}$$

Conversely, by the fundamental factorisation Theorem 1.1.5, we may find a representation  $\pi : M_n \rightarrow B(K)$  and operators  $V, W \in B(\ell_n^2, K)$  for some Hilbert space  $K$ , with  $\|V\| \cdot \|W\| = \|u_\varphi\|_{cb}$ , such that

$$u_\varphi(x) = V^* \pi(x) W .$$

Then,

$$\begin{aligned} |\phi_{ij}| &= |\langle u_\varphi(e_{ij})e_j, e_i \rangle| \\ &= |\langle \pi(e_{ij})We_j, Ve_i \rangle| \\ &= |\langle \pi(e_{i1})\pi(e_{1j})We_j, Ve_i \rangle| \\ &= |\langle \pi(e_{1j})We_j, \pi(e_{i1})Ve_i \rangle| \\ &\leq \|u_\varphi\|_{cb} . \end{aligned}$$

(2) See Theorem 8.3 in [4]. □

It is true, although a bit non-trivial, that  $\|u_\varphi\|_{cb} = \|u_\varphi\|$ , while, however,  $\|T_\varphi\|_{cb} \neq \|T_\varphi\|$ .

## 1.3 Ruan's Theorem and Its Applications

### 1.3.1 Ruan's Theorem

Let us call a complex vector space  $E$  an **abstract operator space** if it comes equipped with a family of norms  $\|\cdot\|_n$  on  $M_n(E)$ ,  $n \geq 1$ , which satisfy the properties:

1.  $\left\| \begin{bmatrix} x & 0 \\ 0 & y \end{bmatrix} \right\|_{m+n} = \max(\|x\|_n, \|y\|_m)$  for all  $x \in M_n(E)$ ,  $y \in M_m(E)$ ;
2.  $\|axb\|_n \leq \|a\|_{M_n \times m} \|x\|_m \|b\|_{M_m \times n}$  for all  $a, b^* \in M_n \times m$ ,  $x \in M_m(E)$ .

In contrast, any vector subspace of a  $B(H)$  is called a **concrete operator space** (with respect to the norms  $\|\cdot\|_n = \|\cdot\|_{M_n(B(H))}$ ).

The equivalence of these two notions is the content of **Ruan's theorem** established in his PhD Thesis [5].

**Theorem 1.3.1 (Ruan)** *Any abstract operator space is completely isometrically isomorphic to a concrete operator space; i.e., for any abstract operator space  $(E, \{\|\cdot\|_n : n \geq 1\})$ , there is a Hilbert space  $H$  and a linear map  $J : E \rightarrow B(H)$  such that, for each  $n \geq 1$  and  $x = [x_{ij}] \in M_n(E)$ ,  $\|x\|_n = \|[J(x_{ij})]\|_{M_n(B(H))}$ .*

The proof of Ruan's theorem bears a strong similarity to the proof of Gelfand Naimark's Theorem that an abstract  $C^*$ -algebra is isomorphic to a norm-closed (concrete)  $*$ -subalgebra of  $B(H)$  for some Hilbert space  $H$ .

*Proof ((Ruan's) Theorem 1.3.1)* The proof of this theorem mainly depends on proving the following:

**Claim** For each  $n \geq 1$  and  $x \in M_n(E)$  with  $\|x\|_n = 1$ , there is a Hilbert space  $H_{n,x}$  and an operator  $J_{n,x} : E \rightarrow B(H_{n,x})$  satisfying

1.  $\|[J_{n,x}(x_{ij})]\|_n = 1$
2.  $\forall m, \forall y = [y_{ij}] \in M_m(E)$ ,  $\|[J_{n,x}(y_{ij})]\|_m \leq \|y\|_m$ .

Indeed, once we have the above, taking  $H := \bigoplus_{n,x} H_{n,x}$  and  $J := \bigoplus_{n,x} J_{n,x} : E \rightarrow B(H)$  we observe that  $J$  satisfies the properties stated in the theorem.

To prove the claim, fix  $n \geq 1$  and an  $x \in M_n(E)$  with  $\|x\|_n = 1$ . Observe first that  $(M_k(E), \|\cdot\|_k)$  embeds (as the north-east corner, with 0's on the  $(n+1)$ th row and column) isometrically in  $(M_{k+1}(E), \|\cdot\|_{k+1})$  for all  $k \geq 1$  (thanks to condition (2) in the requirements for an abstract operator space). Thus, there exists a  $\varphi \in (\bigcup_{k \geq 1} M_k(E))^*$  satisfying  $\|\varphi\| = 1$  and  $\varphi(x) = 1$ . Then,  $\forall y \in M_m(E)$  with  $\|y\|_m \leq 1$  and  $\forall a, b \in M_{m \times n}$ ,  $n \geq 1$ , we have  $|\varphi(a^*yb)| \leq \|a\|_{M_{m \times n}} \|b\|_{M_{m \times n}}$ .

We see, in particular that for each  $m, N, n \geq 1$ , we have  $\forall a_j, b_j \in M_{m \times n}$ ,  $\forall y_j \in M_m(E)$  with  $\|y_j\|_m \leq 1$ ,  $1 \leq j \leq N$ , we have

$$\left| \sum_j \varphi(a_j^* y_j b_j) \right| \leq \left\| \sum_j a_j^* a_j \right\|^{1/2} \left\| \sum_j b_j^* b_j \right\|^{1/2}. \quad (1.2)$$

This is true because taking  $a = \begin{bmatrix} a_1 \\ \vdots \\ a_N \end{bmatrix}$ ,  $y$  to be the diagonal matrix

$y = y_1 \oplus \cdots \oplus y_N \in M_N(M_m(E))$  and  $b = \begin{bmatrix} b_1 \\ \vdots \\ b_N \end{bmatrix}$ , we clearly have

$|\varphi(a^*yb)| \leq \|a\|_{M_mN \times n} \|y\|_N \|b\|_{M_mN \times n}$ , which is precisely the above inequality since  $\| [b_1, \dots, b_N]^T \|_{M_mN \times n} = \| \sum_j b_j^* b_j \|^{1/2}$ .

In fact, by some suitable convexity and Hahn Banach separation arguments, the inequality in (1.2) can be improved (see [5]) to: There exist states  $f_1, f_2$  on  $B(\ell^2)$  such that

$$\left| \sum_j \varphi(a_j^* y_j b_j) \right| \leq \left( \sum_j f_1(a_j^* a_j) \right)^{1/2} \left( \sum_j f_2(b_j^* b_j) \right)^{1/2} \quad (1.3)$$

for all  $N, a_j, b_j, y_j$  as above.

In particular,  $\forall a, b \in M_{1 \times n}, e \in E$  with  $\|e\| \leq 1$ , we have

$$|\varphi(a^*[e]b)| \leq f_1(a^*a)^{1/2} f_2(b^*b)^{1/2} = \|\hat{a}\|_{H_1} \|\hat{b}\|_{H_2},$$

where  $H_i := L^2(B(\ell^2), f_i), i = 1, 2$  and  $\hat{a}$  represents  $a$  as an element of  $H_1$  and likewise  $\hat{b}$ .

Consider  $v : E \rightarrow B(H_2, H_1)$  given by  $\langle v(e)(\hat{b}), \hat{a} \rangle_{H_1} = \varphi(a^*[e]b)$ . Clearly  $\|v(e)\| \leq \|e\|$  for all  $e \in E$  and hence  $\|v\| \leq 1$ . Also, for  $y = [y_{ij}] \in M_m(E)$ ,

$$\begin{aligned} \|v_m(y)\| &= \sup \left\{ \left| \sum_j \langle v(y_{ij}) \hat{b}_j, \hat{a}_j \rangle \right| : \sum_j \|\hat{b}_j\|^2 \leq 1, \sum_j \|\hat{a}_j\|^2 \leq 1 \right\} \\ &= \sup \left\{ \left| \varphi \left( [a_1^*, \dots, a_m^*] y \begin{bmatrix} b_1 \\ \dots \\ b_m \end{bmatrix} \right) \right| : \sum_j \|\hat{b}_j\|^2 \leq 1, \sum_j \|\hat{a}_j\|^2 \leq 1 \right\} \\ &\leq \|y\|_m. \end{aligned}$$

Let  $H_{n,x} = H_1 \oplus H_2$  and  $J_{n,x} : E \rightarrow B(H_{n,x})$  be given by  $J_{n,x}(e) = \begin{bmatrix} 0 & v(e) \\ 0 & 0 \end{bmatrix}$ .

We find that

$$1 = \varphi(x) = \varphi([x_{ij}]) = \sum_{ij} \langle v(x_{ij}) \widehat{e}_{1j}, \widehat{e}_{1i} \rangle = \langle v_n(x) \begin{bmatrix} \widehat{e}_{11} \\ \dots \\ \widehat{e}_{1n} \end{bmatrix}, \begin{bmatrix} \widehat{e}_{11} \\ \dots \\ \widehat{e}_{1n} \end{bmatrix} \rangle,$$



where  $\{e_{ij}\}$  is the system of matrix units in  $M_n$ . Since  $\sum_j \|\widehat{e_{1j}}\|_{H_1}^2 = \sum_j f_i(e_{jj}) \leq f_i(1) = 1$ , we find that

$$1 = |\varphi(x)| \leq \|v_n(x)\| \cdot \left( \sum_j \|\widehat{e_{1j}}\|_{H_1}^2 \right)^{1/2} \cdot \left( \sum_j \|\widehat{e_{1j}}\|_{H_2}^2 \right)^{1/2} \leq \|v_n(x)\|.$$

Thus,  $\|[J_{n,x}(x_{ij})]\| = 1$  and the proof is complete. □

### 1.3.2 Some Applications and Some Basic Facts

- *CB(E, F)*. Let  $E$  and  $F$  be operator spaces and  $G := CB(E, F)$ . For  $[x_{ij}] = x \in M_n(G)$ , let  $\|x\|_n := \|\tilde{x} : E \rightarrow M_n(F)\|_{CB(E, M_n(F))}$ , where  $\tilde{x}(e) := [x_{ij}(e)]$ ,  $e \in E$ . Routine checking shows that the above sequence of norms satisfies Ruan's axioms and, hence, by Theorem 1.3.1,  $CB(E, F)$  admits an operator space structure.
- *Dual operator space*. In particular, the dual space  $E^* = CB(E, \mathbb{C}) = B(E, \mathbb{C})$  also inherits an operator space structure. We now see that this operator space structure on  $E^*$  is the appropriate one, in the sense that many properties of Banach dual space structure carry over to this theory.
- *Adjoint operator*. For  $u \in CB(E, F)$ , its usual adjoint  $u^*$  is also a CB map from  $F^* \rightarrow E^*$  with  $\|u^*\|_{cb} = \|u\|_{cb}$ .
- *Quotient spaces*. Let  $E_2 \subset E_1 \subset B(H)$  be operator spaces. For  $x = [x_{ij}] \in M_n(E_1/E_2)$ , set  $\|x\|_n = \|q_n(\hat{x})\|_{M_n(E_1)/M_n(E_2)}$ , where  $x_{ij} = q\hat{x}_{ij}$  for  $\hat{x}_{ij} \in E_1$  and  $\hat{x} = [\hat{x}_{ij}]$  and  $q : E_1 \rightarrow E_1/E_2$  is the canonical quotient map.

The above sequence of norms are seen to be well-defined and to satisfy Ruan's axioms and thus equip the quotient space  $E_1/E_2$  with the structure of an operator space.

Note there is also an obvious operator space structure on  $E_1/E_2$  via the embedding  $E_1/E_2 \subset B(H)/E_2$ ; but this is not consistent with other properties.

#### Analogy with Banach Space Properties

- With  $E_1$  and  $E_2$  as above,

$$(E_1/E_2)^* \simeq E_2^\perp \text{ and } E_2^* \simeq E_1^*/E_2^\perp$$

as operator spaces, where of course  $E_2^\perp \subset E_1^*$  is defined as  $E_2^\perp = \{f \in E_1^* : E_2 \subset \ker(f)\}$ .

- The row and column operator spaces are dual to each other, i.e.,  $R^* \simeq C$  and  $C^* \simeq R$  as operator spaces.
- The canonical embedding  $E \subset E^{**}$  is a complete isometry—see [2].

### 1.3.3 *min and max Operator Space Structures on a Banach Space*

Given a Banach space  $X$ , two operator space structures attract special attention, which are described by the adjectives *max* and *min*. These structures are characterised as follows.

**Proposition 1.3.2** *There exist operator space structures  $\max(X)$  and  $\min(X)$  on a Banach space  $X$  such that  $\max(X)$  and  $\min(X)$  are isometric to  $X$  and for every operator space  $Z$ ,*

1.  $CB(\max(X), Z) = B(X, Z)$ .
2.  $CB(Z, \min(X)) = B(Z, X)$ .

Moreover, (1) and (2) characterise  $\max(X)$  and  $\min(X)$ , respectively.

*Proof* We discuss only the existence of  $\max(X)$  and its property (1). For  $x \in M_n(X)$ , set

$$\|x\|_{M_n(\max(X))} = \inf\{\|a\|_{M_n \times N} \sup_{j \leq N} \|D_j\|_X \|b\|_{M_N \times n} : a^*, b \in M_{N \times n}, \\ D_j \in X, N \geq 1\}$$

where the infimum runs over all possible decompositions  $x = a \begin{bmatrix} D_1 & 0 & \cdots & 0 \\ 0 & D_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & D_N \end{bmatrix} b$ .

This sequence of norms satisfies Ruan's axioms and hence we have an operator space structure  $\max(X)$  on  $X$ . Now, let  $u \in B(X, Z)$  for an operator space  $Z$  and  $x \in M_n(\max(X))$ . Assume that  $\|x\|_{M_n(\max(X))} < 1$ . Then there exists  $N \geq 1$ ,  $a^*, b \in M_{N \times n}$ ,  $D_j \in X$ ,  $1 \leq j \leq N$  such that  $x =$

$$a \begin{bmatrix} D_1 & 0 & \cdots & 0 \\ 0 & D_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & D_N \end{bmatrix} b \text{ and } \|a\|_{M_n \times N} \sup_{j \leq N} \|D_j\|_X \|b\|_{M_N \times n} < 1. \text{ Then note that}$$

$$u_n(x) = a \begin{bmatrix} u(D_1) & 0 & \cdots & 0 \\ 0 & (D_2) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & u(D_N) \end{bmatrix} b - \text{ and that}$$

$$\|u_n(x)\|_{M_n(\max(X))} \leq \|u\| \|a\|_{M_n \times N} \sup_{j \leq N} \|D_j\|_X \|b\|_{M_N \times n} \leq \|u\|.$$

□

- $(\min(X))^* \simeq \max(X^*)$  and  $(\max(X))^* \simeq \min(X^*)$  as operator spaces—see [1, 2].

## 1.4 Tensor Products of Operator Spaces

We will be mainly interested in the injective and projective tensor products of operator spaces.

### 1.4.1 Injective Tensor Product

Let  $E_i \subset B(H_i)$ ,  $i = 1, 2$  be operator spaces. Then we have a natural embedding  $E_1 \otimes E_2 \subset B(\overline{H_1 \otimes H_2})$  and the *minimal* tensor product of  $E_1$  and  $E_2$  is the space  $E_1 \otimes_{\min} E_2 := \overline{E_1 \otimes E_2} \subset B(H_1 \otimes H_2)$ .

The min tensor product is independent of the embeddings  $E_i \subset B(H_i)$ ,  $i = 1, 2$ ; it depends only upon the sequence of norms on  $E_i$ ,  $i = 1, 2$ . We see this as follows: For simplicity, assume  $H_1 = H_2 = \ell^2$  and let  $t = \sum_{k=1}^r a_k \otimes b_k \in E_1 \otimes E_2$ . Consider the natural embeddings  $\ell_n^2 \subset \ell^2$ ,  $n \geq 1$ —with respect to some choice  $\{\xi_n : n \geq 1\}$  of orthonormal basis for  $\ell^2$  and the corresponding orthogonal projections  $P_n : H_1 \otimes H_2 \rightarrow \ell_n^2 \otimes H_2$ . Then  $\bigcup_n \ell_n^2 \otimes H_2 = H_1 \otimes H_2$  and so  $\|x\| = \sup_n \|P_n x P_n\| \forall x \in B(H_1 \otimes H_2)$ ; and hence,

$$\|t\|_{\min} = \sup_n \|P_n t|_{\ell_n^2 \otimes H_2}\|_{B(\ell_n^2 \otimes H_2)}.$$

Suppose  $\langle a_k \xi_j, \xi_i \rangle = a_k(i, j)$ . Then it is not hard to see that  $P_n t|_{\ell_n^2 \otimes H_2}$  may be identified with the matrix  $t_n \in M_n(E_2)$  given by  $t_n(i, j) = \sum_k (a_k)_{ij} b_k$ . This shows that  $\|t\|_{\min} = \sup_n \|t_n\|_{M_n(E_2)}$ , and hence that the operator space structure of  $E_1 \otimes_{\min} E_2$  depends only on the operator space structure of  $E_2$  and not on the embedding  $E_2 \subset B(H_2)$ . In an entirely similar manner, it can be seen that the operator space structure of  $E_1 \otimes_{\min} E_2$  depends only on the operator space structure of  $E_1$  and not on the embedding  $E_1 \subset B(H_1)$ .

More generally, the same holds for  $\|[t_{ij}]\|_{M_n(B(H_1 \otimes H_2))}$  and hence  $\|\cdot\|_{M_n(E_1 \otimes E_2)}$  depends only on the operator space structures of the  $E_j$ 's and not on the embeddings  $E_j \subset B(H_j)$ .

**Proposition 1.4.1** *Let  $u : E \rightarrow F$  be a CB map between two operator spaces and  $G$  be any operator space. Let  $u_G : G \otimes_{\min} E \rightarrow G \otimes_{\min} F$  be the extension of the map  $id_G \otimes u$ . Then  $\|u\|_{cb} = \sup_G \|u_G\|$ .*

*Proof* For this, first observe that if  $G = M_n$ , then  $M_n = B(\ell_n^2) \Rightarrow M_n \otimes_{\min} E = M_n(E)$  and that  $\|u\|_{M_n} = \|u_n\|$ . (In fact the identification  $M_n \otimes_{\min} E = M_n(E)$  is not only isometric, it is even completely isometric.) It follows—by allowing  $G$  to

range over  $\{M_n : n \in \mathbb{N}\}$ —that

$$\sup\{\|u_G\| : G \text{ an operator space}\} \geq \|u\|_{cb}.$$

The same argument that yielded the conclusion  $M_n \otimes_{\min} E = M_n(E)$  is also seen to show that if  $E \subset B(K)$ , then  $B(H) \otimes_{\min} E = \overline{B(H) \otimes E} \subset B(H \otimes K)$  (as a Banach space); assuming, as before, that  $H = \ell^2$ , and choosing the projections  $P_n = P_{\ell_n^2} \otimes K$ , we may easily deduce from the strong convergence of  $P_n$  to  $\mathbb{I}_{H \otimes K}$  that  $\|u_{B(H)}\| = \|u\|_{cb}$ . Finally, we may conclude that if  $G \subset B(H)$ , then,  $G \otimes_{\min} E$  sits isometrically as a subspace of  $B(H) \otimes_{\min} E$ , and that

$$\begin{aligned} \|u_G\| &= \|u_{B(H)}|_{G \otimes_{\min} E}\| \\ &\leq \|u_{B(H)}\| \\ &= \|u\|_{cb}, \end{aligned}$$

thereby establishing the desired equality,  $G$  was arbitrary.  $\square$

*Remark 1.4.2* 1. For two operator spaces  $E$  and  $F$  with  $F$  being finite dimensional, we have  $CB(E, F) \simeq E^* \otimes_{\min} F$  as operator spaces. This is analogous to the fact that  $B(E, F) \simeq E^* \otimes^{\vee} F$  isometrically for two Banach spaces  $E$  and  $F$ .

2. min tensor product is an associative and commutative tensor product.
3. For operator spaces  $E_i, F_i, i = 1, 2$  and  $u_j \in CB(E_j, F_j)$ , the map  $u_1 \otimes u_2$  extends to a CB map  $u_1 \otimes u_2 : E_1 \otimes_{\min} E_2 \rightarrow F_1 \otimes_{\min} F_2$  such that  $\|u_1 \otimes u_2\|_{cb} = \|u_1\|_{cb} \|u_2\|_{cb}$ . Indeed, one easily shows that the composition of the extensions  $id_{E_1} \otimes u_2 \in CB(E_1 \otimes_{\min} E_2, E_1 \otimes_{\min} F_2)$  and  $u_1 \otimes id_{F_2} \in CB(E_1 \otimes_{\min} F_2, E_2 \otimes_{\min} F_2)$  give us a CB map  $u_1 \otimes u_2 = (u_1 \otimes id_{F_2})(id_{E_1} \otimes u_2)$  with required properties.
4. Another nice consequence of Proposition 1.4.1 and above observation about tensor products of CB maps is that the min tensor product is *injective*, i.e., if  $E_2 \subset E_1$  is an inclusion of operator spaces and  $G$  is any operator space, then the inclusion  $E_2 \otimes_{\min} G \subset E_1 \otimes_{\min} G$  is a completely isometry.

## 1.4.2 Projective Tensor Product

Let  $E_i \subset B(H_i)$  be operator spaces. For  $t = [t_{ij}] \in M_n(E_1 \otimes E_2)$ , set

$$\begin{aligned} \|t\|_n &= \inf\{\|a\|_{M_{n \times N^2}} \|x^1\|_{M_N(E_1)} \|x^2\|_{M_N(E_2)} \|b\|_{M_{N^2 \times n}} : a^*, b \in M_{N^2 \times n}, \\ &\quad x^i \in M_n(E_i), i = 1, 2\} \end{aligned}$$

where the infimum runs over all possible decompositions of the form  $t = a(x^1 \otimes x^2)b$ . This sequence of norms satisfies Ruan's axioms and we obtain an operator space structure on  $E_1 \otimes E_2$ , which after completion is denoted  $\widehat{E_1 \otimes E_2}$  and is called the *projective tensor product* of the operator spaces  $E_1$  and  $E_2$ .

- For operator spaces  $E_1$  and  $E_2$ , we have operator space isomorphisms

$$(E_1 \widehat{\otimes} E_2)^* \simeq CB(E_1, E_2^*) \simeq CB(E_2, E_1^*).$$

*Remark 1.4.3* 1. Projective tensor product of operator spaces is not injective.  
2. Injective tensor product of operator spaces is not projective, in the sense defined below.

**Definition 1.4.4** Let  $E$  and  $F$  be operator spaces and  $u \in CB(E, F)$ . Then  $u$  induces a canonical map  $\tilde{u} : E/\ker u \rightarrow F$ . Then  $u$  is said to be a **complete metric surjection** if  $u$  is surjective and  $\tilde{u}$  is completely isometric.

**Exercise 1.4.5**  $u \in CB(E, F)$  is a complete metric surjection if and only if  $u^* : F^* \rightarrow E^*$  is a completely isometric embedding.

- $\widehat{\otimes}$  is “projective” in the following sense. For any two complete metric surjections  $u_j : E_j \rightarrow F_j$ ,  $j = 1, 2$ , the tensor map  $u_1 \otimes u_2 : E_1 \widehat{\otimes} E_2 \rightarrow F_1 \widehat{\otimes} F_2$  is also a complete metric surjection.

Note that for any two Hilbert spaces  $H$  and  $K$ ,  $B(H, K)$  has a canonical operator space structure via the embedding  $B(H, K) \ni x \mapsto \begin{pmatrix} 0 & 0 \\ x & 0 \end{pmatrix} \in B(H \oplus K)$ .

- Thus, given any Hilbert space  $H$ , we have two new operator spaces, namely,

$$H_c := B(\mathbb{C}, H) \text{ and } H_r := B(\overline{H}, \mathbb{C}).$$

- $(H_c)^* \simeq \overline{H}_r$  as operator spaces.
- If  $H = \ell_2$ , then  $H_c = C$  and  $H_r = R$ .
- For any two Hilbert spaces  $H$  and  $K$ , we have the following operator space isomorphisms:

$$(\overline{K} \widehat{\otimes} H)^* \simeq B(H, K) \simeq CB(H_c, K_c) \simeq (K_c^* \widehat{\otimes} H_c)^*$$

where  $\widehat{\otimes}$  is the Banach space projective tensor product. Note that  $\overline{K} \widehat{\otimes} H$  as well as  $K_c^* \widehat{\otimes} H_c$  are isometrically isomorphic to the space  $S_1(H, K)$  of trace class operators.

### 1.4.3 General Remarks

- Recall the  $\gamma$ -norm in the proof of Fundamental Factorization Theorem. In fact, we have an isometric isomorphism

$$(\overline{K} \otimes E \otimes H)_\gamma \simeq S_1(K, H) \widehat{\otimes} E.$$

- For any measure space  $(X, \mu)$  (unlike in the next remark) the Banach space projective tensor product is injective, at least, if one of the tensor factors is  $L^1(X, \mu)$  and the other is an operator space. This is true because for any inclusion of operator spaces  $E \subset E_1$ , we have

$$L^1(X, \mu) \hat{\otimes} E = L^1(\mu, E) \subset L^1(\mu, E_1) = L^1(X, \mu) \hat{\otimes} E_1.$$

- For a von Neumann algebra  $M$ , its predual  $M_*$ , usually called a *non-commutative  $L^1$ -space*, has a natural operator space structure via the embedding  $M_* \subset M^*$  ( $= (M_*)^{**}$ ). For an inclusion of operator spaces  $E \subset E_1$ , only when  $M$  is injective, there is an isometric embedding  $M_* \widehat{\otimes} E \subset M_* \widehat{\otimes} E_1$ .

### 1.4.4 A Passing Remark on the Haagerup Tensor Product

Apart from the above two tensor products of operator spaces, there is another extremely important tensor product of operator spaces, namely, the *Haagerup* tensor product, usually denoted  $E \otimes_h F$ . Unlike the above two tensor products, the Haagerup tensor product has no analogy in Banach space theory. It is known that the Haagerup tensor product of operator spaces is associative, injective and projective but it is not commutative. Also,  $\overline{K}_r \otimes_h E \otimes_h H_c \simeq (\overline{K} \otimes E \otimes H)_\gamma$  as Banach spaces. We will not get into the details in these notes.

## 1.5 Tensor Products of $C^*$ -Algebras

### 1.5.1 *min and max* Tensor Products of $C^*$ -Algebras

Let  $A$  and  $B$  be unital  $C^*$ -algebras. Then  $A \otimes_{\min} B$  is canonically a  $C^*$ -algebra. For each  $t = \sum_k a_k \otimes b_k \in A \otimes B$ , define

$$\|t\|_{\max} = \sup\{\|\sum_k \pi(a_k)\sigma(b_k)\| : \pi : A \rightarrow B(H), \sigma : B \rightarrow B(H)\}$$

where the infimum runs over all representations  $\pi$  and  $\sigma$  with commuting ranges. Then the completion  $A \otimes_{\max} B$  of  $A \otimes B$  with respect to this norm makes it into a  $C^*$ -algebra. In general, we have  $\|t\|_{\min} \leq \|t\|_{\max}$  for all  $t \in A \otimes B$ .

**Definition 1.5.1** For  $C^*$ -algebras  $A$  and  $B$ , the pair  $(A, B)$  is called a **nuclear pair** if  $\|t\|_{\min} = \|t\|_{\max}$  for all  $t \in A \otimes B$ .

**Definition 1.5.2** A  $C^*$ -algebra  $A$  is said to be **nuclear** if  $(A, B)$  is a nuclear pair for all  $C^*$ -algebras  $B$ .

### 1.5.2 Kirchberg's Theorem

**Theorem 1.5.3 ([3])** *Let  $1 \leq n \leq \infty$ . If  $A = A_n = C^*(\mathbb{F}_n)$ , the full  $C^*$ -algebra on the free group with  $n$ -generators and  $B = B(H)$ , then  $(A, B)$  is a nuclear pair.*

The above result has great significance as both  $C^*(\mathbb{F}_n)$  and  $B(H)$  are universal objects in the sense that every  $C^*$ -algebra is a quotient of  $C^*(\mathbb{F}_n)$  and a  $*$ -subalgebra of  $B(H)$  for suitable choices of  $n$  and  $H$ .

A proof of this remarkable theorem involves a fair bit of analysis.

**Lemma 1.5.4** *Let  $A \subset B(H)$  be a  $C^*$ -algebra and  $\{U_i : i \in I\} \subset \mathcal{U}(H)$  be a family of unitary operators on  $H$  such that  $1 \in E := \overline{\text{span}}\{U_i : i \in I\}$  and  $A = C^*(E)$ . Let  $u : E \rightarrow B(K)$  be a unital c.b map satisfying  $u(U_i) \in \mathcal{U}(K)$  for all  $i \in I$ . If  $\|u\|_{cb} \leq 1$ , then there exists a  $*$ -representation  $\sigma : A \rightarrow B(K)$  such that  $\sigma|_E = u$ .*

*Proof* Since  $u$  is unital and CB with  $\|u\|_{cb} \leq 1$ , there is a Hilbert space  $\hat{H}$ , a  $*$ -representation  $\pi : A \rightarrow B(\hat{H})$ , a subspace  $K \subset \hat{H}$  such that  $u(a) = P_K \pi(a)|_K$  for all  $a \in E$ , where  $P_K : \hat{H} \rightarrow K$  is the orthogonal projection.

Suppose  $\pi(U_i) = \begin{pmatrix} a_i & b_i \\ c_i & d_i \end{pmatrix}$  with respect to  $\hat{H} = K \oplus K^\perp$ , so that  $u(U_i) = a_i$ . Unitarity of  $U_i$  implies that  $a_i^* a_i + b_i^* b_i = a_i a_i^* + c_i c_i^* = 1$ ; while the assumed unitarity of  $a_i$  is seen to imply that  $b_i = c_i = 0$ . Therefore,  $\pi(U_i) = \begin{pmatrix} a_i & 0 \\ 0 & d_i \end{pmatrix}$ , and we see that  $K$  is an invariant subspace for  $\pi(U_i)$  as well as for  $\pi(U_i^*)$  for all  $i \in I$ . Thus,  $K$  is invariant under  $C^*(\{\pi(U_i) : i \in I\})$ . Finally  $\sigma : A \rightarrow B(K)$  defined by  $\sigma(a) = \pi(a)|_K$ ,  $a \in A$  is seen to be a  $*$ -representation which extends  $u$ .  $\square$

**Lemma 1.5.5** *Let  $U_0 = 1, U_1, \dots, U_n$  be unitary generators of  $A_n := C^*(\mathbb{F}_n)$ . If  $\{a_j : 0 \leq j \leq n\} \subset B(H)$  is such that  $\alpha := \|\sum_0^n U_j \otimes a_j\|_{A_n \otimes \min B(H)} \leq 1$ , then there exist  $b_j, c_j \in B(H)$  such that  $a_j = b_j^* c_j$  and  $\|\sum_j b_j^* b_j\| \leq 1$ ,  $\|\sum_j c_j^* c_j\| \leq 1$ .*

*Proof* Consider  $u : \ell_{n+1}^\infty \rightarrow B(H)$  given by  $u(e_j) = a_j$ ,  $0 \leq j \leq n$ , where  $\{e_j : 0 \leq j \leq n\}$  is the standard basis of  $\ell_{n+1}^\infty$ .

**Assertion:**  $u$  is CB with  $\|u\|_{cb} \leq \alpha$ .

Under the identification  $M_m(A) \ni [a_{ij}] \mapsto \sum_{ij} e_{ij} \otimes a_{ij} \in M_m \otimes A$  for any algebra  $A$ , we see that  $u_m : M_m(\ell_{n+1}^\infty) \rightarrow M_m(B(H))$  is given by  $u_m(\sum_{j=0}^n x_j \otimes e_j) = \sum_{j=0}^n x_j \otimes a_j$  for  $x_j \in M_m$ .

It follows that

$$\begin{aligned} \|u_m\| &= \sup\{\|\sum_{j=0}^n x_j \otimes a_j\|_{M_m \otimes \min B(H)} : \max_j \|x_j\|_{M_m} \leq 1\} \\ &\leq \sup\{\|\sum_{j=0}^n V_j \otimes a_j\| : V_j \in M_m \text{ is unitary}\}, \end{aligned}$$

where the first equality is because  $\|\sum_j x_j \otimes e_j\|_{M_m(\ell_\infty^{n+1})} = \max_j \|x_j\|_{M_m}$ , and the second equality is the consequence of the fact that the extreme points of the unit ball of  $M_m(\mathbb{C})$  are precisely the unitary matrices.

By definition of the full free group  $C^*$ -algebra, there exists a  $*$ -homomorphism  $\sigma : A_n \rightarrow M_m$  such that  $\sigma(U_j) = V_j, \forall 0 \leq j \leq n$ . Since any  $*$ -homomorphism is a complete contraction, it follows from Remark 1.4.2 (3) that  $\sigma \otimes id_{B(H)} : A_n \otimes_{\min} B(H) \rightarrow M_m \otimes_{\min} B(H)$  is also a complete contraction, and consequently, we find that indeed  $\|u\|_{cb} = \sup \|u_m\| \leq \alpha \leq 1$ . Thus  $u$  is a complete contraction, and we may deduce from the fundamental factorisation theorem that there exists a representation  $\pi : \ell_{n+1}^\infty \rightarrow B(K)$  and contractions  $V, W : H \rightarrow K$  such that  $(a_j =)u(e_j) = V^*\pi(e_j)W \forall j$ . Pick some isometry  $S : K \rightarrow H$  and define  $b_j = S\pi(e_j)V, c_j = S\pi(e_j)W$ .

Then, notice that indeed  $b_j^*c_j = V^*\pi(e_j)S^*S\pi(e_j)W = a_j$ , and as  $V$  and  $W$  are contractions, we find that

$$\sum_j b_j^*b_j = V^* \left( \sum_j \pi(e_j) \right) V = V^*V \leq 1$$

and

$$\sum_j c_j c_j^* = \sum_j S\pi(e_j)W^*W\pi(e_j)S^* \leq \sum_j S\pi(e_j)S^* = SS^* \leq 1,$$

thereby completing the proof of the Lemma.  $\square$

We shall need the following version of the Cauchy-Schwarz inequality. We say nothing about its proof but it is easy and may be found, for instance, in any treatment of Hilbert  $C^*$ -algebras.

**Lemma 1.5.6** *For elements  $b_j, c_j$  of any  $C^*$ -algebra, we have*

$$\left\| \sum_j b_j^*c_j \right\|^2 \leq \left\| \sum_j b_j^*b_j \right\| \cdot \left\| \sum_j c_j^*c_j \right\|$$

*Proof of Theorem 1.5.3* We have  $A = C^*(\mathbb{F}_n)$  and  $B = B(H)$ . We need to show  $\|A \otimes_{\min} B \rightarrow A \otimes_{\max} B\| \leq 1$ . Let  $E = \text{span}\{U_j \otimes B(H) : 0 \leq j \leq n\} = \text{span}\{U_j \otimes V : V \in \mathcal{U}(H)\} \subset A \otimes_{\min} B$ . Note that  $1 \otimes 1 \in E$ . Consider  $u = id_E : E \rightarrow E \subset A \otimes_{\max} B$ .

We shall be done if we show that  $\|u\| \leq 1$ . Suppose  $t = \sum_j U_j \otimes a_j \in E$  with  $\|t\|_{\min} \leq 1$ . Then, pick  $b_j, c_j \in B(H)$  as in Lemma 1.5.5 such that  $a_j = b_j^*c_j$ . For any two representations  $\pi : A \rightarrow B(K)$  and  $\sigma : B \rightarrow B(K)$  with commuting ranges, consider  $\pi \cdot \sigma : E \rightarrow B(K)$  given by  $(\pi \cdot \sigma)(t) = \sum_j \pi(U_j)\sigma(a_j) = \sum_j \sigma(b_j^*)\pi(U_j)\sigma(c_j)$ .



Then, by Lemma 1.5.6, we find that  $\|(\pi \cdot \sigma)(t)\| \leq \|\sum_j \sigma(b_j^*)\sigma(b_j)\|^{1/2} \|\sum_j \sigma(c_j^*)\sigma(c_j)\|^{1/2} \leq 1$ . As  $\pi$  and  $\sigma$  were arbitrary, we find thus that  $\|t\|_{\max} \leq 1$ . So, we indeed have  $\|u\| \leq 1$ , thus completing the proof of Kirchberg's theorem.  $\square$

## References

1. D. Blecher, Tensor products of operator spaces II. *Can. J. Math.* **44**, 75–90 (1992)
2. D. Blecher, V. Paulsen, Tensor products of operator spaces. *J. Funct. Anal.* **99**, 262–292 (1991)
3. E. Kirchberg, On non-semisplit extensions, tensor products and exactness of group  $C^*$ -algebras. *Invent. Math.* **112**, 449–489 (1993)
4. G. Pisier, *Introduction to Operator Space Theory*. LMS Lecture Note Series, vol. 294 (Cambridge University Press, Cambridge, 2003)
5. Z.-J. Ruan, PhD Thesis, UCLA (1987)

# Chapter 2

## Entanglement in Bipartite Quantum States

**Abstract** Entanglement is a central feature of quantum mechanical systems that lies at the heart of most quantum information processing tasks. In this chapter, based on lectures by K.R. Parthasarathy, we examine the important question of quantifying entanglement in bipartite quantum systems. We begin with a brief review of the mathematical framework of quantum states and probabilities and formally define the notion of entanglement in composite Hilbert spaces via the Schmidt number and Schmidt rank. For pure bipartite states in a finite-dimensional Hilbert space, we study the question of how many subspaces exist in which all unit vectors are of a certain minimal Schmidt number  $k$ . For mixed states, we describe the role of  $k$ -positive maps in identifying states whose Schmidt rank exceeds  $k$ . Finally, we compute the Schmidt rank of an interesting class of states called generalized Werner states.

### 2.1 Quantum States, Observables and Probabilities

We begin with a few introductory remarks on notation. We consider finite-level quantum systems labeled by  $A, B, C$  etcetera. Let  $\mathcal{H}_A, \mathcal{H}_B, \mathcal{H}_C$ , denote the finite-dimensional Hilbert spaces associated with them. The elements of  $\mathcal{H}$  are called *ket vectors* denoted as  $|u\rangle, |v\rangle$ , while the elements of the dual  $\mathcal{H}^*$  are called *bra vectors* denoted by  $\langle u|, \langle v|$ . The *bra-ket*  $\langle u|v\rangle$  denotes the sesquilinear form, which is linear in  $|v\rangle$  and anti-linear in  $|u\rangle$ . If  $\mathcal{H}$  is a  $n$ -dimensional Hilbert space,  $|u\rangle$  can be represented as a column vector and  $\langle u|$  as a row vector, with complex entries.

$$|u\rangle = \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{bmatrix}; \quad \langle u| = [\bar{z}_1 \ \bar{z}_2 \ \dots \ \bar{z}_n]$$

The bra vector  $\langle u| \equiv |u\rangle^\dagger$  is thus the *adjoint* of the ket vector. The adjoint of any operator  $X$  is denoted as  $X^\dagger$ . Also note that for vectors  $|u\rangle, |v\rangle$  and operator  $X$ ,  $\langle v|Xu\rangle = \langle v|X|u\rangle = \langle X^\dagger v|u\rangle$ . If  $|u\rangle \in \mathcal{H}_1$  and  $|v\rangle \in \mathcal{H}_2$  are vectors in two

different Hilbert spaces  $\mathcal{H}_1$  and  $\mathcal{H}_2$  respectively,  $|u\rangle\langle v|$  is an operator from  $\mathcal{H}_2 \rightarrow \mathcal{H}_1$ . That is, for any  $|w\rangle \in \mathcal{H}_2$ ,  $(|u\rangle\langle v|)|w\rangle = \langle v|w\rangle|u\rangle$ .

Given a Hilbert space  $\mathcal{H}$  associated with some physical system, let  $\mathcal{P}(\mathcal{H})$  denote the set of all orthogonal projectors in  $\mathcal{H}$ ,  $\mathcal{B}(\mathcal{H})$  denote the set of all bounded linear operators in  $\mathcal{H}$ ,  $\mathcal{O}(\mathcal{H})$  be the set of self-adjoint operators in  $\mathcal{H}$  and  $\mathcal{S}(\mathcal{H})$  denote the set of positive operators of unit trace in  $\mathcal{H}$ . The elements of  $\mathcal{S}(\mathcal{H})$  are the allowed *states* of the system, the elements of  $\mathcal{O}(\mathcal{H})$  are the *observables*, and, the elements of  $\mathcal{P}(\mathcal{H})$  are *events*.

In classical probability theory, recall that a probability distribution  $p$  over a finite set  $\Omega$  is defined by assigning a probability  $p(w) \geq 0$  to each  $w \in \Omega$  such that  $\sum_w p(w) = 1$ .

**Definition 2.1.1 (Expectation Value of a Classical Random Variable)** The expectation value of any real-valued random variable  $f(w)$  under the distribution  $p$  is evaluated as  $\mathbb{E}[f] = \sum_w f(w)p(w)$ .

The sample space  $\Omega$  is simply the finite set of all possible outcomes, and an *event* is any subset  $E \subset \Omega$ . The probability of event  $E$  is then given by  $P(E) = \sum_{w \in E} p(w)$ .

In quantum probability theory, an event  $E$  is a projection. For any  $\rho \in \mathcal{S}(\mathcal{H})$ ,  $\text{Tr}[E\rho]$  gives the probability of  $E$  in the state  $\rho$ . The states  $\rho$  thus plays the role of the probability distribution in the quantum setting. Just as the set of all probability distributions form a convex set, the set of all states  $\mathcal{S}(\mathcal{H})$  is a convex set. The extreme points of  $\mathcal{S}(\mathcal{H})$  are one-dimensional projections, which are called *pure states*. For any unit vector  $|v\rangle$ , the operator  $|v\rangle\langle v|$  is a one-dimensional projection. It is, in fact, the projection on the subspace (one-dimensional ray) generated by  $|v\rangle$ . Up to multiplication by a scalar of unit modulus, the unit vector  $|v\rangle$  is uniquely determined by the pure state  $|v\rangle\langle v|$  and hence we often refer to  $|v\rangle$  itself as the pure state. Note that  $|v\rangle\langle v|$  is an operator of unit trace,<sup>1</sup> usually called the *density operator* corresponding to the state  $|v\rangle$ .

Just as quantum states are the analogues of classical probability distributions, observables correspond to random variables. We next define the concept of **expectation value** of an observable. Since observables are self-adjoint operators, by the spectral theorem, every observable  $X$  has a spectrum  $\sigma(X)$  and a spectral decomposition:

$$X = \sum_{\lambda \in \sigma(X)} \lambda E_\lambda,$$

where,  $E_\lambda$  is the spectral projection for  $\lambda$ .  $E_\lambda$  is the event that the observable  $X$  takes the value  $\lambda$ . The *values* of any observable are simply the points of its spectrum. The probability that the observable  $X$  assumes the value  $\lambda$  in state  $\rho$  is  $\text{Tr}[\rho E_\lambda]$ . This naturally leads to the following definition.

---

<sup>1</sup>In the bra-ket notation,  $\text{Tr}[|u\rangle\langle v|] \equiv \langle v|u\rangle$ .

**Definition 2.1.2 (Expectation Value of an Observable)** The expectation value of observable  $X$  in state  $\rho$  is given by

$$\mathbb{E}_\rho(X) = \sum_{\lambda \in \sigma(X)} \lambda \text{Tr}[\rho E_\lambda] = \text{Tr}[\rho X]. \quad (2.1)$$

The set of projections  $\{E_\lambda\}$  constitutes a resolution of identity, that is,  $\sum_\lambda E_\lambda = I$ .

In a similar fashion, we can compute expectation of any real-valued function of observable  $X$ . If  $f$  is a real-valued function,  $f(X)$  is also a self-adjoint operator and it follows from the spectral theorem that,  $f(X) = \sum_{\lambda \in \sigma(X)} f(\lambda) E_\lambda$ . Therefore,  $\mathbb{E}_\rho(f(X)) = \text{Tr}[\rho f(X)]$ . In particular, the variance of an observable  $X$  in state  $\rho$  is given by

$$\text{Var}_\rho X = \text{Tr}[\rho X^2] - (\text{Tr}[\rho X])^2.$$

If  $m = \text{Tr}[\rho X]$  denotes the mean value of observable  $X$ , the variance can also be written as  $\text{Var}_\rho X = \text{Tr}[\rho(X - mI)^2]$ . It is then easy to see that the variance of  $X$  vanishes in state  $\rho$  iff  $\rho$  is supported on the eigenspace of  $X$  corresponding to eigenvalue  $m$ .

Now, consider the variance of  $X$  in a pure state  $|\psi\rangle$ .

$$\begin{aligned} \text{Var}_{|\psi\rangle} X &= \text{Tr}[|\psi\rangle\langle\psi|(X - mI)^2] \\ &= \langle\psi|(X - mI)^2|\psi\rangle = \|(X - mI)|\psi\rangle\|^2. \end{aligned} \quad (2.2)$$

Thus, in any pure state  $|\psi\rangle$ , there will always exist an observable  $X$  such that the variance of  $X$  in  $|\psi\rangle$  is non-zero! Contrast this with the situation in classical probability theory. The classical analogues of the pure states are extreme points of the convex set of probability measures on  $X$ , namely, the point measures. And the variance of any classical random variable vanishes at these extreme points. We thus have an important point of departure between classical and quantum probabilities. In the quantum case, corresponding to the extreme points, namely the pure states, we can always find an observable with non-vanishing variance. The indeterminacy is thus much stronger in the case of quantum probabilities. In fact, examining the variances of pairs of observables leads to a formulation of Heisenberg's uncertainty principle. For a more detailed study of finite-level quantum probability theory, see [9].

We conclude this section with a brief discussion on the concept of **measurement** in quantum theory. Let the eigenvalues of the observable  $X$  be labeled as  $\{\lambda_1, \lambda_2, \dots, \lambda_k\}$ , with associated projections  $\{E_{\lambda_i}\}$ . Then, the measurement process associated with an observable  $X = \sum_{\lambda_i \in \sigma(X)} \lambda_i E_{\lambda_i}$  essentially specifies which value  $\lambda_i$  is realized for a given state  $\rho$ . The probability that label  $i$  is observed is  $\text{Tr}[\rho E_{\lambda_i}]$ . Furthermore, the measurement process transforms the state in a non-trivial fashion. The von Neumann collapse postulate states that the measurement process

collapses the state  $\rho$  to a state  $\rho' = \sum_i E_{\lambda_i} \rho E_{\lambda_i}$ . Note that  $\rho'$  is also positive and has unit trace.

Associated with every observable  $X$  is thus a *projection-valued measurement* characterized by the set of spectral projections  $\{E_{\lambda_i}\}$ . More generally, quantum measurement theory also considers generalized measurements.

**Definition 2.1.3 (Generalized Measurements)** A generalized quantum measurement  $\mathcal{L}$  for a finite-level system is characterized by a finite family of operators  $\{L_j\}$  with the property that  $\sum_{j=1}^k L_j^\dagger L_j = I$ . In state  $\rho$ , the label  $j$  is observed with probability  $\text{Tr}[\rho L_j^\dagger L_j]$  and the post measurement state is given by  $\rho' = \sum_{j=1}^k L_j \rho L_j^\dagger$ .

The transformation  $\rho \rightarrow \rho'$  effected by a generalized measurement is indeed a completely positive trace-preserving (CPTP) map.

## 2.2 Entanglement

Given two quantum systems  $\mathcal{H}_A$  and  $\mathcal{H}_B$ , the Hilbert space associated with the *composite* system  $AB$  is denoted as  $\mathcal{H}_{AB}$ . In classical probability theory, if  $\Omega_1$  and  $\Omega_2$  are the sample spaces for two experiments, the sample space for the joint experiment is given by the Cartesian product  $\Omega_1 \times \Omega_2$ . The analogue of that in the quantum case is the tensor product of the two Hilbert spaces, that is,  $\mathcal{H}_{AB} = \mathcal{H}_A \otimes \mathcal{H}_B$ . In other words, if  $|u\rangle \in \mathcal{H}_A$  and  $|v\rangle \in \mathcal{H}_B$ , then,  $|u\rangle \otimes |v\rangle \in \mathcal{H}_{AB}$ . The tensor symbol is often omitted and the elements of  $\mathcal{H}_{AB}$  are simply denoted as linear combinations of  $|u\rangle|v\rangle$ .

If  $\dim(\mathcal{H}_A) = d_A$  and  $\dim(\mathcal{H}_B) = d_B$ , then,  $\dim(\mathcal{H}_{AB}) = d_A d_B$ . Furthermore, if the vectors  $\{|e_1^A\rangle, |e_2^A\rangle, \dots, |e_m^A\rangle\}$  constitute an orthonormal basis for  $\mathcal{H}_A$  and  $\{|f_1^B\rangle, |f_2^B\rangle, \dots, |f_n^B\rangle\}$  form an orthonormal basis for  $\mathcal{H}_B$ , then the vectors  $\{|e_i^A\rangle \otimes |f_j^B\rangle\}$  constitute an orthonormal basis for  $\mathcal{H}_{AB}$ . Once such a basis is fixed, the vectors are often simply denoted as  $|e_i^A\rangle|f_j^B\rangle \equiv |ij\rangle$ . Similarly, we can denote the composite system corresponding to several such Hilbert spaces  $\mathcal{H}_{A_1}, \mathcal{H}_{A_2}, \dots, \mathcal{H}_{A_N}$  as  $\mathcal{H}_{A_1 A_2 \dots A_N} = \mathcal{H}_{A_1} \otimes \mathcal{H}_{A_2} \otimes \dots \otimes \mathcal{H}_{A_N}$ . The basis vectors of such a composite system are simply denoted as  $|ijk\dots\rangle \equiv |e_i^{A_1}\rangle \otimes |f_j^{A_2}\rangle \otimes |g_k^{A_3}\rangle \otimes \dots$

For example, consider the two-dimensional complex Hilbert space  $\mathbb{C}^2$ . This corresponds to a **one-qubit system** in quantum information theory. The two basis vectors of  $\mathbb{C}^2$  are commonly denoted as

$$|0\rangle = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad |1\rangle = \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

The  $n$ -fold tensor product space  $(\mathbb{C}^2)^{\otimes n}$  is the  **$n$ -qubit Hilbert space**. The basis vectors of  $(\mathbb{C}^2)^{\otimes n}$  can thus be written in terms of binary strings, as  $|x_1 x_2 \dots x_n\rangle$ , where  $x_i \in \{0, 1\}$ .

Classically, the extreme points of the joint probability distributions over  $\Omega_1 \times \Omega_2$  are Dirac measures on the joint sample space, of the form,  $\delta_{(\omega_1, \omega_2)} = \delta_{\omega_1} \otimes \delta_{\omega_2}$ , for points  $\omega_1 \in \Omega_1$  and  $\omega_2 \in \Omega_2$ . The extreme points of the set of joint distributions are in fact products of Dirac measures on the individual sample spaces. The situation is however drastically different for composite Hilbert spaces: there exist pure states of a bipartite Hilbert space which are not product states, but are sums of product states. Such states are said to be **entangled**. We will formalize this notion in the following section.

### 2.2.1 Schmidt Decomposition

We first state and prove an important property of pure states in a bipartite Hilbert space  $\mathcal{H}_{AB}$ .

**Theorem 2.2.1 (Schmidt Decomposition)** *Every pure state  $|\psi\rangle \in \mathcal{H}_{AB}$  can be written in terms of non-negative real numbers  $\{\lambda_k (k \leq r)\}$ , and two orthonormal sets  $\{|\phi_k^A\rangle\} \subset \mathcal{H}_A$  and  $\{|\psi_k^B\rangle\} \subset \mathcal{H}_B$ , as,*

$$|\psi\rangle = \sum_{k=1}^r \lambda_k |\phi_k^A\rangle |\psi_k^B\rangle, \quad (2.3)$$

where,  $\lambda_k$  satisfy  $\sum_{k=1}^r \lambda_k^2 = 1$ .

*Proof* Consider a pure state  $|\psi\rangle \in \mathcal{H}_{AB}$ . Let  $\mathcal{H}_A$  be of dimension  $d_A$  and  $\mathcal{H}_B$  of dimension  $d_B$ . In terms of the orthonormal bases for  $\mathcal{H}_A$  and  $\mathcal{H}_B$ , the state  $|\psi\rangle$  can be written as

$$|\psi\rangle = \sum_{i=1}^{d_A} \sum_{j=1}^{d_B} a_{ij} |i_A\rangle |j_B\rangle.$$

The matrix of coefficients  $[A]_{ij} = a_{ij}$  completely characterizes the pure state  $|\psi\rangle$ . Let  $r = \text{rank}(A)$ .  $A$  can be represented via singular value decomposition, as  $A = UDV^\dagger$ , where  $U, V$  are unitaries of order  $d_A, d_B$  respectively, and  $D$  is a diagonal matrix of rank  $r$ , whose entries are the singular values ( $\lambda_k$ ) of  $A$ . Thus,

$$\begin{aligned} |\psi\rangle &= \sum_{i,j,k} u_{ik} \lambda_k v_{kj} |i_A\rangle |j_B\rangle \\ &= \sum_{k=1}^r \lambda_k |\phi_k^A\rangle |\psi_k^B\rangle, \end{aligned} \quad (2.4)$$

where, the vectors  $\{|\phi_k^A\rangle\}$  constitute an orthonormal set in  $\mathcal{H}_A$  and  $\{|\psi_j^B\rangle\}$  constitute an orthonormal set in  $\mathcal{H}_B$ , since  $U$  and  $V$  are unitary matrices.

Since the state  $|\psi\rangle \in \mathcal{H}_{AB}$  is normalized, the corresponding coefficient matrix  $A$  is such that  $\sum_{i,j} |a_{ij}|^2 = 1$ . This in turn implies that  $\sum_{k=1}^r \lambda_k^2 = 1$ .  $\square$

Given any density operator (state)  $\rho_{AB} \in \mathcal{H}_{AB}$ , the *reduced state* on  $\mathcal{H}_A$ , denoted as  $\rho_A$ , is obtained by tracing out over an orthonormal basis in  $\mathcal{H}_B$ :  $\rho_A = \text{Tr}_B[\rho_{AB}]$ . Thus,  $\rho_A$  is the operator on  $\mathcal{H}_A$  which satisfies, for all  $|u\rangle, |v\rangle \in \mathcal{H}_A$ , the condition

$$\langle v|\rho_A|u\rangle = \sum_j \langle v|\langle f_j|\rho_{AB}|u\rangle|f_j\rangle,$$

for some (and in fact any) orthonormal basis  $\{|f_j\rangle\}$  of  $\mathcal{H}_B$ . Similarly, the reduced state  $\rho_B$  on  $\mathcal{H}_B$ , is obtained by tracing out over an orthonormal basis in  $\mathcal{H}_A$ :  $\rho_B = \text{Tr}_A[\rho_{AB}]$ . The states  $\rho_A, \rho_B$  are also called *marginal states*, since they are analogues of the marginal distributions of a joint distribution in classical probability theory.

Now, consider the marginals of a pure state  $|\psi\rangle \in \mathcal{H}_{AB}$ . It is a simple exercise to show that the marginals of  $|\psi\rangle$  are not pure in general, they are *mixed* states.

**Exercise 2.2.2** Show that the reduced states of the density operator  $|\psi\rangle\langle\psi|$  corresponding to the pure state  $|\psi\rangle = \sum_{k=1}^r \lambda_k |\phi_k^A\rangle|\psi_k^B\rangle$ , are given by

$$\rho_A = \sum_{k=1}^r \lambda_k^2 |\phi_k^A\rangle\langle\phi_k^A|; \quad \rho_B = \sum_{k=1}^r \lambda_k^2 |\psi_k^B\rangle\langle\psi_k^B| \quad (2.5)$$

Thus, the marginals of a pure state  $|\psi\rangle \in \mathcal{H}_{AB}$  are no longer pure, they are mixed states, of rank equal to the rank of the coefficient matrix  $A$ . Contrast this with the classical case, where the marginals of the extreme points of the set of joint distributions are in fact extreme points of the set of distributions over the individual sample spaces. This important departure from classical probability theory, leads naturally to the notion of **entanglement**.

We can associate with any pure state  $|\psi\rangle \in \mathcal{H}_{AB}$  a probability distribution  $\{\lambda_1^2, \lambda_2^2, \dots, \lambda_k^2\}$ , via the Schmidt decomposition proved in Theorem 2.2.1. Recall that the Shannon entropy of a probability distribution  $\{p_1, p_2, \dots, p_k\}$  is defined as  $H(p) = -\sum_i p_i \log p_i$ . Correspondingly, the **von Neumann entropy** of a quantum state is defined as  $S(\rho) = -\text{Tr} \rho \log \rho$ . We will study the properties of this and other quantum entropies in greater detail in Sect. 4.2. For now, it suffices to note that for a pure state  $|\psi\rangle \in \mathcal{H}_{AB}$  of a joint system,  $S(|\psi\rangle\langle\psi|) = 0$ , whereas for the marginals  $\rho_A, \rho_B$  in Eq. (2.5),

$$S(\rho_A) = S(\rho_B) = -\sum_{k=1}^r (\lambda_k^2) \log(\lambda_k^2).$$

Again, note the quantitative departure from classical probability theory: given a joint distribution over two sample spaces  $A, B$ , the Shannon entropy of the joint distribution is always greater than the Shannon entropy of the marginals, that is,  $H(AB) \geq H(A)$ . But in quantum systems, while the pure states of a joint system will always have zero entropy, their reduced states will have non-zero entropy in general.

For a bipartite pure state  $|\psi\rangle$ , the Schmidt decomposition provides a way of quantifying the deviation of  $|\psi\rangle$  away from a product pure state. The number of product states in the Schmidt decomposition and the relative weights assigned to the different product states, quantify the entanglement of state  $|\psi\rangle$ .

**Definition 2.2.3 (Schmidt Rank)** Given a bipartite pure state  $|\psi\rangle \in \mathcal{H}_{AB}$  with a Schmidt decomposition  $|\psi\rangle = \sum_{k=1}^r \lambda_k |\phi_k^A\rangle |\psi_k^B\rangle$ ,

- (i) The number  $r$  of non-zero coefficients  $\lambda_k$  is defined to be the **Schmidt rank** of the state  $|\psi\rangle$ .
- (ii) A bipartite pure state  $|\psi\rangle$  is said to be **entangled** if it has Schmidt rank greater than one.
- (iii)  $S(\rho_A) = S(\rho_B)$  is a measure of the entanglement of the pure state  $|\psi\rangle$ , where  $\rho_A = \text{Tr}_B[|\psi\rangle\langle\psi|]$ .

For a  $d$ -level system, the von Neumann entropy of any state is bounded from above by  $\log d$ . Therefore, the maximum entanglement of a bipartite pure state  $|\psi\rangle \in \mathcal{H}_{AB}$  is  $\log \min(d_A, d_B)$ .

**Exercise 2.2.4** If  $\min(d_A, d_B) = m$ , the **maximally entangled state** in  $\mathcal{H}_{AB}$  is

$$|\psi\rangle = \frac{1}{\sqrt{m}} \sum_{i=1}^m |i_A\rangle |i_B\rangle, \quad (2.6)$$

where  $\{|i_A\rangle\}$  and  $\{|i_B\rangle\}$  are orthonormal bases in  $\mathcal{H}_A$  and  $\mathcal{H}_B$ .

We have thus far restricted our discussion to *pure* bipartite states. We can formally define product and entangled states in general as follows: a state  $\rho \in \mathcal{S}(\mathcal{H}_{AB})$  is called a *product state* if it is of the form  $\rho = \rho_A \otimes \rho_B$ , where  $\rho_A \in \mathcal{S}(\mathcal{H}_A)$  and  $\rho_B \in \mathcal{S}(\mathcal{H}_B)$ ; if not,  $\rho$  is said to be *entangled*. There are several interesting questions that arise in this context. For example, given a pair of states  $\rho_A \in \mathcal{S}(\mathcal{H}_A)$ ,  $\rho_B \in \mathcal{S}(\mathcal{H}_B)$ , consider the following convex set.

$$\mathcal{C}(\rho_A, \rho_B) = \{\rho \in \mathcal{S}(\mathcal{H}_{AB}) \mid \text{Tr}_B[\rho] = \rho_A, \text{Tr}_A[\rho] = \rho_B\}. \quad (2.7)$$

Then, what is the least value  $\min\{S(\rho) \mid \rho \in \mathcal{C}(\rho_A, \rho_B)\}$  of the von Neumann entropy of states  $\rho$  that belong to this convex set? The maximum value is attained when the systems  $A$  and  $B$  are unentangled, that is,  $\rho = \rho_A \otimes \rho_B$ , so that  $S(\rho) = S(\rho_A) + S(\rho_B)$ . We know this is indeed the maximum possible value because the von Neumann entropy satisfies the **strong subadditivity property**:  $S(\rho) \leq S(\rho_A) + S(\rho_B)$  for any  $\rho \in \mathcal{H}_{AB}$ . The strong subadditivity of the von



Neumann entropy is discussed in greater detail in Sect. 4.2. Estimating the minimum value  $\min\{S(\rho)|\rho \in \mathcal{C}(\rho_A, \rho_B)\}$  is in general a hard problem, though some estimates have been obtained for special cases [7]. Interestingly, the analogous problem in classical probability theory also remains an open!

### 2.2.2 Unitary Bases, EPR States and Dense Coding

For a finite-dimensional Hilbert space  $\mathcal{H}$  of dimension  $d$ , consider  $\mathcal{B}(\mathcal{H})$ , the set of all bounded linear operators on  $\mathcal{H}$ .  $\mathcal{B}(\mathcal{H})$  is also a Hilbert space with the inner product between any two operators  $X, Y \in \mathcal{B}(\mathcal{H})$  defined as  $\langle X|Y \rangle = \text{Tr}[X^\dagger Y]$ .  $\mathcal{B}(\mathcal{H})$  has a *unitary orthogonal basis*, that is, it admits a family of unitary operators  $\{W_\alpha, \alpha = 1, 2, 3, \dots, d^2\}$ , such that  $\text{Tr}[W_\alpha^\dagger W_\beta] = d \delta_{\alpha\beta}$ . In coding theory, such a basis is called a unitary error basis.

Consider the bipartite system  $\mathcal{H} \otimes \mathcal{H}$  composed of two copies of  $\mathcal{H}$ . Let  $|\psi\rangle$  denote a maximally entangled pure state in  $\mathcal{H} \otimes \mathcal{H}$ , which is written as per Eq. (2.6) as,

$$|\psi\rangle = \frac{1}{\sqrt{d}} \sum_i |i\rangle|i\rangle.$$

Such a maximally entangled state is often referred to as an **Einstein-Podolsky-Rosen (EPR) state**, after the famous trio—Einstein, Podolsky and Rosen—who were the first to note that states like these exhibit strange properties which are uniquely quantum. We now show how a unitary orthogonal basis for  $\mathcal{B}(\mathcal{H})$  can in fact be used to construct a basis of EPR states, which lie at the heart of several quantum communication and cryptographic tasks.

Consider the states  $|\psi_\alpha\rangle$  generated by applying the unitaries  $\{W_\alpha\}$  to one half of the maximally entangled state  $|\psi\rangle$ , as follows:

$$|\psi_\alpha\rangle = \frac{1}{\sqrt{d}} \sum_{i=1}^d (W_\alpha|i\rangle)|i\rangle, \quad \alpha = 1, 2, \dots, d^2. \quad (2.8)$$

In other words, the state  $|\psi\rangle$  is being acted upon by the operators  $\{W_\alpha \otimes I\}$ . It is then a simple exercise to show that the states  $|\psi_\alpha\rangle$  are mutually orthogonal and maximally entangled.

**Exercise 2.2.5 (Basis of EPR States)** *The states  $\{|\psi_\alpha\rangle, \alpha = 1, 2, \dots, d^2\}$  defined in Eq. (2.8) constitute an orthonormal basis—called the basis of EPR states—for  $\mathcal{H} \otimes \mathcal{H}$ . Furthermore, each  $|\psi_\alpha\rangle$  is maximally entangled, that is,*

$$\text{Tr}_{1,2} [|\psi_\alpha\rangle\langle\psi_\alpha|] = \frac{I}{d}, \quad \forall \alpha,$$

*whether the partial trace is taken over the first or the second Hilbert space.*

A remark on uniqueness and existence of unitary orthogonal bases: note that a given unitary orthogonal basis is invariant under scalar multiplication (with a scalar of unit modulus), permutation, and conjugation by a fixed operator ( $W_\alpha \rightarrow \Gamma W_\alpha \Gamma^\dagger$ ). It is an interesting question to classify the distinct unitary orthogonal bases upto this equivalence. A simple construction of such a unitary error basis in any dimension  $d$  is as follows. Consider the group  $\mathbb{Z}_d$  and identify  $\mathcal{H} \equiv \ell_2(\mathbb{Z}_d)$ . Then, consider the translation and rotation operators  $U_a, V_b$  ( $a, b \in \mathbb{Z}_d$ ), namely,

$$U_a \delta_x = \delta_{x+a}; \quad V_b \delta_x = b(x) \delta_x,$$

where  $\{\delta_x \mid \forall x \in \mathbb{Z}_d\}$  is the standard orthonormal basis for  $\mathcal{H}$  and  $b$  is a character of the group  $\mathbb{Z}_d$ . The products  $\{U_a V_b \mid a, b \in \mathbb{Z}_d\}$ , which are the Weyl operators of the abelian group  $\mathbb{Z}_d$ , form a unitary error basis.

**Superdense coding** is a simple yet important example of a quantum communication task that is made possible using EPR states. The goal of the task is for one party  $A$  (*Alice*) to communicate  $d^2$  classical messages to another party  $B$  (*Bob*), by physically sending across just one quantum state. Say  $\mathcal{H}_A$  is the  $d$ -dimensional Hilbert space associated with Alice's system and  $\mathcal{H}_B$  the  $d$ -dimensional Hilbert space corresponding to Bob's system. We assume that Alice and Bob share a bipartite EPR state, namely,  $|\psi_{AB}\rangle = \frac{1}{\sqrt{d}} \sum_{i=1}^d |i_A\rangle |i_B\rangle$ . In other words, the joint system  $\mathcal{H}_{AB}$  is in state  $|\psi_{AB}\rangle$ . Now, if Alice wants to send the message  $\alpha \in \{1, 2, \dots, d^2\}$ , she simply applies the *unitary gate*<sup>2</sup>  $W_\alpha$  on her half of the state. The joint state of  $\mathcal{H}_{AB}$  is then transformed to  $|\psi_\alpha\rangle$ .

Alice then sends across her half of the state to Bob. To *decode* the message, Bob has to perform a measurement and obtain a classical output. As discussed in Sect. 2.1, corresponding to the orthonormal basis  $\{|\psi_\alpha\rangle\}$ , we have a projective measurement characterized by the collection of one-dimensional projections  $\{|\psi_\alpha\rangle\langle\psi_\alpha|\}$ . When Bob performs this measurement on any state  $\rho$ , the probability of obtaining outcome  $\alpha$  is  $\text{Tr}[\rho |\psi_\alpha\rangle\langle\psi_\alpha|]$ . Thus, given the state  $|\psi_\alpha\rangle$ , Bob will correctly decode the message  $\alpha$  with probability one. The idea of dense coding was first proposed in [1]; for recent pedagogical discussions, see [6, 8].

As an aside, another interesting application of the unitary group is in proving *universality* of a set of quantum gates. For a composite system made up of  $k$  finite-dimensional Hilbert spaces  $\mathcal{H}_1 \otimes \mathcal{H}_2 \otimes \dots \otimes \mathcal{H}_k$ , consider the unitary group  $\mathcal{U}(\mathcal{H}_1 \otimes \mathcal{H}_2 \otimes \dots \otimes \mathcal{H}_k)$ . Let  $\mathcal{U}_{i,j}$  be the set of unitary operators  $\{U_{ij}\}$  that act only on Hilbert spaces  $\mathcal{H}_i \otimes \mathcal{H}_j$ , leaving the other Hilbert spaces unaffected.  $\mathcal{U}_{i,j}$  is a subgroup of  $\mathcal{U}$ . The *universality theorem* states that every unitary operator  $U \in \mathcal{U}$  can be decomposed as a product  $U = U_1 U_2 \dots U_N$ , where each  $U_m$  is an element of  $\mathcal{U}_{i,j}$  for some  $i, j \in 1, \dots, k$ . That is, any unitary operator acting on the composite system of  $k$  Hilbert spaces can be written as a product of unitaries that act non-trivially only on two of the  $k$  Hilbert spaces.

<sup>2</sup>Since deterministic state changes are effected by unitary operators in quantum theory, unitaries are often referred to as *quantum gates* in the quantum computing literature.

## 2.3 Schmidt Rank of Bipartite Entangled States

Using the concepts defined in Sect. 2.2.1, we will now explore a few interesting problems relating to the Schmidt rank of bipartite entangled states.

### 2.3.1 Subspaces of Minimal Schmidt Rank

Given a pair of finite-dimensional Hilbert spaces  $\mathcal{H}_1 \otimes \mathcal{H}_2$ , a question of interest is to construct subspaces  $\mathcal{S} \subset \mathcal{H}_1 \otimes \mathcal{H}_2$  such that every pure state  $|\psi\rangle \in \mathcal{S}$  has Schmidt rank  $\geq k$ . In particular, what is the maximum dimension of a subspace  $\mathcal{S}$  whose pure states are all of Schmidt rank greater than or equal to  $k$ ? Note that any state that has support on such a subspace  $\mathcal{S}$  will necessarily be a highly entangled state. Since entangled states are communication resources (as seen in the case of dense coding, for example), this question assumes importance in quantum information theory.

Let  $\mathbb{M}(m, n)$  denote the set of  $m \times n$  complex matrices. For elements  $X, Y \in \mathbb{M}(m, n)$ , the inner product is defined as  $\langle X|Y\rangle = \text{Tr}[X^\dagger Y]$ . Suppose the Hilbert spaces  $\mathcal{H}_1$  and  $\mathcal{H}_2$  are of dimensions  $m$  and  $n$  respectively, there is a natural identification between  $\mathcal{H}_1 \otimes \mathcal{H}_2$  and  $\mathbb{M}(m, n)$ . To see the explicit correspondence between elements of  $\mathcal{H}_1 \otimes \mathcal{H}_2$  and  $\mathbb{M}(m, n)$ , define a *conjugation*  $J$  on  $\mathcal{H}_2$ . For product vectors  $|u\rangle|v\rangle \in \mathcal{H}_1 \otimes \mathcal{H}_2$ , the conjugation acts as follows:

$$\Gamma(J) : |u\rangle|v\rangle \rightarrow |u\rangle\langle Jv|. \quad (2.9)$$

Since the product vectors  $|u\rangle|v\rangle$  are total in  $\mathcal{H}_1 \otimes \mathcal{H}_2$ ,  $\Gamma(J)$  defines the correspondence  $\mathcal{H}_1 \otimes \mathcal{H}_2 \rightarrow \mathbb{M}(m, n)$ . The corresponding operators  $|u\rangle\langle Jv|$  span  $\mathbb{M}(m, n)$ . Furthermore,  $\Gamma(J)$  is inner product preserving, and is therefore a unitary isomorphism between  $\mathcal{H}_1 \otimes \mathcal{H}_2$  and  $\mathbb{M}(m, n)$ . The following exercise is a simple consequence of this isomorphism.

**Exercise 2.3.1** *Show that the Schmidt rank of any pure state  $|\psi\rangle \in \mathcal{H}_1 \otimes \mathcal{H}_2$  is equal to the rank of  $\Gamma(J)(|\psi\rangle)$ .*

The problem of finding subspaces for which every vector has a Schmidt rank greater than or equal to  $k$ , thus reduces to the following matrix-theoretic problem: Construct subspaces  $\mathcal{S} \subset \mathbb{M}(m, n)$ , with the property that every non-zero element of these subspaces is of rank greater than or equal to  $k$ . In particular, the problem is to find the maximum dimension of the subspace  $\mathcal{S} \subset \mathbb{M}(m, n)$ , whose non-zero elements are all of rank greater than or equal to  $k$ . While the problem remains open for a general  $k$ , we will present an example of such a construction for  $k = 2$ .

Let  $\{W_\alpha, \alpha = 1, \dots, n^2\}$  denote a unitary orthogonal basis for the space of square matrices  $\mathbb{M}_n \equiv \mathbb{M}(n, n)$ . Consider a positive, self-adjoint matrix  $\Phi \in \mathbb{M}_n$  with the

spectral resolution,

$$\Phi = \sum_{\alpha=1}^{n^2} p_{\alpha} |W_{\alpha}\rangle\langle W_{\alpha}|,$$

with distinct eigenvalues  $p_{\alpha} > 0$ . Now, we can construct a subspace  $\mathcal{S} \subset \mathbb{M}_{2n}$ , of square matrices in  $\mathbb{M}_{2n} \equiv \mathbb{M}(2n, 2n)$ , whose non-zero elements are of rank  $k \geq 2$ .

**Theorem 2.3.2** Consider  $\mathcal{S} \subset \mathbb{M}(2n, 2n)$  defined as follows:

$$\mathcal{S} = \left\{ X \equiv \begin{pmatrix} A & \Phi(B) \\ B & A \end{pmatrix}, A, B \in \mathbb{M}_n \right\}. \quad (2.10)$$

The elements  $X \in \mathcal{S}$  satisfy  $\text{rank}(X) \geq 2$ , for arbitrary  $A, B \in \mathbb{M}_n$ .

*Proof* Suppose  $B = 0$  and  $A \neq 0$ . Then, the off-diagonal blocks are zero, but the diagonal blocks are non-zero and  $\text{rank}(X) = 2$  for all  $A \in \mathbb{M}_n$ . If  $A = 0$  but  $B \neq 0$ , then, since  $\Phi$  is non-singular,  $\Phi(B) \neq 0$ . Therefore,  $\text{rank}(X) \geq 2, \forall B \in \mathbb{M}_n$ .

Now, it remains to show that  $\text{rank}(X) \geq 2$ , for all  $A, B \neq 0$  in  $\mathbb{M}_n$ . Suppose  $\text{rank}(X(A, B)) = 1$ . Such a matrix  $X$  is of the form

$$\begin{bmatrix} |u\rangle \\ |v\rangle \end{bmatrix} \begin{bmatrix} \langle u'| \\ \langle v'| \end{bmatrix} = \begin{pmatrix} |u\rangle\langle u| & |u\rangle\langle v'| \\ |v\rangle\langle u'| & |v\rangle\langle v'| \end{pmatrix},$$

where,  $|u\rangle, |v\rangle$  are column vectors of length  $n$ . Comparing with the definition of  $X$  in Eq. (2.10), we see that,

$$|u\rangle\langle u'| = |v\rangle\langle v'| = A, \quad \Phi(|v\rangle\langle u'|) = |u\rangle\langle v'|. \quad (2.11)$$

The first equality implies that there exists scalars  $c, c' \neq 0$  such that  $|v\rangle = c|u\rangle$  and  $\langle v'| = c'\langle u'|$ . The second equality in Eq. (2.11) thus becomes

$$\Phi(|v\rangle\langle u'|) = c'c^{-1}|v\rangle\langle u'|.$$

This implies that  $|v\rangle\langle u'|$  an eigenvector of  $\Phi$ . However, the eigenvectors of  $\Phi$  belong to the unitary error basis, and thus  $\Phi$  cannot have such a rank-one matrix as its eigenvector. Thus, the assumption that  $X(A, B)$  as defined in Eq. (2.10) if of rank one leads to a contradiction. This proves the claim.  $\square$

The above construction in fact leads to a more interesting property of the set of matrices in  $\mathbb{M}_{2n}$ .

**Theorem 2.3.3** The space of  $2n \times 2n$  matrices  $\mathbb{M}_{2n}$  is a direct sum of the form  $\mathbb{M}_{2n} = \mathcal{S} \oplus \mathcal{S}^{\perp}$ , where,  $\text{rank}(X) \geq 2$ , for every  $0 \neq X \in \mathcal{S} \cup \mathcal{S}^{\perp}$ .

*Proof* Consider an element of  $\mathcal{S}^\perp$ , the orthogonal complement of  $\mathcal{S}$ , written in block matrix form as:

$$Y = \begin{pmatrix} K & L \\ M & N \end{pmatrix}.$$

Since  $\text{Tr}[Y^\dagger X] = 0$  for any  $X \in \mathcal{S}$  and  $Y \in \mathcal{S}^\perp$ , we have,

$$\begin{aligned} \text{Tr} \left[ \begin{pmatrix} K^\dagger & M^\dagger \\ L^\dagger & N^\dagger \end{pmatrix} \begin{pmatrix} A & \Phi(B) \\ B & A \end{pmatrix} \right] &= 0 \\ \Rightarrow \text{Tr}[K^\dagger A + M^\dagger B + L^\dagger \Phi(B) + N^\dagger A] &= 0, \quad \forall A, B. \end{aligned} \quad (2.12)$$

Setting  $B = 0$ , we have,

$$\text{Tr}[(K^\dagger + N^\dagger)A] = 0, \quad \forall A \Rightarrow K + N = 0.$$

Similarly, setting  $A = 0$ , we have,

$$\text{Tr}[(M + \Phi(L))^\dagger B] = 0, \quad \forall B \Rightarrow M + \Phi(L) = 0.$$

Thus, every element of the orthogonal complement of  $\mathcal{S}$  is of the form,

$$Y = \begin{pmatrix} K & L \\ -\Phi(L) & -K \end{pmatrix}.$$

By the same argument as in the proof of Theorem 2.3.2,  $\text{rank}(Y) \geq 2$ , for all  $Y \in \mathcal{S}^\perp$ .  $\square$

The above result for  $\mathbb{M}_{2n}$  naturally leads to the following problem for the more general matrix spaces  $\mathbb{M}(m, n)$ .

**Question 2.3.4** *Identify the quadruples  $(m, n, r, s)$  for which it is possible to find a direct sum decomposition  $\mathbb{M}(m, n) = \mathcal{S} \oplus \mathcal{S}^\perp$  of  $\mathbb{M}(m, n)$ , into a direct sum of subspace  $\mathcal{S}$  and its orthogonal complement  $\mathcal{S}^\perp$ , such that,*

$$\text{rank}(X) \geq r, \quad \forall 0 \neq X \in \mathcal{S}, \quad \text{rank}(Y) \geq s, \quad \forall 0 \neq Y \in \mathcal{S}^\perp.$$

Finally, we return to the question of the maximum dimension of the subspace  $\mathcal{S}$  whose elements have Schmidt rank greater than or equal to  $k$ .

**Theorem 2.3.5** *Let  $\mathcal{L} \subset \mathbb{M}(m, n)$  be such that  $\forall 0 \neq X \in \mathcal{L}, \text{rank}(X) \geq k + 1$ . Then,  $\dim(\mathcal{L}) \leq (m - k)(n - k)$ .*

*Proof* Let  $S_k$  denote the *variety* of matrices of rank less than or equal to  $k$  in  $\mathbb{M}(m, n)$ . Then, it is known [3] that  $\dim(S_k) = k(m + n - k)$ . Let  $\mathcal{L}$  be a subspace

( $\mathcal{L} \subset \mathbb{M}(m, n)$ ) such that for all  $0 \neq X \in \mathcal{L}$ ,  $\text{rank}(X) \geq k + 1$ . Note that,  $S_k \cap \mathcal{L} = 0$ . Using a generalization of Azoff's Theorem in algebraic geometry [2], we have,

$$\dim(S_k) + \dim(\mathcal{L}) \leq mn.$$

The dimension of any such subspace  $\mathcal{L}$  is therefore given by

$$\dim(\mathcal{L}) \leq (m - k)(n - k). \quad (2.13)$$

□

Furthermore, we can in fact explicitly construct a subspace  $\mathcal{L}_0$  whose dimension is exactly equal to the maximum value  $(m - k)(n - k)$ . Consider polynomials  $P$  such that  $\deg(P) \leq p - k - 1$ . Construct diagonal matrices of order  $p \times p$  with entries  $D = \text{diag}(P(z_1), P(z_2), \dots, P(z_p))$ , where  $z_1, z_2, \dots, z_p$  are all distinct. The linear space of such diagonal matrices, denoted by  $\mathcal{D}_{p,k}$  is of dimension  $p - k$ . At most  $p - k - 1$  entries of such a diagonal matrix  $D$  can be zero. Therefore,  $\text{rank}(D) \geq k + 1$ ,  $\forall D \in \mathcal{D}_{p,k}$ . The diagonal complementary space  $\mathcal{D}_{p,k}^\perp$  is of dimension  $k$ .

Now the construction proceeds as follows. Any matrix  $X \in \mathcal{L}_0$  has  $m + n - 1$  diagonals. Fill all diagonals of length less than or equal to  $k$  with zeros. Let the least length of a non-vanishing diagonal be  $p$ . Choose the entries of this diagonal of length  $p$  from some matrix  $D \in \mathcal{D}_{p,k}$ . Consider the  $p \times p$  minor of such an  $X$ . This is either a lower triangular or an upper triangular matrix. The diagonal of this  $p \times p$  minor has the property that at most  $p - k - 1$  entries are zero. Therefore, rank of such a  $p \times p$  minor is greater than or equal to  $k + 1$ . By this construction, every non-zero element  $X \in \mathcal{L}_0$  is of rank at least  $k + 1$ . Therefore,

$$\dim(\mathcal{L}_0) = mn - k(k + 1) - (m + n - 1 - 2k)k = (m - k)(n - k),$$

where the second term enumerates the zero entries of the diagonals of length  $k$  and the final term enumerates the zero entries of the non-vanishing diagonals.

This construction and the result of Theorem 2.3.5 thus imply the following result for the quantum information theoretic problem of finding subspaces of high entanglement.

**Theorem 2.3.6** *Given a composite system  $\mathcal{H}_1 \otimes \mathcal{H}_2$  with  $\dim(\mathcal{H}_1) = m$  and  $\dim(\mathcal{H}_2) = n$  and subspaces  $\mathcal{S} \subset \mathcal{H}_1 \otimes \mathcal{H}_2$ ,*

$$\max \{ \dim(\mathcal{S}) \mid \text{Schmidt rank}(|\psi\rangle) \geq k, \forall \psi \in \mathcal{S} \} = (m - k)(n - k). \quad (2.14)$$

## 2.4 Schmidt Number of Mixed States

We now move beyond pure states and study entanglement in mixed states. Recall that pure states are the extreme points of the set  $\mathcal{S}(\mathcal{H})$  of positive operators of unit-trace. A general state  $\rho \in \mathcal{S}(\mathcal{H})$  is thus a mixture of pure states; it is a mixed state, with  $\rho \geq 0$  and  $\text{Tr}[\rho] = 1$ . For the composite system  $\mathcal{H}_1 \otimes \mathcal{H}_2$ , with  $\dim(\mathcal{H}_1) = m$  and  $\dim(\mathcal{H}_2) = n$ , consider the sets

$$\mathcal{S}_k = \{|\psi\rangle\langle\psi| \in \mathcal{S}(\mathcal{H}_1 \otimes \mathcal{H}_2) \mid \text{Schmidt No.}(|\psi\rangle) \leq k\}, \quad (2.15)$$

$$\tilde{\mathcal{S}}_k = \left\{ \int_{\text{Schmidt No.}(|\psi\rangle) \leq k} |\psi\rangle\langle\psi| \mu d(\psi) \right\}, \quad (2.16)$$

where  $\mu$  is a probability distribution.  $\tilde{\mathcal{S}}_k$  is thus the set of all mixed states that are convex combinations of the pure states in  $\mathcal{S}_k$ . We know from the Schmidt decomposition that for  $|\psi\rangle \in \mathcal{H}_1 \otimes \mathcal{H}_2$ ,  $1 \leq k \leq \min(m, n)$ . Note that  $\tilde{\mathcal{S}}_k$  is a compact convex set in the real linear space of Hermitian operators on  $\mathcal{H}_1 \otimes \mathcal{H}_2$ , of dimension  $m^2 n^2$ . It then follows from Carathéodory's Theorem that every  $\rho \in \tilde{\mathcal{S}}_k$  can be expressed as

$$\rho = \sum_{j=1}^{m^2 n^2 + 1} p_j |\psi_j\rangle\langle\psi_j|,$$

with  $|\psi_j\rangle \in \mathcal{S}_k$ ,  $p_j \geq 0$ , and  $\sum_j p_j = 1$ . It is therefore enough to consider finite convex combinations of pure states to represent the elements of  $\tilde{\mathcal{S}}_k$ .

**Definition 2.4.1 (Schmidt Number)** A state  $\rho$  is defined to have Schmidt number  $k$  if  $\rho \in \tilde{\mathcal{S}}_{k+1} \setminus \tilde{\mathcal{S}}_k$ .

Estimating the Schmidt number for arbitrary mixed states is in general a hard problem. A related question of interest is whether it is possible to construct a *test* to identify states  $\rho \notin \tilde{\mathcal{S}}_k$ . Such a test would be a special case of the class of **entanglement witnesses** [4] which are well-studied in the quantum information literature (see [5] for a recent review).

In the following Section, we prove the Horodecki-Terhal criterion [11] which shows that  $k$ -positive maps can act as entanglement witnesses for states with Schmidt number exceeding  $k$ . Such a result was first proved for  $k = 1$  [4] and later extended to general  $k$ .

### 2.4.1 Test for Schmidt Number $k$ Using $k$ -Positive Maps

By the geometric Hahn-Banach Theorem, in the real linear space of Hermitian operators we can construct a linear functional  $\Lambda$  such that there exists a hyperplane corresponding to  $\Lambda(\cdot) = c$  (for some constant  $c$ ) that separates  $\rho$  and the convex set  $\tilde{\mathcal{S}}_k$ . In other words, there exists a linear functional  $\Lambda$  such that,

$$\Lambda(\rho) < c \leq \Lambda(\sigma), \quad \forall \sigma \in \tilde{\mathcal{S}}_k. \quad (2.17)$$

Any linear functional  $\Lambda$  on the Banach space of Hermitian operators can be written as  $\Lambda(X) = \text{Tr}[XA]$ , where  $A$  is a Hermitian operator. Therefore Eq. (2.17) implies, for any  $\rho \notin \tilde{\mathcal{S}}_k$ ,

$$\text{Tr}[\rho A] < c \leq \text{Tr}[\sigma A] \Rightarrow \text{Tr}[\rho H] < 0 \leq \text{Tr}[\sigma H], \quad \forall \sigma \in \tilde{\mathcal{S}}_k, \quad (2.18)$$

where,  $H = A - cI$ , is a Hermitian operator on  $\mathcal{H}_1 \otimes \mathcal{H}_2$ . Defining the map  $\Gamma_H : \mathcal{B}(\mathcal{H}_1) \rightarrow \mathcal{B}(\mathcal{H}_2)$  such that,

$$\Gamma_H(X) = \text{Tr}_{\mathcal{H}_1}[H(X \otimes I_{\mathcal{H}_2})], \quad \forall X \in \mathcal{B}(\mathcal{H}_1),$$

we can rewrite  $H$  as follows [4].

**Proposition 2.4.2** *Given an orthonormal basis of rank-1 operators  $\{E_\alpha\}$  in  $\mathcal{B}(\mathcal{H}_1)$ ,  $H$  can be written as,*

$$H = \sum_{\alpha} E_{\alpha} \otimes \Gamma_H(E_{\alpha}^{\dagger}). \quad (2.19)$$

*Proof* Choose orthonormal bases  $\{E_{\alpha}\} \in \mathcal{B}(\mathcal{H}_1)$  and  $\{F_{\beta}\} \in \mathcal{B}(\mathcal{H}_2)$ . Then,  $\{E_{\alpha} \otimes F_{\beta}\}$  constitutes an orthonormal basis for  $\mathcal{B}(\mathcal{H}_1 \otimes \mathcal{H}_2)$ . Therefore, the operator  $H$  can be written as,

$$\begin{aligned} H &= \sum_{\alpha, \beta} \text{Tr} \left[ H(E_{\alpha}^{\dagger} \otimes F_{\beta}^{\dagger}) \right] (E_{\alpha} \otimes F_{\beta}) \\ &= \sum_{\alpha, \beta} \text{Tr}_{\mathcal{H}_2} \left[ \text{Tr}_{\mathcal{H}_1} [H(E_{\alpha}^{\dagger} \otimes I_{\mathcal{H}_2})] F_{\beta}^{\dagger} \right] (E_{\alpha} \otimes F_{\beta}) = \sum_{\alpha} E_{\alpha} \otimes \Gamma_H(E_{\alpha}^{\dagger}). \end{aligned}$$

□

We now show (following [11]) that the separation in Eq. (2.18) between the set  $\tilde{\mathcal{S}}_k$  and states  $\rho \notin \tilde{\mathcal{S}}_k$  can be realized using a  $k$ -positive map from  $\mathcal{B}(\mathcal{H}_1) \rightarrow \mathcal{B}(\mathcal{H}_2)$ .



**Theorem 2.4.3** *There exists a  $k$ -positive map  $\Lambda_0 : \mathcal{B}(\mathcal{H}_1) \rightarrow \mathcal{B}(\mathcal{H}_2)$  such that*<sup>3</sup>

$$\text{Tr} \left[ (\mathbb{I} \otimes \Lambda_0) \left( \sum_{\alpha} E_{\alpha} \otimes E_{\alpha}^{\dagger} \right) \rho \right] < 0, \quad \rho \notin \tilde{\mathcal{S}}_k, \quad (2.20)$$

$$\text{Tr} \left[ (\mathbb{I} \otimes \Lambda_0) \left( \sum_{\alpha} E_{\alpha} \otimes E_{\alpha}^{\dagger} \right) \sigma \right] \geq 0, \quad \forall \sigma \in \tilde{\mathcal{S}}_k, \quad (2.21)$$

where  $\{E_{\alpha}\}$  is an orthonormal basis of rank-1 operators in  $\mathcal{B}(\mathcal{H}_1)$ .

*Proof* First we show that the operator  $H$  that defines the separation in Eq. (2.18) can be written in terms of an operator  $\Lambda_0 : \mathcal{B}(\mathcal{H}_1) \rightarrow \mathcal{B}(\mathcal{H}_2)$ .

**Lemma 2.4.4** *If  $H$  is an operator with the property*

$$\text{Tr}[\rho H] < 0 \leq \text{Tr}[\sigma H], \quad \forall \sigma \in \tilde{\mathcal{S}}_k, \quad (2.22)$$

then  $H$  is of the form

$$H = (\mathbb{I} \otimes \Lambda_0) \left( \sum_{\alpha} E_{\alpha} \otimes E_{\alpha}^{\dagger} \right) = (\mathbb{I} \otimes \Lambda_0)(mP), \quad (2.23)$$

for some operator  $\Lambda_0 : \mathcal{B}(\mathcal{H}_1) \rightarrow \mathcal{B}(\mathcal{H}_2)$ , and a projection  $P$ .

*Proof* Suppose we consider an orthonormal basis  $\{|e_i\rangle\}$  for  $\mathcal{H}_1$  and choose the rank-1 operators  $E_{\alpha}$  to be  $E_{\alpha} = |e_i\rangle\langle e_j|$ . Then,  $E_{\alpha}^{\dagger} = |e_j\rangle\langle e_i|$ . Let  $T$  denote the transpose operation with respect to the  $\{|e_i\rangle\}$  basis, then,  $E_{\alpha}^{\dagger} = T(E_{\alpha})$ . Therefore, for this specific choice of basis, Eq. (2.19) implies,

$$H = \sum_{\alpha} E_{\alpha} \otimes (\Gamma_H \circ T)(E_{\alpha}) = (\mathbb{I} \otimes \Lambda_H) \left( \sum_{\alpha} E_{\alpha} \otimes E_{\alpha}^{\dagger} \right), \quad (2.24)$$

where we have defined  $\Lambda_0(\cdot) \equiv (\Gamma_H \circ T)(\cdot)$ . It is now a simple exercise to check that  $\sum_{\alpha} E_{\alpha} \otimes E_{\alpha}^{\dagger} = mP$ , where  $P$  is a projection operator.

**Exercise 2.4.5** *Define the rank-1 operators  $E_{\alpha} = |e_i\rangle\langle e_j| \in \mathcal{B}(\mathcal{H}_1)$ , where  $\{|e_i\rangle\}$  is an orthonormal basis for the space  $\mathcal{H}_1$ . Then, the operator*

$$P = \frac{1}{m} \sum_{\alpha} E_{\alpha} \otimes E_{\alpha}$$

is a projection ( $P^{\dagger} = P = P^2$ ). □

---

<sup>3</sup>A remark on notation: we use  $\mathbb{I}$  to denote the identity map  $\mathbb{I} : \mathcal{B}(\mathcal{H}) \rightarrow \mathcal{B}(\mathcal{H})$ , with  $\mathbb{I}(\rho) = \rho$  for any  $\rho \in \mathcal{B}(\mathcal{H})$ .

The theorem is proved once we show that the operator  $\Lambda_0$  defined above is indeed  $k$ -positive.

**Lemma 2.4.6** *The operator  $\Lambda_0$  defined in Eq. (2.23) corresponding to a  $H$  that satisfies Eq. (2.22), is  $k$ -positive on  $\mathcal{B}(\mathcal{H}_1) \rightarrow \mathcal{B}(\mathcal{H}_2)$ .*

*Proof* Lemma 2.4.4 implies that for a chosen basis  $\{|e_i\rangle\} \in \mathcal{H}_1$ , any state  $\sigma \in \tilde{\mathcal{S}}_k$  satisfies,

$$\mathrm{Tr} \left[ \left( \mathbb{I} \otimes \Lambda_0 \right) \left( \sum_{i,j} |e_i\rangle\langle e_j| \otimes |e_i\rangle\langle e_j| \right) \sigma \right] \geq 0.$$

If we choose  $\sigma$  to be a pure state  $\sigma \equiv |\psi\rangle\langle\psi|$ , the corresponding vector has the Schmidt decomposition  $|\psi\rangle = \sum_{r=1}^k |u_r\rangle|v_r\rangle$ . Then, the above condition becomes,

$$\begin{aligned} & \sum_{i,j} \mathrm{Tr} \left[ |e_i\rangle\langle e_j| \otimes \Lambda_0(|e_i\rangle\langle e_j|) \left( \sum_{r,s=1}^k |u_r\rangle\langle u_s| \otimes |v_r\rangle\langle v_s| \right) \right] \\ &= \sum_{i,j} \sum_{r,s} \langle e_j|u_r\rangle\langle u_s|e_i\rangle \langle v_s|\Lambda_0(|e_i\rangle\langle e_j|)|v_r\rangle \geq 0. \end{aligned} \quad (2.25)$$

Now, note that  $\sum_i \langle u_s|e_i\rangle|e_i\rangle = \sum_i |e_i\rangle\langle e_i|J|u_s\rangle$ , where  $J$  is a conjugation. Then, since  $\{|e_i\rangle\}$  is an orthonormal basis,  $\sum_i \langle u_s|e_i\rangle|e_i\rangle = |J|u_s\rangle = |J|u_s\rangle$ , the inequality in Eq. (2.25) becomes

$$\sum_{r,s=1}^k \langle v_s|(\Lambda_0|u'_s\rangle\langle u'_r|)|v_r\rangle \geq 0, \quad (2.26)$$

for all such sets of vectors  $\{|u_r\rangle\}, \{|v_r\rangle\}$ . Recalling the definition of  $k$ -positivity (see Definition 1.1.4), this is true if and only if  $\Lambda_0$  is  $k$ -positive.  $\square$

Finally, Theorem 2.4.3 implies the existence of a  $k$ -positive map that can act as an *entanglement witness* for Schmidt number  $k$  in the following sense [11].

**Theorem 2.4.7 (Schmidt Number Witness)**

- If  $\Lambda : \mathcal{B}(\mathcal{H}_2) \rightarrow \mathcal{B}(\mathcal{H}_1)$  is a  $k$ -positive operator, then for any  $\sigma \in \tilde{\mathcal{S}}_k$ ,

$$(\mathbb{I}_{\mathcal{H}_1} \otimes \Lambda)(\sigma) \geq 0. \quad (2.27)$$

- If  $\rho \notin \tilde{\mathcal{S}}_k$ , there exists a  $k$ -positive map  $\Lambda : \mathcal{B}(\mathcal{H}_2) \rightarrow \mathcal{B}(\mathcal{H}_1)$ , such that,

$$(\mathbb{I}_{\mathcal{H}_1} \otimes \Lambda)(\rho) \not\geq 0. \quad (2.28)$$

Such a map  $\Lambda$  is an **entanglement witness** for Schmidt number  $k$ .

*Proof* The first statement of the theorem is easily proved, and is left as an exercise. To prove the second statement, we note that if a map  $\Lambda$  is  $k$ -positive, then so is its adjoint.

**Exercise 2.4.8** *If  $\Lambda : \mathcal{B}(\mathcal{H}_1) \rightarrow \mathcal{B}(\mathcal{H}_2)$  is  $k$ -positive, then,  $\Lambda^\dagger : \mathcal{B}(\mathcal{H}_2) \rightarrow \mathcal{B}(\mathcal{H}_1)$  is also  $k$ -positive.*

Since we know from Lemma 2.4.6 that the map  $\Lambda_0$  defined in Eq. (2.23) is  $k$ -positive, the above exercise implies that the map  $\Lambda_0^\dagger : \mathcal{B}(\mathcal{H}_2) \rightarrow \mathcal{B}(\mathcal{H}_1)$  is also  $k$ -positive. Then, it follows that

$$(\mathbb{I}_{\mathcal{H}_1} \otimes \Lambda_0^\dagger)(\sigma) \geq 0, \quad \forall \sigma \in \tilde{\mathcal{S}}_k. \quad (2.29)$$

Going back to Eq. (2.20) and taking the adjoint of the operator in the trace, we have,

$$\text{Tr} \left[ (\mathbb{I} \otimes \Lambda_0^\dagger)(\rho) \left( \sum_{\alpha} E_{\alpha} \otimes E_{\alpha}^{\dagger} \right) \right] < 0.$$

This implies,  $(\mathbb{I} \otimes \Lambda_0^\dagger)(\rho) \not\geq 0$ . We have thus constructed a  $k$ -positive map  $\Lambda_0^\dagger$  with the desired property.  $\square$

In order to check if a given state  $\rho$  is of Schmidt number strictly greater than  $k$ , we need a  $k$ -positive map that satisfies Eq. (2.28). Finding such  $k$ -positive maps is indeed a hard problem. There is however a vast body of work on positive but not completely positive maps, which can act as entanglement witnesses in the quantum information literature [5]. The classical example of such a test for entanglement is using the transpose operation which is positive, but not completely positive [10].

## 2.4.2 Schmidt Number of Generalized Werner States

Given a  $d$ -dimensional Hilbert space  $\mathcal{H}$ , consider the bipartite system  $\mathcal{H} \otimes \mathcal{H}$  formed using two copies of  $\mathcal{H}$ . The **Werner states** [12] which we define below, are an interesting single parameter family of bipartite states in  $\mathcal{H} \otimes \mathcal{H}$ .

For any unitary operator  $U$  in  $\mathcal{H}$ , let  $\Pi(U) : U \rightarrow U \otimes \bar{U}$  be a representation of the unitary group  $\mathcal{U}(\mathcal{H})$ .

**Proposition 2.4.9** *Any state  $\rho$  commuting with every  $U \otimes \bar{U}$  has the form*

$$\rho = \rho_F \equiv F |\psi_0\rangle\langle\psi_0| + (1 - F) \frac{(I - |\psi_0\rangle\langle\psi_0|)}{d^2 - 1}, \quad (2.30)$$

where  $|\psi_0\rangle = \frac{1}{\sqrt{d}} \sum_{i=1}^d |ii\rangle$  is the maximally entangled state in  $\mathcal{H} \otimes \mathcal{H}$ .

*Proof* This follows from the isomorphism between  $\mathcal{H} \otimes \mathcal{H}$  and  $\mathcal{B}(\mathcal{H})$ . Let  $\mathcal{J}$  denote conjugation with respect to a chosen basis  $\{|i\rangle\} \in \mathcal{H}$ . Defining the map,

$$\Gamma(\mathcal{J}) : |u\rangle|v\rangle \rightarrow |u\rangle\langle\mathcal{J}v|,$$

we see that,

$$\Gamma(\mathcal{J})\Pi(U)\Gamma(\mathcal{J})^{-1}(\cdot) = U(\cdot)U^\dagger.$$

Thus the action of the  $U \otimes \bar{U}$  corresponds to the unitary action  $X \rightarrow UXU^\dagger$  for all  $X \in \mathcal{B}(\mathcal{H})$ . We denote this representation of the unitary group as  $\tilde{\Pi}(U)$ .

Now note that the commutant  $\{\tilde{\Pi}(U), U \in \mathcal{U}(\mathcal{H})\}'$  is of dimension two, and is spanned by  $I$  and  $X \rightarrow \text{Tr}(X)I = \langle I|X\rangle|I\rangle$ . In turn, the commutant  $\{\Pi(U), U \in \mathcal{U}(\mathcal{H})\}'$  is spanned by  $I_{\mathcal{H} \otimes \mathcal{H}}$  and  $|\psi_0\rangle\langle\psi_0|$ , where  $|\psi_0\rangle = \frac{1}{\sqrt{d}} \sum_{i=1}^d |ii\rangle$ , is the maximally entangled pure state in the canonical basis. This in turn implies that a state  $\rho$  that commutes with all the unitaries  $\Pi(U)$  must be a convex combination as given in Eq. (2.30).  $\square$

The state  $\rho_F$  is the generalized **Werner State**. When  $F \geq \frac{1}{d^2}$ ,  $\rho_F$  can be rewritten as

$$\rho_F = p|\psi_0\rangle\langle\psi_0| + (1-p)\frac{I}{d^2},$$

where  $I$  is the identity operator on  $\mathcal{H} \otimes \mathcal{H}$ . Note that while the Schmidt rank of the maximally entangled state is  $d$ , the Schmidt rank of the *maximally mixed state*  $I/d^2$  is 1. Such a mixture of  $|\psi_0\rangle\langle\psi_0|$  and  $I/d^2$  can be realized by the averaging the action of the unitary group  $\Pi(U)$  on state  $\rho$ , that is, by performing the following operation:

$$\int dU(U \otimes \bar{U})\rho(U \otimes \bar{U})^\dagger.$$

Our goal is to now evaluate the Schmidt number of  $\rho_F$ . We first note that when  $F = 1$ ,  $\rho_F$  has Schmidt number  $d$ . Furthermore, for any state  $|\psi\rangle \in \mathcal{H} \otimes \mathcal{H}$ ,

$$\langle\psi|\rho_F|\psi\rangle = \frac{1}{d^2 - 1} [(d^2 F - 1)|\langle\psi_0|\psi\rangle|^2 + 1 - F]. \quad (2.31)$$

It is then an easy exercise to evaluate the maximum value of  $\langle\psi|\rho_F|\psi\rangle$ .

**Exercise 2.4.10** If  $F \geq \frac{1}{d^2}$ ,

$$\max_{|\psi\rangle \in \mathcal{H} \otimes \mathcal{H}} \langle\psi|\rho_F|\psi\rangle = F,$$

which is attained when  $|\psi\rangle = |\psi_0\rangle$ . If  $F \leq \frac{1}{d^2}$ ,

$$\max_{|\psi\rangle \in \mathcal{H} \otimes \mathcal{H}} \langle \psi | \rho_F | \psi \rangle = \frac{1 - F}{d^2 - 1},$$

and this value is attained when  $|\psi\rangle$  is orthogonal to  $|\psi_0\rangle$ .

The central result we prove in this section is the following result due to Terhal-Horodecki [11].

**Theorem 2.4.11** *Let  $\frac{k-1}{d} \leq F \leq \frac{k}{d}$ , with  $F \geq \frac{1}{d^2}$ . Then, the Schmidt number of  $\rho_F$  is equal to  $k$ .*

Thus, even though the parameter  $F$  changes continuously, the Schmidt number of  $\rho_F$  changes at discrete values of  $F$ , remaining constant for the intermediate values.

Before venturing to prove Theorem 2.4.11, we first note the following general result which proves a lower bound on the Schmidt number of a general state  $\rho \in \mathcal{H} \otimes \mathcal{H}$ , given a bound on its expectation value with a maximally entangled state.

**Proposition 2.4.12** *Suppose for a state  $\rho \in \mathcal{H} \otimes \mathcal{H}$  and a maximally entangled state  $|\psi\rangle \in \mathcal{H} \otimes \mathcal{H}$ ,  $\langle \psi | \rho | \psi \rangle > \frac{k}{d}$ . Then,*

$$\text{Schmidt number}(\rho) \geq k + 1. \quad (2.32)$$

The proof follows from the following Lemma and its corollary discussed below.

**Lemma 2.4.13** *Let  $|\psi\rangle = \sum_{j=1}^d \lambda_j |x_j\rangle |y_j\rangle$  be a state in Schmidt form. Then,*

$$\max\{|\langle \psi | \psi' \rangle|, |\psi'\rangle : \text{max. entangled state}\} = \frac{1}{\sqrt{d}} \sum_j \lambda_j. \quad (2.33)$$

*Proof* Choose  $|\psi'\rangle = \frac{1}{\sqrt{d}} \sum_{j=1}^d |\xi_j\rangle |\eta_j\rangle$ , where,  $\{|\xi_j\rangle\}$  and  $\{|\eta_j\rangle\}$  are any two orthonormal bases. Then,

$$\langle \psi | \psi' \rangle = \frac{1}{\sqrt{d}} \sum_{j,l} \lambda_j \langle x_j | \xi_l \rangle \langle y_j | \eta_l \rangle.$$

Choose unitaries  $U, V$ , such that  $U|\xi_l\rangle = |l\rangle$ ,  $V|\eta_l\rangle = |l\rangle$ . Let  $C$  denote conjugation with respect to the standard basis  $\{|l\rangle\}$ . Then,

$$\begin{aligned} \langle \psi | \psi' \rangle &= \frac{1}{\sqrt{d}} \sum_{j,l} \lambda_j \langle U x_j | l \rangle \langle V y_j | l \rangle \\ &= \frac{1}{\sqrt{d}} \sum_{j,l} \lambda_j \langle U x_j | l \rangle \langle l | C V y_j \rangle \\ &= \frac{1}{\sqrt{d}} \sum_j \lambda_j \langle U x_j | C V y_j \rangle. \end{aligned} \quad (2.34)$$

Note that, since  $|U x_j\rangle, |CV y_j\rangle$  are unit vectors and  $\lambda_j \geq 0$ ,  $|\langle\psi|\psi'\rangle| \leq \frac{1}{\sqrt{d}} \sum_j \lambda_j$ .

Now, let  $|\xi_l\rangle = |l\rangle$ , for all  $l = 1, \dots, d$ , so that  $U \equiv I$ . We choose the unitary  $V$  that satisfies  $VC|y_j\rangle = |x_j\rangle$ , for all  $j = 1, \dots, d$ . Then, we have,

$$\begin{aligned} \langle\psi|\psi'\rangle &= \frac{1}{\sqrt{d}} \sum_{j,l} \lambda_j \langle x_j|l\rangle \langle C \eta_l | C y_j \rangle \\ &= \frac{1}{\sqrt{d}} \sum_{j,l} \lambda_j \langle x_j|l\rangle \langle VC \eta_l | VC y_j \rangle \\ &= \frac{1}{\sqrt{d}} \sum_{j,l} \lambda_j \langle x_j|l\rangle \langle l|x_j\rangle = \frac{1}{\sqrt{d}} \sum_{j=1}^d \lambda_j, \end{aligned} \quad (2.35)$$

where we have used,  $C|\eta_l\rangle = V^\dagger|l\rangle$ , so that,  $|\eta_l\rangle = CV^\dagger|l\rangle$ .  $\square$

We now have the following Corollary to Lemma 2.4.13

**Corollary 2.4.14** *Let Schmidt Number( $|\psi\rangle$ ) =  $k$ . Then,*

$$\max\{|\langle\psi|\psi'\rangle|, |\psi'\rangle : \text{max. entangled state}\} \leq \sqrt{\frac{k}{d}}. \quad (2.36)$$

*Proof*  $|\psi\rangle = \sum_{j=1}^k \lambda_j |u_j\rangle |v_j\rangle$ , in Schmidt form, with  $\lambda_j = 0$ , if  $j \geq k + 1$ . By Lemma 2.4.13,

$$\begin{aligned} \max\{|\langle\psi|\psi'\rangle|, |\psi'\rangle : \text{max. entangled state}\} &\leq \frac{1}{\sqrt{d}} (\lambda_1 + \lambda_2 + \dots + \lambda_k) \\ &\leq \frac{1}{\sqrt{d}} k^{1/2} (\lambda_1^2 + \lambda_2^2 + \dots + \lambda_k^2)^{1/2} \\ &= \sqrt{\frac{k}{d}}. \end{aligned} \quad (2.37)$$

$\square$

It is now easy to prove Proposition 2.4.12.

*Proof* Assume, to the contrary, that the Schmidt number of  $\rho$  is less than  $k$ . Recall that such a state  $\rho$  can be written as a convex combination of pure states  $|\phi\rangle\langle\phi| \in \mathcal{S}_k$ , where the convex set  $\mathcal{S}_k$  was defined in Eq. (2.15), as follows:

$$\rho = \int_{\mathcal{S}_k} |\phi\rangle\langle\phi| \mu d(\phi).$$

If  $|\psi\rangle$  is maximally entangled, then, Corollary 2.4.14 implies

$$\langle\psi|\rho|\psi\rangle = \int_{S_k} |\langle\phi|\psi\rangle|^2 \mu d\phi \leq \frac{k}{d}. \quad (2.38)$$

Therefore, if  $\rho$  has the property that  $\langle\psi|\rho|\psi\rangle \geq \frac{k}{d}$  for any maximally entangled state  $|\psi\rangle$ ,

$$\text{Schmidt number}(\rho) \geq k.$$

□

Now, consider the Werner state  $\rho_F$  with  $F = \frac{k}{d}$  for some  $k \leq d$ . It is left as an exercise for the reader to check that such a state can be obtained by averaging the action of the unitary group  $\Pi(U)$  on an entangled state of Schmidt rank  $k$ .

**Exercise 2.4.15** Show that for  $|\psi_k\rangle = \frac{1}{\sqrt{k}} \sum_{i=1}^k |ii\rangle$ ,

$$\int (U \otimes \bar{U}) |\psi_k\rangle \langle\psi_k| (U \otimes \bar{U})^\dagger = \rho_{k/d}, \quad (2.39)$$

where  $\rho_{k/d}$  is the Werner state with  $F = \frac{k}{d}$ .

Clearly,  $\text{Schmidt Number}(\rho_{k/d}) \leq k$ . We will now prove that the Schmidt number of  $\rho_{k/d}$  is in fact exactly equal to  $k$ .

**Proposition 2.4.16** The state  $\rho_{k/d}$  defined in Eq. (2.39) has Schmidt number equal to  $k$ .

*Proof* We know from Exercise 2.4.10, that the state  $\rho_{k/d}$  satisfies

$$\max\{ \langle\psi|\rho_{k/d}|\psi\rangle, |\psi\rangle : \text{max. entangled state}\} = \frac{k}{d} \geq \frac{k-1}{d}.$$

Then, it is a direct consequence of Proposition 2.4.12 that,

$$\text{Schmidt number}(\rho) \geq k.$$

But we already know that  $\text{Schmidt number}(\rho) \leq k$ , thus proving that  $\text{Schmidt number}(\rho) = k$ . □

The final ingredient in proving Theorem 2.4.11 is the following property of Werner states with  $\frac{1}{d^2} \leq F \leq \frac{k}{d}$ .

**Proposition 2.4.17** For  $\frac{1}{d^2} \leq F \leq \frac{k}{d}$ , the Werner state  $\rho_F$  is a convex combination of  $\frac{1}{d^2}$  ( $I$  being the identity operator on  $\mathcal{H} \otimes \mathcal{H}$ ) and  $\rho_{k/d}$ .

*Proof* Let,

$$\rho_F = (1 - \theta) \frac{I}{d^2} + \theta \rho_{k/d}.$$

Recall that  $\rho_{k/d}$  is itself a convex combination of the maximally entangled state  $|\psi_0\rangle\langle\psi_0|$  and the state orthogonal to it, as stated in Eq. (2.30). Therefore,

$$\begin{aligned} \rho_F &= (1 - \theta) \frac{I}{d^2} + \theta \left[ \left(1 - \frac{k}{d}\right) \frac{I - |\psi_0\rangle\langle\psi_0|}{d^2 - 1} + \frac{k}{d} |\psi_0\rangle\langle\psi_0| \right] \\ &= \left[ \frac{1 - \theta}{d^2} + \theta \frac{1 - k/d}{d^2 - 1} \right] (I - |\psi_0\rangle\langle\psi_0|) + \left[ \frac{1 - \theta}{d^2} + \theta \frac{k}{d} \right] |\psi_0\rangle\langle\psi_0|. \end{aligned} \quad (2.40)$$

Comparing with the canonical form of the Werner state given in Eq. (2.30), we see that the parameter  $F$  corresponding to  $\rho_F$  is,

$$F = \frac{1 - \theta}{d^2} + \theta \frac{k}{d}.$$

In other words, the parameter  $\theta$  satisfies

$$\theta = \frac{F - \frac{1}{d^2}}{\frac{k}{d} - \frac{1}{d^2}}.$$

Clearly,  $0 \leq \theta \leq 1$  iff

$$\frac{1}{d^2} \leq F \leq \frac{k}{d}.$$

□

Finally, we note an important corollary of Proposition 2.4.17.

**Corollary 2.4.18** Schmidt number( $\rho_F$ )  $\leq k$  if  $\frac{1}{d^2} \leq F \leq \frac{k}{d}$ .

The proof of Theorem 2.4.11 now follows easily. Recall that the Werner state  $\rho_F$  for  $\frac{k-1}{d} \leq F \leq \frac{k}{d}$  satisfies (Exercise 2.4.10)

$$\langle\psi|\rho_F|\psi\rangle \geq \frac{k-1}{d}.$$

Then, Proposition 2.4.12 implies

$$\text{Schmidt number}(\rho_F) \geq k.$$



However, Proposition 2.4.17 and Corollary 2.4.18 imply that

$$\text{Schmidt number}(\rho_F) \leq k,$$

thus proving Theorem 2.4.11.

An interesting open question is to estimate the Schmidt number of the Werner state  $\rho_F$  for the parameter range  $0 \leq F \leq \frac{1}{d^2}$ .

## References

1. C.H. Bennett, S.J.Wiesner, Communication via one-and two-particle operators on Einstein-Podolsky-Rosen states. *Phys. Rev. Lett.* **69**, 2881 (1992)
2. K.R. Davidson, L.W. Marcoux, H. Radjavi, Transitive spaces of operators. *Integr. Equ. Oper. Theory* **61**, 187–210 (2008)
3. J. Harris, *Algebraic Geometry: A First Course* (Springer, New York, 1992)
4. M. Horodecki, P. Horodecki, R. Horodecki, Separability of mixed states: necessary and sufficient conditions. *Phys. Lett. A* **223**(1), 1–8 (1996)
5. R. Horodecki, P. Horodecki, M. Horodecki, K. Horodecki, Quantum entanglement. *Rev. Mod. Phys.* **81**, 865 (2009)
6. M.A. Nielsen, I.L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, 2000)
7. K.R. Parthasarathy, Extremal quantum states in coupled systems. *Ann. Inst. Henri Poincaré B Probab. Stat.* **41**(3), 257–268 (2005)
8. K.R. Parthasarathy, *Lectures on Quantum Computation, Quantum Error Correcting Codes and Information Theory* (Tata Institute of Fundamental Research, Mumbai, 2006)
9. K.R. Parthasarathy, *Coding Theorems of Classical and Quantum Information Theory* (Hindustan Book Agency, New Delhi, 2013)
10. A. Peres, Separability criterion for density matrices. *Phys. Rev. Lett.* **77**, 1413 (1996)
11. B.M. Terhal, P. Horodecki, Schmidt number for density matrices. *Phys. Rev. A* **61**(4), 040301 (2000)
12. R.F. Werner, Quantum states with Einstein-Podolsky-Rosen correlations admitting a hidden-variable model. *Phys. Rev. A* **40**, 4277 (1989)

# Chapter 3

## Operator Systems

**Abstract** In the first two sections of this chapter, which is based on lectures by Vern Paulsen, we will take a closer look at the results of Choi and how they explain some results on quantum error correction. In particular, we will look at the Knill-Laflamme result and Shor’s code. One novel aspect of our approach, is that we will introduce Douglas’ Factorization Theorem, which can be used to replace many calculations.

A quantum channel is always defined as a completely positive, trace preserving mapping. The natural setting to discuss completely positive mappings is an *operator system*. So our last three sections will be an introduction to some topics in the theory of operator systems that we believe are important for anyone interested in quantum information theory. The chapter concludes with some remarks on the equivalence of the Connes Embedding Problem, Kirchberg’s conjecture on  $C^*(\mathbb{F}_\infty)$  having the Weak Expectation Property, and other statements, including even a nuclearity assertion about some finite dimensional operator systems.

### 3.1 Theorems of Choi

#### Notations and Conventions

Vectors  $v \in \mathbb{C}^n$  will be treated as column vectors  $v = \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix}$ . In short,

$\mathbb{C}^n \equiv M_{n \times 1}$ . And for such a column vector  $v$ , its adjoint gives a row vector  $v^* := [\bar{\alpha}_1, \dots, \bar{\alpha}_n]$ .

*Bra and Ket Notations:* For  $v, w \in \mathbb{C}^n$ , we shall, following the physicists, write

$|v\rangle := v$  and  $\langle w| := w^*$ . Thus, if  $v = \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix}$  and  $w = \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_n \end{bmatrix}$ , then  $v^*w =$

$\sum_j \bar{\alpha}_j \beta_j = \langle v|w\rangle$  and  $wv^* = [\beta_j \bar{\alpha}_i] =: |w\rangle\langle v|$ . (In particular, our inner products will be linear in the second variable and conjugate linear in first.)

### 3.1.1 Douglas Factorization

We know by the spectral theorem that for positive semidefinite matrix  $P \in M_n^+$ , if  $u_1, \dots, u_r$  are non-zero eigenvectors ( $r = \text{rank}(P)$ ) with eigenvalues  $p_1, \dots, p_r$  and if  $v_i = \sqrt{p_i}u_i$ , then  $P = \sum_{i=1}^r v_i v_i^*$ . This is known as the *spectral decomposition*, but there are many ways to decompose a positive semi definite matrix as a sum of rank-1 matrices. Let  $P = \sum_{i=1}^m w_i w_i^*$  be another such. To find a relation between the  $v_i$ 's and  $w_i$ 's, we need the following proposition:

**Proposition 3.1.1 (Douglas' Factorization Theorem)** *Let  $A, B$  be two bounded operators on a Hilbert space  $H$  such that  $B^*B \leq A^*A$ . Then there exists a unique  $C \in B(R(A), R(B))$  such that  $\|C\| \leq 1$  and  $CA = B$  (where  $R(A)$  is range of  $A$ ).*

Observe that the conclusion in the above proposition of Douglas says loosely that  $B^*B \leq A^*A$  implies  $A$  "divides"  $B$  (in the above sense).

*Proof* Define  $C(Ah) = Bh$  for  $h \in H$ . Then well-definedness of the map  $C$  and the other assertions of the proposition follow from the inequality in the hypothesis.  $\square$

**Proposition 3.1.2** *Let  $P \in M_n^+$  with  $\text{rank}(P) = r$  and representations  $P = \sum_{i=1}^r v_i v_i^* = \sum_{i=1}^m w_i w_i^*$ . Then:*

1. *there is an isometric 'change of basis', that is, there exists a unique isometry  $U = [u_{i,j}]_{m \times r}$ , i.e.,  $U^*U = 1_r$ , such that  $w_i = \sum_{j=1}^r u_{i,j} v_j$  for all  $1 \leq i \leq m$ .*
2.  *$\text{span}\{v_1, \dots, v_r\} = \text{span}\{w_1, \dots, w_m\} = R(P)$ .*

*Proof* (1) Let  $V = [v_1 \vdots v_2 \vdots \dots \vdots v_r]$  and  $W = [w_1 \vdots w_2 \vdots \dots \vdots w_m]$ . Note that each vector  $v_i$  (or  $w_i$ ) is an  $n \times 1$  column vector, thus  $V \in M_{n \times r}$  and  $W \in M_{n \times m}$ .

The set  $\{v_i : i = 1, \dots, r\}$  is linearly independent since  $\text{rank}(P) = r$ ; in particular,  $V : \mathbb{C}^r \rightarrow \mathbb{C}^n$  is injective and  $V^* : \mathbb{C}^n \rightarrow \mathbb{C}^r$  is thus surjective.

Also  $P = VV^* = WW^*$ ; so, by Douglas' factorization, there exists a contraction  $U_1 : R(V^*) = \mathbb{C}^r \rightarrow R(W^*) \subset \mathbb{C}^m$  such that  $W^* = U_1 V^*$ . For the same reason, there also exists a contraction  $C : R(W^*) \rightarrow R(V^*) = \mathbb{C}^r$  such that  $V^* = C W^*$ . Hence  $(C U_1) V^* = V^*$ . As  $V^*$  is surjective, this shows that  $C U_1 = 1_r$  everywhere.

Hence  $C$  and  $U_1$  must be partial isometries and  $C U_1 = 1_r$  implies the injectivity of  $U_1$ ; we, therefore, see that  $U_1$  must be an isometry. The complex conjugate matrix  $U = \bar{U}_1$  is seen to satisfy the requirements of the proposition.

(2) follows from the fact that  $U$  as above is an isometry.  $\square$

### 3.1.2 Choi-Kraus Representation and Choi Rank

**Theorem 3.1.3 (Choi's First Theorem)** *Let  $\Phi : M_n \rightarrow M_d$  be linear. The following conditions on  $\Phi$  are equivalent:*

1.  $\Phi$  is completely positive.
2.  $\Phi$  is  $n$ -positive.
3.  $P_\Phi = (\Phi(E_{i,j})) \in M_n(M_d)^+$  where  $E_{i,j}$  are the standard matrix units of  $M_n$ .
4.  $\Phi(X) = \sum_{i=1}^r A_i X A_i^*$  for some  $r, d$  and  $A_i \in M_{d \times n}$  (Choi-Kraus representation).

*Proof* It is easy to see that (1)  $\Rightarrow$  (2)  $\Rightarrow$  (3) and (4)  $\Rightarrow$  (1). So we will only prove (3)  $\Rightarrow$  (4).

Let  $r = \text{rank}(P_\Phi)$ . Since,  $P_\Phi \in M_n(M_d)^+ = M_{nd}^+$ , as above, there exist vectors

$$v_i \in \mathbb{C}^{nd}, 1 \leq i \leq r \text{ such that } P_\Phi = \sum_{i=1}^r v_i v_i^*. \text{ Suppose } v_i = \begin{pmatrix} \alpha_1^i \\ \vdots \\ \alpha_n^i \end{pmatrix} \in \mathbb{C}^{nd},$$

$$\text{where each } \alpha_j^i = \begin{pmatrix} \alpha_j^i(1) \\ \vdots \\ \alpha_j^i(d) \end{pmatrix} \in \mathbb{C}^d.$$

Let  $A_i = [\alpha_1^i \ \vdots \ \dots \ \vdots \ \alpha_n^i]_{d \times n}$ . It is now easy to check that  $\Phi(E_{i,j}) = \sum_l \alpha_l^i (\alpha_l^j)^* = \sum_{l=1}^n A_i E_{i,j} A_l^*$  for all  $1 \leq i, j \leq n$ . Hence (4) holds.  $\square$

*Remark 3.1.4* If we write  $P_\Phi$  as  $\sum_{i=1}^m w_i w_i^*$ , where  $w_i = \begin{pmatrix} \beta_1^i \\ \vdots \\ \beta_n^i \end{pmatrix} \in \mathbb{C}^{nd}$ , with  $\beta_j^i \in$

$\mathbb{C}^d$ , then  $\Phi(X) = \sum_{i=1}^m B_i X B_i^*$ , where  $B_i := [\beta_1^i \ \vdots \ \dots \ \vdots \ \beta_n^i]_{d \times n}$ . Conversely, if

$$\Phi(X) = \sum_{i=1}^m B_i X B_i^* \text{ for some } B_i = [\beta_1^i \ \vdots \ \dots \ \vdots \ \beta_n^i]_{d \times n}, \text{ and if we let } w_i = \begin{pmatrix} \beta_1^i \\ \vdots \\ \beta_n^i \end{pmatrix},$$

then  $P_\Phi = \sum_{i=1}^m w_i w_i^*$ .

We now apply the Choi-Kraus result to characterize quantum channels.

**Proposition 3.1.5** *Let  $\mathcal{E} : M_n \rightarrow M_n$  be CPTP (i.e., CP and trace preserving). Then there exist  $E_i \in M_n$ ,  $1 \leq i \leq r$  satisfying  $\sum_{i=1}^r E_i^* E_i = 1_n$  such that  $\mathcal{E}(X) = \sum_{i=1}^r E_i X E_i^*$  for all  $X \in M_n$ .*

*Proof* Since  $\mathcal{E}$  is CP, from Theorem 3.1.3, we know of the existence of matrices  $E_i \in M_n$  such that  $\mathcal{E}$  is decomposed as a sum as above. However,  $\mathcal{E}$  is trace preserving  $\Rightarrow \text{tr}(\mathcal{E}(X)) = \text{tr}(X) \ \forall X \Rightarrow \text{tr}(\sum_{i=1}^r E_i X E_i^*) = \text{tr}(X) \ \forall X \Rightarrow \text{tr}(\sum_{i=1}^r E_i^* E_i X) = \text{tr}(X) \ \forall X \Rightarrow \langle \sum_{i=1}^r E_i^* E_i, X^* \rangle = \langle 1_n, X^* \rangle, \ \forall X$  and hence  $\sum_{i=1}^r E_i^* E_i = 1_n$ .  $\square$

*Remark 3.1.6* It is easy to see that conversely,  $\sum_{i=1}^r E_i^* E_i = 1_n \Rightarrow \mathcal{E}$  is trace preserving.

**Definition 3.1.7 (Choi Rank of a CP Map)** Let  $\Phi : M_n \rightarrow M_d$  be CP. Then, by the above theorem,  $\Phi(X) = \sum_{i=1}^q B_i X B_i^*$  for some matrices  $B_i \in M_{d,n}$ ,  $1 \leq i \leq q$ . The Choi rank of  $\Phi$  is given by  $\text{cr}(\Phi) := \min\{q : \Phi(X) = \sum_{i=1}^q B_i X B_i^*\}$ .

**Proposition 3.1.8 (Choi)** In the above set up,  $\text{cr}(\Phi) = \text{rank}(P_\Phi)$ .

*Proof* This follows from Theorem 3.1.3, Remark 3.1.4 and Proposition 3.1.2.  $\square$

**Theorem 3.1.9 (Choi's Second Theorem)** Let  $\Phi \in CP(M_n, M_d)$  with  $\text{cr}(\Phi) = r$ . Suppose  $\Phi(X) = \sum_{i=1}^r A_i X A_i^* = \sum_{i=1}^m B_i X B_i^*$  are two Choi-Kraus representations of  $\Phi$ . Then there exists a unique matrix  $U = (u_{i,j}) \in M_{m \times r}$  such that  $U^* U = 1_r$ ,  $B_i = \sum_{j=1}^r u_{i,j} A_j$  and  $\text{span}\{A_1, \dots, A_r\} = \text{span}\{B_1, \dots, B_m\}$ .

*Proof* This follows from Proposition 3.1.2, Theorem 3.1.3 and Remark 3.1.4.  $\square$

### Example: Binary Case Quantum Error Detection/Correction

This is an introductory binary example to motivate the quantum error correction/detection of the next section. Let us take a binary string of 0s and 1s of length, say 5, e.g.  $s = (0, 1, 0, 1, 1) \in \mathbb{Z}_2^5$ . We want to transmit this. Some errors may occur in transmission. We want to detect/correct that error.

One way to detect/correct error is to encode the given string into a larger vector. One famous binary error correcting codes is the **Majority Rule Code**.

We start with our original vector  $s$  of length  $r$  (here  $r = 5$ ) and encode it within a vector of length  $nr$  for some odd  $n$ , where each digit gets repeated  $n$  times consecutively.

Say  $n = 3$  and  $s = (0, 1, 0, 1, 1)$  then this vector gets encoded into

$$s_1 = (0, 0, 0; 1, 1, 1; 0, 0, 0; 1, 1, 1; 1, 1, 1)$$

(note that the encoded vectors form a five-dimensional subspace of a 15-dimensional space).

After transmission, suppose the output vector turns out to be, say

$$s_2 = (1, 0, 0; 1, 1, 1; 0, 0, 1; 0, 0, 1; 1, 1, 0).$$

We decode/recover the string by choosing the digit (0 or 1) that appears as a majority among each block of  $n (= 3)$  consecutive digits. Thus, in each block of three consecutive digits, the majority rules. Here the recovered string will be  $s_3 = (0, 1, 0, 0, 1)$ .

Schematically:  $s \xrightarrow{\text{encode}} s_1 \xrightarrow{\text{transmit}} \dots \xrightarrow{\text{receive}} s_2 \xrightarrow{\text{decode}} s_3$ .

Note that there is one error and in fact to have one error in the output there has to be at least two (i.e., more than  $n/2$ ) errors within some block of  $n$  digits. Knowing probabilities of those errors will help to understand how effective a code is.

Alternatively, instead of encoding cleverly, one could pick a clever five-dimensional subspace of the 15-dimensional space and then any embedding of  $\mathbb{Z}_2^5$  onto that subspace would be the encoding..

Note that this code requires that we “clone” each digit three times. Thus, it violates the “no cloning” rule and could not be implemented on a quantum machine.

In the next section, we will present Shor’s code, which is a quantum code that is robust to many types of errors.

### 3.2 Quantum Error Correction

If we assume that any errors that occur must also be the result of some quantum event, then it is natural to assume that errors are also the result of the action of a CPTP map acting on the states. Thus, an *error operator* will be a CPTP map  $\mathcal{E}$ . The usual strategy/protocol in quantum error correction is the following:

As in the above example, we don’t expect to be able to correct all errors that occur. If  $v \in V$  is a state in a ‘protected’ subspace  $V$ , we want to correct  $\mathcal{E}(|v\rangle\langle v|)$ , which is the error that has happened to the state. To do this, we seek a recovery operator  $\mathcal{R} : M_n \xrightarrow{CPTP} M_n$  such that  $\mathcal{R}(\mathcal{E}(|v\rangle\langle v|)) = |v\rangle\langle v|, \forall v \in V$ .

The approach appears somewhat naive since it appears that it assumes that we know the error explicitly. However using Choi’s theorem it turns out, that correcting  $V$  against one assumed error leads to its correction against a whole family of errors. I believe that this is the key element of Knill-Laflamme error correction.

#### 3.2.1 Applications of Choi’s Theorems to Error Correction

**Theorem 3.2.1 (Knill-Laflamme)** *Let  $V \subset \mathbb{C}^n$  be a subspace and  $P : \mathbb{C}^n \rightarrow V$  be its orthogonal projection. Let  $\mathcal{E} : M_n \rightarrow M_n$  be CPTP (an error map) given by  $\mathcal{E}(X) = \sum_{i=1}^m E_i X E_i^*$  for all  $X \in M_n$ . Then there exists a (recovery) CPTP map  $\mathcal{R} : M_n \rightarrow M_n$  such that  $\mathcal{R}(\mathcal{E}(PXP)) = PXP, \forall X \in M_n$  if, and only if,  $P E_i^* E_j P = \alpha_{i,j} P$  for some  $\alpha_{i,j} \in \mathbb{C}, 1 \leq i, j \leq m$ .*

*Proof* ( $\Rightarrow$ ) Suppose a recovery map  $\mathcal{R}$  has a representation  $\mathcal{R}(W) = \sum_{l=1}^q A_l W A_l^*$  for all  $W \in M_n$  with  $\sum A_l^* A_l = 1$ . Since  $\mathcal{R}$  and  $\mathcal{E}$  are both CPTP, so is  $\mathcal{R} \circ \mathcal{E}$ . Further, the equality  $(\mathcal{R} \circ \mathcal{E})(PXP) = \mathcal{R}(\mathcal{E}(PXP)) = \sum_{i=1}^m \sum_{l=1}^q (A_l E_i P) X (P E_i^* A_l^*) = PXP$ , gives two Choi-Kraus representations for  $\mathcal{R} \circ \mathcal{E}$ , and  $PXP$  is of minimal length. Hence, by Theorem 3.1.9, there

exists a unique isometry  $U = [\beta_{ij}]_{mq \times 1}$  such that  $A_l E_l P = \beta_{li} P$ . Since

$$U = \begin{pmatrix} \beta_{11} \\ \vdots \\ \beta_{1m} \\ \beta_{21} \\ \vdots \\ \beta_{2m} \\ \vdots \\ \beta_{q1} \\ \vdots \\ \beta_{qm} \end{pmatrix} \text{ is an isometry, we have } \sum_{l,i} |\beta_{li}|^2 = U^* U = 1. \text{ In particular,}$$

$(PE_i^* A_l^*)(A_l E_j P) = (\bar{\beta}_{li} P)(\beta_{lj} P) = \bar{\beta}_{li} \beta_{lj} P$  and, using  $\sum_l A_l^* A_l = 1_n$ , we have

$$\sum_{l=1}^q \bar{\beta}_{li} \beta_{lj} P = \sum_{l=1}^q PE_i^* (A_l^* A_l) E_j P = PE_i^* E_j P.$$

Simply take  $\alpha_{i,j} = \sum_l \bar{\beta}_{li} \beta_{lj}$  for the conclusion.

( $\Leftarrow$ ) We may clearly assume  $P \neq 0$  so  $tr(P) > 0$ , as the Theorem is vacuously true in the contrary case! Suppose there exist  $\alpha_{i,j} \in \mathbb{C}$ ,  $1 \leq i, j \leq m$  such that  $PE_i^* E_j P = \alpha_{i,j} P$  for all  $1 \leq i, j \leq m$ . Then  $(\alpha_{i,j}) = (PE_i^* E_j P) = \begin{pmatrix} PE_1^* \\ \vdots \\ PE_m^* \end{pmatrix} (E_1 P, \dots, E_m P) \geq 0$ . This implies  $(\alpha_{i,j}) \in M_m^+$  (since  $[\alpha_{i,j}] = \frac{1}{tr_n(P)}(id \otimes tr_n)([\alpha_{i,j}] \otimes P)$ ); also,  $\sum_i \alpha_{i,i} P = \sum_i PE_i^* E_i P = P^2 = P \Rightarrow tr_m([\alpha_{i,j}]) = \sum_i \alpha_{i,i} = 1$ , i.e.,  $[\alpha_{ij}]$  is a density matrix.

Thus,  $(\alpha_{i,j})$  is unitarily diagonalizable, i.e., there exists a unitary  $U = (u_{i,j}) \in M_m$  such that  $U(\alpha_{i,j})U^* = D = \text{diag}(d_{11}, \dots, d_{mm})$ , with  $d_{ii} \geq 0 \forall i$  and  $\sum_{i=1}^m d_{ii} = 1$ . Set  $F_i = \sum_{j=1}^m \bar{u}_{i,j} E_j$ ,  $1 \leq i \leq m$ . Hence, by Choi (or by direct calculation),  $\sum_{i=1}^m F_i X F_i^* = \mathcal{E}(X)$  for all  $X \in M_n$ . Also, note that

$$\begin{aligned} PF_i^* F_j P &= P \left( \sum_{k,l} u_{i,k} E_k^* \bar{u}_{j,l} E_l \right) P \\ &= \sum_{k,l} u_{i,k} \bar{u}_{j,l} P E_k^* E_l P \\ &= \sum_{k,l} u_{i,k} \alpha_{k,l} \bar{u}_{j,l} P \\ &= d_{ij} P. \end{aligned}$$

So we see that if we define

$$V_i = \begin{cases} 0 & \text{if } d_{ii} = 0 \\ \frac{1}{\sqrt{d_{ii}}} F_i P & \text{otherwise} \end{cases},$$

we find that the  $V_i$ 's are partial isometries with  $V_i^* V_j = \delta_{ij} P \forall i, j$  (i.e., all non-zero ones among them having initial space equal to  $V$ ) and with  $V_i V_i^* = F_i(V)$  (i.e., with pairwise orthogonal final spaces). Hence we see that  $R = \sum_i V_i V_i^* = \sum \frac{1}{d_{ii}} F_i P F_i^*$  is an orthogonal projection. Let  $Q = 1 - R$ .

Define  $\mathcal{R} : M_n \rightarrow M_n$  by  $\mathcal{R}(X) = \sum V_i^* X V_i + Q X Q$ . Clearly,  $\mathcal{R}$  is a unital CPTP map.

We want to show that  $\mathcal{R}(\mathcal{E}(PXP)) = PXP \forall X \in M_n$ . To see this, it is enough to show that for  $v \in V$ ,  $\mathcal{R}(\mathcal{E}(vv^*)) = vv^*$  (since  $\{PXP : X \in M_n\} =$  all matrices living on  $V$  while  $\{vv^* : v \in V\} =$  all rank-1 projections living on  $V$ ). For this, compute thus:

$$\begin{aligned} \mathcal{R}(\mathcal{E}(vv^*)) &= \mathcal{R}(\mathcal{E}(Pvv^*P)) \\ &= \sum_i V_i^* \left( \sum_j F_j P v v^* P F_j^* \right) V_i + Q \left( \sum_j F_j P v v^* P F_j^* \right) Q \\ &= \sum_{i,j} V_i^* F_j P v v^* P F_j^* V_i + \sum_j Q F_j P v v^* P F_j Q \\ &= \sum_{i,j} d_{ij} V_i^* V_j v v^* V_i + \sum_j d_{jj} Q V_j P v v^* P V_j Q \\ &= \sum_j d_{jj} P v v^* P + 0 \\ &= vv^*, \end{aligned}$$

as desired.  $\square$

**Theorem 3.2.2** *Let  $\mathcal{R}$  be the recovery operator constructed as in Theorem 3.2.1 and  $\tilde{\mathcal{E}} : M_n \rightarrow M_n$  be another error operator admitting a representation  $\tilde{\mathcal{E}}(X) = \sum_{i=1}^p \tilde{E}_i X \tilde{E}_i^*$  with  $\tilde{E}_i \in \text{span}\{E_1, \dots, E_m\}$ ,  $1 \leq i \leq p$ . Then, also  $\mathcal{R}\tilde{\mathcal{E}}(PXP) = PXP$ ,  $\forall X \in M_n$ .*

*Proof* We have  $\tilde{E}_i \in \text{span}\{E_1, \dots, E_m\}$ ,  $1 \leq i \leq p$  and since  $\tilde{\mathcal{E}}$  is CPTP, they must satisfy  $\sum \tilde{E}_i^* \tilde{E}_i = 1$ . Recall that—with the  $F_i$  as in the proof of Theorem 3.2.1—we may deduce from Theorem 3.1.9 that  $\text{span}\{E_1, \dots, E_m\} = \text{span}\{F_1, \dots, F_m\}$  and  $P F_i^* F_j P = \delta_{i,j} d_{ii} P$ ,  $1 \leq i, j \leq m$ ; so,

$$V_i^* F_j P = \delta_{i,j} \sqrt{d_{ii}} P. \quad (3.1)$$



Write  $\tilde{E}_i = \sum_{l=1}^m \beta_{i,l} F_l$ . Then,  $1 = \sum_k \tilde{E}_k^* \tilde{E}_k = \sum_k (\sum_{l=1}^m \bar{\beta}_{k,l} F_l^*) (\sum_{j=1}^m \beta_{k,j} F_j)$ . So,

$$\begin{aligned} P &= P1P = P \left( \sum_k \left( \sum_{l=1}^m \beta_{k,l} F_l^* \right) \left( \sum_{j=1}^m \beta_{k,j} F_j \right) \right) P \\ &= \sum_k \sum_{l,j} \bar{\beta}_{k,l} \beta_{k,j} P F_l^* F_j P \\ &= \sum_k \sum_{j=1}^m |\beta_{k,j}|^2 d_{jj} P. \end{aligned}$$

Hence,

$$\sum_k \sum_{j=1}^m |\beta_{k,j}|^2 d_{jj} = 1. \quad (3.2)$$

Hence,

$$\begin{aligned} \mathcal{R}(\tilde{\mathcal{E}}(PXP)) &= \sum_{i,j} V_i^* \tilde{E}_j PXP \tilde{E}_j^* V_i \\ &= \sum_{i,j,k,l} \bar{\beta}_{lj} \beta_{kj} V_i^* F_l PXP F_k^* V_i \\ &= \sum_{i,j} |\beta_{ij}|^2 d_{ii} PXP \quad (\text{by Eq. (3.1)}) \\ &= PXP \quad \forall X \in M_n \quad (\text{by Eq. (3.2)}). \end{aligned}$$

□

### 3.2.2 Shor's Code: An Example

We consider the 'protected subspace'  $V \subset \mathbb{C}^2 \otimes \cdots \otimes \mathbb{C}^2$  (9 copies) with  $\dim(V) = 2$  given by

$$V = \text{span}\{|0_L\rangle, |1_L\rangle\},$$

where  $0_L = \frac{1}{2\sqrt{2}}((|000\rangle + |111\rangle) \otimes (|000\rangle + |111\rangle) \otimes (|000\rangle + |111\rangle))$  and  $1_L = \frac{1}{2\sqrt{2}}((|000\rangle - |111\rangle) \otimes (|000\rangle - |111\rangle) \otimes (|000\rangle - |111\rangle))$ . (Notation: Fixing an orthonormal basis  $\{|0\rangle, |1\rangle\}$  for  $\mathbb{C}^2$ , we write  $|000\rangle$  for  $|0\rangle \otimes |0\rangle \otimes |0\rangle$  and so on.)

We consider the Pauli basis of  $\mathbb{C}^2 \otimes \dots \otimes \mathbb{C}^2$  (9 copies) constructed as follows:

Take the basis of  $M_2$  (which is orthonormal with respect to the normalised trace-inner-product) consisting of  $X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ ,  $Y = \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix}$ ,  $Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$  and  $1_2$ . (These are regarded as maps on  $\mathbb{C}^2$  with respect to the orthonormal basis  $\{|0\rangle, |1\rangle\}$ .)

1-Pauli elements: For  $i = 1, \dots, 9$ , these are the  $2^9 \times 2^9$  unitary self-adjoint matrices defined by

$$\begin{aligned} 1 &= 1_2 \otimes 1_2 \otimes \dots \otimes 1_2 \\ X_i &= 1_2 \otimes 1_2 \otimes \dots \otimes X \otimes \dots \otimes 1_2 \text{ (} X \text{ at } i^{\text{th}} \text{ position)} \\ Y_i &= 1_2 \otimes 1_2 \otimes \dots \otimes Y \otimes \dots \otimes 1_2 \text{ (} Y \text{ at } i^{\text{th}} \text{ position)} \\ Z_i &= 1_2 \otimes 1_2 \otimes \dots \otimes Z \otimes \dots \otimes 1_2 \text{ (} Z \text{ at } i^{\text{th}} \text{ position)}. \end{aligned}$$

Let us list the above 1-Paulis as  $U_1, \dots, U_{28}$ . Define  $\mathcal{E} : M_{2^9} \rightarrow M_{2^9}$  by  $\mathcal{E}(X) = \frac{1}{28} \sum_{i=1}^{28} U_i X U_i^*$ . Then, it is easily seen that  $\mathcal{E}$  is a CPTP map (being an average of \*-automorphisms).

**Proposition 3.2.3** *With this notation, we have that*

$$V, X_1V, \dots, X_9V, Y_1V, \dots, Y_9V, Z_1V$$

are all mutually orthogonal and  $Z_{3k+i}|_V = Z_{3k+j}|_V$  for  $k = 0, 1, 2$ , and  $0 \leq i, j < 3$ .

*Proof* Exercise. □

- Notice that  $P U_i^* U_j P \in \{0, P\} \forall 1 \leq i, j \leq 28$  and hence Theorem 3.2.1 ensures the existence of a recovery operator  $\mathcal{R}$  satisfying  $\mathcal{R} \circ \mathcal{E}(PXP) = PXP \forall X \in M_n$ .

By Theorem 3.2.2 above, for this protected space  $V$  and error map  $\mathcal{E}$ , we have  $\mathcal{R}(\tilde{\mathcal{E}}(PXP)) = PXP$  for any error map  $\tilde{\mathcal{E}} : M_n \rightarrow M_n$  given by  $\tilde{\mathcal{E}}(X) = \sum \tilde{E}_i X \tilde{E}_i^*$ , where  $\tilde{E}_i \in \text{span}\{1\text{-Pauli basis}\}$ .

- The 1-Paulis contain in their span any operator of the form  $1 \otimes \dots \otimes 1 \otimes A \otimes 1 \dots \otimes 1$ . So while the Shor code may not fix all errors, it does fix all errors in  $\text{span}\{1\text{-Paulis}\}$ . Thus, if each term in the error operator acts on only one of the qubits, then this subspace will be protected from this error and the decoding map will recover the original encoded qubit.

*Remark 3.2.4* See the work of David Kribs, et al. for various generalizations of the Knill-Laflamme theory, including infinite dimensional versions of this theory.

A more refined version of the Knill-Laflamme theory than we have stated is to consider protected *operator subsystems* of the matrices and encodings of states into such subsystems.

In the next lecture we will introduce this concept.

### 3.3 Matrix Ordered Systems and Operator Systems

Recall that a typical quantum channel on  $M_n$  looks like  $\mathcal{E}(X) = \sum_{i=1}^r E_i X E_i^*$  for some matrices  $E_i \in M_n$ , which are not unique. However, given any such representation of  $\mathcal{E}$ , the space  $\mathcal{S} := \text{span} \{E_j : 1 \leq j \leq r\}$  remains the same. Moreover, the space  $\mathcal{S}$  contains 1 and is closed under taking adjoints.

Duan, Severini and Winter have argued that various concepts of quantum capacities of the channel  $\mathcal{E}$  really only depend on this subspace  $\mathcal{S}$ , i.e., if two channels generate the same subspace then their capacities should be the same. Thus, capacities are naturally functions of such subspaces.

Moreover, in the extensions of the Knill-Laflamme theory, it is exactly such subspace that are the protected subspaces, i.e., the subspaces that one wants to encode states into so that they can be recovered after the actions of some error operators.

A subspace of  $M_n$  (or more generally,  $B(\mathcal{H})$ ) that contains 1 and is closed under the taking of adjoints is called an *operator system*. These are also the natural domains and ranges of completely positive maps.

Thus, the concept of an operator system plays an important role in the study of completely positive maps and, in particular, in QIT. For this reason we want to introduce their general theory and axiomatic definitions.

Finally, when we study  $M_n = B(\mathbb{C}^n)$  we know that the positive operators of rank one, represent the states of the underlying space and that positive operators of trace one represent the mixed states. But when we focus on a more general operator system, what exactly is it the states of? One viewpoint is to just regard it as a restricted family of states of the underlying space. But this is very problematical since many operator subsystems of  $M_n$  have no rank one positives and others that have plenty of rank one positives, still have trace one positives that cannot be written as sums of rank ones! The correct answer to what is an operator system the states of involves introducing the concept of (*ordered*) *duals*.

Motivated by these issues and some structural properties of  $M_n$ , we introduce a bunch of abstract definitions.

**Definition 3.3.1** A  $*$ -vector space is a complex vector space  $\mathcal{V}$  with a map  $*$  :  $\mathcal{V} \rightarrow \mathcal{V}$  satisfying

1.  $(v + w)^* = v^* + w^*$ ;
2.  $(\lambda v)^* = \bar{\lambda} v^*$ ; and
3.  $(v^*)^* = v$

for all  $v, w \in \mathcal{V}$ ,  $\lambda \in \mathbb{C}$ . The self adjoint elements of such a space is denoted by

$$\mathcal{V}_h = \{v \in \mathcal{V} : v = v^*\}.$$

- As usual, we have a *Cartesian decomposition* in a  $*$ -vector space  $\mathcal{V}$  given by  $v = \frac{v+v^*}{2} + i \frac{v-v^*}{2i}$  for all  $v \in \mathcal{V}$  and we call these the *real* and *imaginary* parts of  $v$ .

- Given a  $*$ -vector space  $\mathcal{V}$  and  $n \geq 1$ , its  $n$ th-amplification  $M_n(\mathcal{V})$ , which is just the set of  $n \times n$  matrices with entries from  $\mathcal{V}$  inherits a canonical  $*$ -vector space structure and is naturally identified with  $M_n \otimes \mathcal{V}$ .

**Definition 3.3.2** A matrix ordering on a  $*$ -vector space  $\mathcal{V}$  is a collection  $\mathcal{C}_n \subset M_n(\mathcal{V})_h$ ,  $n \geq 1$  satisfying:

1.  $\mathcal{C}_n$  is a cone for all  $n \geq 1$ ;
2.  $\mathcal{C}_n \cap (-\mathcal{C}_n) = (0)$  for all  $n \geq 1$ ; and
3.  $APA^* \in \mathcal{C}_k$  for all  $A \in M_{k,n}$ ,  $P \in \mathcal{C}_n$ ,  $k, n \geq 1$ .

A  $*$ -vector space  $\mathcal{V}$  with a matrix ordering as above is called a matrix ordered space.

*Remark 3.3.3* Some authors also add the following axiom in the definition of a matrix ordering:

$$(\tilde{4}) \quad \mathcal{C}_n - \mathcal{C}_n = M_n(\mathcal{V})_h.$$

However, we will abstain from its use because of reasons that will get clear while discussing ‘dual of an operator system’.

**Exercise 3.3.4** Let  $P \in \mathcal{C}_n$  and  $Q \in \mathcal{C}_k$ . Then  $\begin{bmatrix} P & 0 \\ 0 & Q \end{bmatrix} \in \mathcal{C}_{n+k}$ .

*Example 3.3.5* 1.  $\mathcal{V} = B(H)$  with usual adjoint structure and  $\mathcal{C}_n := M_n(B(H))^+ = B(H \otimes \mathbb{C}^n)^+$ ,  $n \geq 1$  is matrix ordered.

2. Any subspace  $\mathcal{V} \subset B(H)$  such that  $\mathcal{V}$  is closed under taking adjoints provides  $\mathcal{V}$  with the natural induced matrix ordering  $\mathcal{C}_n := M_n(B(H))^+ \cap M_n(\mathcal{V})$ ,  $n \geq 1$ .

**Definition 3.3.6** An operator system is a subspace  $\mathcal{S} \subset B(H)$  that is closed under taking adjoints and contains the unit  $e_{\mathcal{S}} := id_H$ , for some Hilbert space  $H$ , together with the matrix ordering given in Example 3.3.5(2).

- Usually, whenever the matrix ordering is clear from the context, we simply write  $M_n(\mathcal{V})^+$  for  $\mathcal{C}_n$ ,  $n \geq 1$ .

**Definition 3.3.7** Given matrix ordered spaces  $\mathcal{V}$  and  $\mathcal{W}$ , a linear map  $\varphi : \mathcal{V} \rightarrow \mathcal{W}$  is said to be  $n$ -positive if its  $n$ th-application  $\varphi^{(n)} : M_n(\mathcal{V}) \rightarrow M_n(\mathcal{W})$ ,  $[v_{ij}] \mapsto [\varphi(v_{ij})]$  is positive, i.e.,  $\varphi^{(n)}(M_n(\mathcal{V})^+) \subset M_n(\mathcal{W})^+$ .  $\varphi$  is said to be completely positive (in short, CP) if  $\varphi$  is  $n$ -positive for all  $n \geq 1$ .

- Clearly, if we include Axiom  $(\tilde{4})$  in the axioms of matrix ordering, then every CP map is  $*$ -preserving.

**Definition 3.3.8** Two matrix ordered spaces  $\mathcal{V}$  and  $\mathcal{W}$  are said to be completely order isomorphic if there is a completely positive linear isomorphism  $\varphi : \mathcal{V} \rightarrow \mathcal{W}$  such that  $\varphi^{-1}$  is also completely positive.

Given two operator systems  $\mathcal{S}_1, \mathcal{S}_2$  on possibly different Hilbert spaces, we identify them as the “same” operator system when they are completely order isomorphic via a unital complete order isomorphism.

### 3.3.1 Duals of Matrix Ordered Spaces

Let  $\mathcal{V}$  be a matrix ordered space. Let  $\mathcal{V}^d$  be the space of linear functionals on the vector space  $\mathcal{V}$ .

*\*-structure:* For each  $f \in \mathcal{V}^d$  define  $f^* \in \mathcal{V}^d$  by  $f^*(v) = \overline{f(v^*)}$ ,  $v \in \mathcal{V}$ . This makes  $\mathcal{V}^d$  into a \*-vector space.

*Matrix ordering:* Given a matrix of linear functionals  $[f_{ij}] \in M_n(\mathcal{V}^d)$ , identify it with the map  $\Phi : \mathcal{V} \rightarrow M_n$  given by  $\Phi(v) = [f_{ij}(v)]$ ,  $v \in \mathcal{V}$ . Let

$$\mathcal{C}_n = \{[f_{ij}] \in M_n(\mathcal{V}^d) : \Phi : \mathcal{V} \rightarrow M_n \text{ is CP}\}.$$

Then,  $\mathcal{V}^d$  together with above cones forms a matrix ordered space.

This matrix ordered space is what is meant by the *matrix-ordered dual* of  $\mathcal{V}$ .

- There are many other ways of making  $\mathcal{V}^d$  into a matrix ordered space. However, the above structure has better compatibility with respect to some important operations on matrix ordered spaces.

*Remark 3.3.9* In general, if we require  $\mathcal{V}$  to also satisfy axiom  $(\tilde{4})$  in the definition of a matrix ordering, then it can still be the case that  $\mathcal{V}^d$  does not satisfy  $(\tilde{4})$ .

An immediate compatibility of the above dual structure is seen in the following:

**Proposition 3.3.10** *Let  $\varphi : \mathcal{V} \rightarrow \mathcal{W}$  be a CP map between two matrix ordered spaces. Then the usual dual map  $\varphi^d : \mathcal{W}^d \rightarrow \mathcal{V}^d$  is also CP.*

*Proof* Let  $n \geq 1$  and  $[f_{ij}] \in M_n(\mathcal{W}^d)^+$ . Then  $(\varphi^d)^{(n)}([f_{ij}]) = [f_{ij} \circ \varphi]$ . Now,  $[f_{ij} \circ \varphi] : \mathcal{V} \rightarrow M_n$  is given by  $v \mapsto [f_{ij} \circ \varphi(v)]$ , which being a composite of the CP maps  $\mathcal{V} \xrightarrow{\varphi} \mathcal{W} \xrightarrow{[f_{ij}]} M_n$  is again CP. Thus,  $(\varphi^d)^{(n)}([f_{ij}]) \in M_n(\mathcal{V}^d)^+$ . In particular,  $\varphi^d$  is CP.  $\square$

Consider the matrix algebra operator system  $\mathcal{V} = M_p = \mathcal{L}(\mathbb{C}^p)$ . Let  $\{E_{ij} : 1 \leq i, j \leq p\}$  be the system of matrix units for  $\mathcal{V}$ . Via this basis, we can in fact identify  $\mathcal{V}^d$  with  $\mathcal{V}$  itself. Formally, let  $\{\delta_{ij}\} \subset \mathcal{V}^d$  be the dual basis for  $\{E_{ij}\}$ . Given  $A = [a_{ij}] \in \mathcal{V}$ , define  $f_A \in \mathcal{V}^d$  by  $f_A = \sum_{ij} a_{ij} \delta_{ij}$ . Thus,  $a_{ij} = f_A(E_{ij})$  and we see that  $A$  is the usual “density” matrix of the linear functional  $f_A$ . Note that, for  $B = [b_{ij}] \in \mathcal{V}$ , we have  $f_A(B) = \sum_{ij} a_{ij} b_{ij} = \text{tr}_p(A^* B)$ . Clearly,  $f_A(B) \geq 0$  for all  $B \geq 0$  if and only if  $A \geq 0$ . Define  $\Gamma : \mathcal{V} \rightarrow \mathcal{V}^d$  by  $\Gamma(A) = f_A$ . Then, we have the following:

**Theorem 3.3.11** ([14, Theorem 6.2]) *The map  $\Gamma : M_p(\mathbb{C}) \rightarrow M_p(\mathbb{C})^d$  as constructed above is a complete order isomorphism.*

A natural question to ask at this stage would be whether any other basis for  $M_p$  works equally well or not? And, quite surprisingly, the answer is not very clear!

Let  $\mathcal{B} = \{B_{rs} : 1 \leq r, s \leq p\}$  be any other basis for  $M_p$  and  $\{\eta_{rs} : 1 \leq r, s \leq p\}$  its dual basis. Define  $\Gamma_{\mathcal{B}} : M_p \rightarrow (M_p)^d$  by  $\Gamma_{\mathcal{B}}(A) = \sum_{rs} a_{rs} \eta_{rs}$ , where  $A = \sum_{rs} a_{rs} B_{rs}$ . Then, it is not difficult to find a basis  $\mathcal{B}$  such that  $\Gamma_{\mathcal{B}}$  is not a complete order isomorphism. However, it will be interesting to answer the following:

- Question 3.3.12** 1. *What are the necessary and sufficient conditions on the basis  $\mathcal{B}$  such that the map  $\Gamma_{\mathcal{B}}$  as above is a complete order isomorphism?*  
 2. *If the map  $\Gamma_{\mathcal{B}}$  is positive, then is it automatically completely positive?*

See [13] for some work on this problem.

Note that, under the above complete order isomorphism  $\Gamma$ , we have  $\Gamma(1) = tr_p$ . We will soon see that 1 and  $tr_p$  have some interesting significance on the structures of  $M_p$  and  $(M_p)^d$ , respectively.

### 3.3.2 Choi-Effros Theorem

Arveson introduced the concept of an operator system in 1969. Choi-Effros were the first to formally axiomatize the theory. Their axiomatic characterization follows.

**Theorem 3.3.13** ([1, Theorem 4.4]) *Let  $\mathcal{V}$  be a matrix ordered space with an element  $e \in \mathcal{V}_h$  satisfying:*

1. *For each  $n \geq 1$  and  $H \in M_n(\mathcal{V})_h$ , there exists an  $r > 0$  such that  $r \text{diag}(e, e, \dots, e)_{n \times n} + H \in M_n(\mathcal{V})^+$ . (Such an  $e$  is called a matrix order unit.)*
2. *If  $H \in M_n(\mathcal{V})_h$  satisfies  $r \text{diag}(e, e, \dots, e)_{n \times n} + H \in M_n(\mathcal{V})^+$  for all  $r > 0$ , then  $H \in M_n(\mathcal{V})^+$ . (Such a matrix order unit is called an Archimedean matrix order unit.)*

*Then there exists a Hilbert space  $H$  and a CP map  $\varphi : \mathcal{V} \rightarrow B(H)$  such that  $\varphi(e) = i d_H$  and  $\varphi$  is a complete order isomorphism onto its range.*

- The converse of this theorem is clear: Every operator system clearly is matrix ordered with  $e_{\mathcal{S}}$  as an Archimedean matrix order unit. And, the above theorem of Choi and Effros allows us to realize every Archimedean matrix ordered space with an operator system. We will thus use the terminology operator system for an Archimedean matrix ordered space and vice versa.

**Theorem 3.3.14** ([1, §4]) *Let  $S \subset B(H)$  be a finite dimensional operator system. Then,*

1. *there exists an  $f \in (S^d)^+$  such that  $f(p) > 0$  for all  $p \in S^+ \setminus \{0\}$ ; and*
2. *any such  $f$  is an Archimedean matrix order unit for the matrix ordered space  $S^d$ .*

- Remark 3.3.15* 1. When  $\mathcal{S} \subset M_p$  is an operator system for some  $p \geq 1$ , then  $tr_p$  also works as an Archimedean matrix order unit for the matrix ordering on  $\mathcal{S}^d$ ; thus,  $(\mathcal{S}^d, tr_p)$  is an operator system. However, the above theorem becomes important as, quite surprisingly, there exist finite dimensional operator systems ([1, §7]) that cannot be embedded in matrix algebras as operator sub-systems.
2. In general, for an infinite dimensional operator system  $\mathcal{S}$ , it is not clear whether  $\mathcal{S}^d$  admits an Archimedean matrix order unit or not.

For a finite dimensional operator system  $\mathcal{S}$ , it is easily checked that  $(\mathcal{S}^d)^d = \mathcal{S}$ . Thus, returning to our motivating question: if we start with an operator system, then the object that it is naturally the set of states on is  $\mathcal{S}^d$ .

### 3.4 Tensor Products of Operator Systems

In the usual axioms for quantum mechanics, if Alice has a quantum system represented as the states on a (finite dimensional) Hilbert space  $H_A$  and Bob has a quantum system represented as the states on a Hilbert space  $H_B$  then when we wish to consider the combined system it has states represented by the Hilbert space  $H_A \otimes_2 H_B$  where we've introduced the subscript 2 to indicate that this is the unique Hilbert space with inner product satisfying,

$$\langle h_A \otimes h_B | k_A \otimes k_B \rangle = \langle h_A | k_A \rangle \cdot \langle h_B | k_B \rangle.$$

As vector spaces we have that

$$\mathcal{B}(H_A \otimes_2 H_B) = \mathcal{B}(H_A) \otimes \mathcal{B}(H_B),$$

and since the left hand side is an operator system, this tells us exactly how to make an operator system out of the two operator systems appearing on the right hand side.

If  $P \in \mathcal{B}(H_A)^+$  and  $Q \in \mathcal{B}(H_B)^+$  then  $P \otimes Q \in \mathcal{B}(H_A \otimes_2 H_B)^+$ . But there are many positive operators in  $\mathcal{B}(H_A \otimes_2 H_B)^+$ , even of rank one, i.e., vector states, that can not be expressed in such a simple fashion and this is what leads to the important phenomenon known as *entanglement* which you've undoubtedly heard about in other lectures.

Now suppose that we are in one of the scenarios, such as in coding theory or capacity theory, where Alice and Bob do not both have all the operators on their respective Hilbert spaces but instead are constrained to certain operator subsystems,  $\mathcal{S}_A \subseteq \mathcal{B}(H_A)$  and  $\mathcal{S}_B \subseteq \mathcal{B}(H_B)$ . When we wish to consider the bivariate system that includes them both then as a vector space it should be  $\mathcal{S}_A \otimes \mathcal{S}_B$ , but which elements should be the states? More importantly, since we want to study quantum channels on this bivariate system, we need to ask: What should be the operator system structure on this bivariate system?

There is an easy answer, one could identify

$$\mathcal{S}_A \otimes \mathcal{S}_B \subseteq \mathcal{B}(H_A \otimes_2 H_B),$$

and when one does this we see that it is an operator subsystem, i.e., it contains the identity operator and is closed under the taking of adjoint. This operator system is denoted by  $\mathcal{S}_A \otimes_{sp} \mathcal{S}_B$  and is called their *spatial tensor product*.

Unfortunately, in general,

$$(\mathcal{S}_A \otimes_{sp} \mathcal{S}_B)^d \neq \mathcal{S}_A^d \otimes_{sp} \mathcal{S}_B^d,$$

so we need at least one other tensor product to explain what are the states on a tensor product.

Attempts by researchers such as Tsirelson [15, 16] to determine the sets of density matrices that are the outcomes of various multipartite quantum settings, and various works in operator algebras, argue for several other ways to form the tensor product of operator systems.

Thus, we are lead to consider more general ways that we can form an operator system out of a bivariate system.

Given operator systems  $(\mathcal{S}, e_{\mathcal{S}})$  and  $(\mathcal{T}, e_{\mathcal{T}})$ , we wish to take the vector space tensor product  $\mathcal{S} \otimes \mathcal{T}$ , endow it with a matrix ordering  $\{\mathcal{C}_n \subset M_n(\mathcal{S} \otimes \mathcal{T}) : n \geq 1\}$  such that  $\mathcal{S} \otimes \mathcal{T}$  together with these cones and  $e_{\mathcal{S}} \otimes e_{\mathcal{T}}$  forms an operator system.

**Definition 3.4.1** Given operator systems  $(\mathcal{S}, e_{\mathcal{S}})$  and  $(\mathcal{T}, e_{\mathcal{T}})$ , by an operator system tensor product  $\tau$  we mean a family of cones  $\mathcal{C}_n^{\tau} \subset M_n(\mathcal{S} \otimes \mathcal{T})$ ,  $n \geq 1$  such that  $(\mathcal{S} \otimes \mathcal{T}, \{\mathcal{C}_n^{\tau}\}, e_{\mathcal{S}} \otimes e_{\mathcal{T}})$  is an operator system satisfying:

1.  $P \otimes Q = [p_{ij} \otimes q_{kl}] \in \mathcal{C}_{nm}^{\tau}$  for all  $P = [p_{ij}] \in M_n(\mathcal{T})^+$ ,  $Q = [q_{kl}] \in M_m(\mathcal{T})^+$ ,  $n, m \geq 1$ .
2.  $\varphi \otimes \psi \in CP(\mathcal{S} \otimes_{\tau} \mathcal{T}, M_n \otimes M_m = M_{nm})$  for all  $\varphi \in CP(\mathcal{S}, M_n)$ ,  $\psi \in CP(\mathcal{T}, M_m)$ ,  $n, m \geq 1$ .

*Remark 3.4.2* Condition (2) in the above definition is analogous to the reasonableness axiom of Grothendieck for Banach space tensor products.

**Definition 3.4.3** A tensor product  $\tau$  of operator systems is said to be

1. *functorial* if

- (a) it is defined for any two operator systems; and
- (b)  $\varphi \otimes \psi \in CP(\mathcal{S}_1 \otimes_{\tau} \mathcal{S}_2, \mathcal{T}_1 \otimes_{\tau} \mathcal{T}_2)$  for all  $\varphi \in CP(\mathcal{S}_1, \mathcal{T}_1)$ ,  $\psi \in CP(\mathcal{S}_2, \mathcal{T}_2)$ .

2. *associative* if  $(\mathcal{S}_1 \otimes_{\tau} \mathcal{S}_2) \otimes_{\tau} \mathcal{S}_3$  is canonically completely order isomorphic to  $\mathcal{S}_1 \otimes_{\tau} (\mathcal{S}_2 \otimes_{\tau} \mathcal{S}_3)$  for any three operator systems  $\mathcal{S}_i$ ,  $i = 1, 2, 3$ .

3. *symmetric* if the flip map gives a complete order isomorphism  $\mathcal{S} \otimes_{\tau} \mathcal{T} \simeq \mathcal{T} \otimes_{\tau} \mathcal{S}$  for any two operator systems  $\mathcal{S}$  and  $\mathcal{T}$ .



- Suppose a  $*$ -vector space  $\mathcal{W}$  has two matrix orderings  $\{\mathcal{C}_n\}$  and  $\{\mathcal{C}'_n\}$ ; then  $(\mathcal{W}, \{\mathcal{C}_n\})$  is thought of as to be “bigger” than  $(\mathcal{W}, \{\mathcal{C}'_n\})$  if the identity map  $id_{\mathcal{W}} : (\mathcal{W}, \{\mathcal{C}_n\}) \rightarrow (\mathcal{W}, \{\mathcal{C}'_n\})$  is CP; or, equivalently, if  $\mathcal{C}_n \subset \mathcal{C}'_n$  for all  $n \geq 1$ .

Note that this notion is parallel to the fact that if  $\|\cdot\|_1$  and  $\|\cdot\|_2$  are two norms on a complex vector space  $X$ , then  $\|\cdot\|_1 \leq \|\cdot\|_2$  if and only if the (closed) unit balls with respect to these norms satisfy  $B_1(X, \|\cdot\|_2) \subset B_1(X, \|\cdot\|_1)$ .

### 3.4.1 Minimal Tensor Product of Operator Systems

Let  $(\mathcal{S}, e_{\mathcal{S}})$  and  $(\mathcal{T}, e_{\mathcal{T}})$  be two operator systems. For each  $p \geq 1$ , set

$$\mathcal{C}_p^{\min} = \left\{ [u_{ij}] \in M_p(\mathcal{S} \otimes \mathcal{T}) : [(\varphi \otimes \psi)(u_{ij})] \in M_{nmp}^+, \forall \varphi \in CP(\mathcal{S}, M_n), \right. \\ \left. \psi \in CP(\mathcal{T}, M_m), n, m \geq 1 \right\}.$$

**Theorem 3.4.4 ([8])** *With above setup,*

1.  $\{\mathcal{C}_p^{\min}\}$  is an operator system tensor product on  $\mathcal{S} \otimes \mathcal{T}$  and we denote the consequent operator system by  $\mathcal{S} \otimes_{\min} \mathcal{T}$ .
2.  $\otimes_{\min}$  is the smallest operator system tensor product in the sense that, if  $\{\mathcal{C}_p^{\tau}\}$  is any other operator system tensor product on  $\mathcal{S} \otimes \mathcal{T}$ , then  $\mathcal{C}_p^{\tau} \subset \mathcal{C}_p^{\min}$  for all  $p \geq 1$ .
3. if  $\mathcal{S} \subset B(H)$  and  $\mathcal{T} \subset B(K)$  for some Hilbert spaces  $H$  and  $K$ , then the spatial tensor product  $\mathcal{S} \otimes_{sp} \mathcal{T} \subset B(H \otimes K)$  is completely order isomorphic to the minimal tensor product  $\mathcal{S} \otimes_{\min} \mathcal{T}$ .
4.  $\otimes_{\min}$  is functorial, associative and symmetric.
5. if  $A$  and  $B$  are unital  $C^*$ -algebras, then their minimal tensor product as operator systems is completely order isomorphic to the image of  $A \otimes B$  in  $A \otimes_{C^*-\min} B$ .

### 3.4.2 Maximal Tensor Product of Operator Systems

Let  $(\mathcal{S}, e_{\mathcal{S}})$  and  $(\mathcal{T}, e_{\mathcal{T}})$  be two operator systems. For each  $n \geq 1$ , consider

$$\mathcal{D}_n^{\max} = \left\{ X^*([p_{ij}] \otimes [q_{kl}])X : [p_{ij}] \in M_r(\mathcal{S})^+, [q_{kl}] \in M_s(\mathcal{T})^+, \right. \\ \left. X \in M_{rs,n}(\mathbb{C}), r, s \geq 1 \right\}.$$

It can be seen that  $\{\mathcal{D}_n^{\max}\}$  gives a matrix ordering on  $\mathcal{S} \otimes \mathcal{T}$  and that  $e_{\mathcal{S}} \otimes e_{\mathcal{T}}$  is a matrix order unit for this ordering. However, there exist examples where  $e_{\mathcal{S}} \otimes e_{\mathcal{T}}$  fails to be an Archimedean matrix order unit. We, therefore, Archimedeanize the

above ordering, by considering:

$$\mathcal{C}_n^{\max} = \left\{ [u_{ij}] \in M_n(\mathcal{S} \otimes \mathcal{T}) : \delta \operatorname{diag}(e_{\mathcal{S}} \otimes e_{\mathcal{T}}, e_{\mathcal{S}} \otimes e_{\mathcal{T}}, \dots, e_{\mathcal{S}} \otimes e_{\mathcal{T}}) \right. \\ \left. + [u_{ij}] \in \mathcal{D}_n^{\max}, \forall \delta > 0 \right\}.$$

**Theorem 3.4.5 ([8])** *With above set up,*

1.  $\{\mathcal{C}_n^{\max}\}$  is an operator system tensor product on  $\mathcal{S} \otimes \mathcal{T}$  and we denote the consequent operator system by  $\mathcal{S} \otimes_{\max} \mathcal{T}$ .
2.  $\otimes_{\max}$  is the largest operator system tensor product in the sense that if  $\{\mathcal{C}_n^{\tau}\}$  is any other operator system tensor product on  $\mathcal{S} \otimes \mathcal{T}$ , then  $\mathcal{C}_n^{\max} \subset \mathcal{C}_n^{\tau}$  for all  $n \geq 1$ , i.e. max tensor product is the largest operator system tensor product.
3.  $\otimes_{\max}$  is functorial, associative and symmetric.
4. if  $A$  and  $B$  are unital  $C^*$ -algebras, then their maximal tensor product as operator systems is completely order isomorphic to the image of  $A \otimes B$  in  $A \otimes_{C^*-\max} B$ .

*Remark 3.4.6* For any  $C^*$ -algebra  $A \subset B(H)$ , the matrix ordering that it inherits does not depend (upto complete order isomorphism) upon the embedding or the Hilbert space  $H$ .

- At this point, we must mention that there is a big difference between the operator space maximal tensor product and the operator system maximal tensor product. This can be illustrated by an example:

For  $n, m \geq 1$ , the operator system maximal tensor product of  $M_n$  and  $M_m$  equals their  $C^*$ -maximal tensor product, whereas their operator space maximal tensor product does not.

**Theorem 3.4.7 (CP Factorisation Property [6])** *Let  $\mathcal{S} \subset B(H)$  be an operator system. Then  $\mathcal{S} \otimes_{\min} \mathcal{T} = \mathcal{S} \otimes_{\max} \mathcal{T}$  for all operator systems  $\mathcal{T}$  if and only if there exist nets of UCP maps  $\varphi_{\lambda} : \mathcal{S} \rightarrow M_{n_{\lambda}}$  and  $\psi_{\lambda} : M_{n_{\lambda}} \rightarrow \mathcal{S}$ ,  $\lambda \in \Lambda$  such that*

$$\|\psi_{\lambda} \circ \varphi_{\lambda}(s) - s\| \rightarrow 0, \forall s \in \mathcal{S}.$$

*Remark 3.4.8* One can even avoid the above embedding  $\mathcal{S} \subset B(H)$ , and give a characterization for CFPF, alternately, by considering the norm

$$\|s\| := \inf \left\{ r > 0 : \begin{pmatrix} re & s \\ s^* & re \end{pmatrix} \in M_2(\mathcal{S})^+ \right\}, s \in \mathcal{S}.$$

- We had remarked earlier that there exist finite dimensional operator systems which can not be embedded in matrix algebras. The surprise continues as, unlike  $C^*$ -algebras, not all finite dimensional operator systems are (min, max)-nuclear in the sense of Sect. 3.6.3—[8, Theorem 5.18]. However, we have the following useful fact.

**Lemma 3.4.9** ([6, 8]) *Matrix algebras are (min, max)-nuclear as operator systems, i.e.,  $M_n \otimes_{\min} \mathcal{S} = M_n \otimes_{\max} \mathcal{T}$  for any operator system  $\mathcal{T}$ .*

*Proof (Sketch!)* One basically identifies  $M_k(M_n \otimes \mathcal{S})$  naturally with  $M_{kn} \otimes \mathcal{S}$  and then with some serious calculation shows that  $\mathcal{D}_k^{\max}(M_n \otimes \mathcal{S}) = M_{kn}(\mathcal{S})^+ = \mathcal{C}_k^{\min}(M_n \otimes \mathcal{S})$ .  $\square$

*Remark 3.4.10* ([6]) In fact, the only finite dimensional (min, max)-nuclear operator systems are the matrix algebras and their direct sums.

*Remark 3.4.11* For a finite dimensional operator system  $\mathcal{S}$ , the canonical isomorphism  $\mathcal{S} \ni x \mapsto \hat{x} \in (\mathcal{S}^d)^d$ , where  $\hat{x}(f) := f(x)$  for all  $f \in \mathcal{S}^d$ , is a complete order isomorphism and  $\hat{e}_{\mathcal{S}}$  is an Archimedean matrix order unit for  $(\mathcal{S}^d)^d$ .

The requirement for the complete order isomorphism of the above map is that  $[x_{ij}] \in M_n(\mathcal{S})^+$  if and only if  $[\hat{x}_{ij}] \in M_n((\mathcal{S}^d)^d)^+$ ; and this can be deduced readily from the following fact:

**Lemma 3.4.12** ([8, Lemma 4.1]) *For any operator system  $\mathcal{S}$  and  $P \in M_n(\mathcal{S})$ ,  $P \in M_n(\mathcal{S})^+$  if and only if  $\varphi^{(n)}(P) \in M_{nm}^+$  for all  $\varphi \in \text{UCP}(\mathcal{S}, M_m)$  and  $m \geq 1$ .*

For vector spaces  $\mathcal{S}$  and  $\mathcal{T}$  with  $\mathcal{S}$  finite dimensional, we have an identification between  $\mathcal{S} \otimes \mathcal{T}$  and the space of all linear maps from  $\mathcal{S}^d$  into  $\mathcal{T}$  by identifying the element  $u = \sum_i s_i \otimes t_i \in \mathcal{S} \otimes \mathcal{T}$  with the map  $\mathcal{S}^d \ni f \mapsto \sum_i f(s_i)t_i \in \mathcal{T}$ .

**Lemma 3.4.13** ([8, Lemma 8.4]) *Let  $\mathcal{S}$  and  $\mathcal{T}$  be operator systems with  $\mathcal{S}$  finite dimensional and let  $[u_{ij}] \in M_n(\mathcal{S} \otimes \mathcal{T})$ . Then  $[u_{ij}] \in M_n(\mathcal{S} \otimes_{\min} \mathcal{T})^+$  if and only if the map  $\mathcal{S}^d \ni f \mapsto [L_{u_{ij}}(f)] \in M_n(\mathcal{T})$  is CP.*

*Proof* ( $\Rightarrow$ ) Let  $[u_{ij}] \in M_n(\mathcal{S} \otimes_{\min} \mathcal{T})^+$ ,  $k \geq 1$  and  $[f_{rs}] \in M_k(\mathcal{S}^d)^+$ . Suppose  $u_{ij} = \sum_p s_p^{ij} \otimes t_p^{ij}$ . We need to show that  $X := [L_{u_{ij}}]^{(k)}([f_{rs}]) \in M_k(M_n(\mathcal{T}))^+$ . We will again appeal to Lemma 3.4.12. Let  $m \geq 1$  and  $\varphi \in \text{UCP}(\mathcal{T}, M_m)$ . Then, for each  $1 \leq k, l \leq m$ , there exists a unique  $\varphi_{kl} \in \mathcal{T}^d$  such that  $\varphi(t) = [\varphi_{kl}(t)]$  for all  $t \in \mathcal{T}$ ; and, thus

$$\begin{aligned} \varphi^{(kn)}(X) &= [[\varphi \circ L_{u_{ij}}(f_{rs})]_{ij}]_{rs} \\ &= [[[\varphi_{kl} \circ L_{u_{ij}}(f_{rs})]_{kl}]_{ij}]_{rs} \\ &= \left[ \left[ \left[ \varphi_{kl} \left( \sum_p f_{rs}(s_p^{ij}) t_p^{ij} \right) \right]_{kl} \right]_{ij} \right]_{rs} \\ &= \left[ \left[ [f_{rs} \left( \sum_p s_p^{ij} \varphi_{kl}(t_p^{ij}) \right)]_{kl} \right]_{ij} \right]_{rs} \\ &= \tau \left( [f_{rs}]^{(nm)} \left( \left[ \left[ \sum_p s_p^{ij} \varphi_{kl}(t_p^{ij}) \right]_{kl} \right]_{ij} \right) \right), \end{aligned}$$

where  $[f_{rs}]^{(nm)}$  denotes the  $nm$ -amplification of the CP map  $\mathcal{S} \ni s \mapsto [f_{rs}(s)] \in M_k$  and  $\tau$  is the canonical flip  $*$ -isomorphism  $M_{nmk} \simeq M_n \otimes M_m \otimes M_k \simeq M_k \otimes M_n \otimes M_m \simeq M_{knm}$ . Next, since  $\otimes_{\min}$  is functorial,  $id \otimes \varphi : \mathcal{S} \otimes_{\min} \mathcal{T} \rightarrow M_m(\mathcal{S})$  is UCP, and, note that under the complete order isomorphism  $\theta : M_n \otimes M_m \otimes_{\min} \mathcal{S} \simeq M_n \otimes_{\min} \mathcal{S} \otimes M_m$ ,  $\theta\left(\left[\left[\sum_p s_p^{ij} \varphi_{kl}(t_p^{ij})\right]_{kl}\right]_{ij}\right) = (id \otimes \varphi)^{(n)}([u_{ij}])$ . Thus,  $\left[\left[\sum_p s_p^{ij} \varphi_{kl}(t_p^{ij})\right]_{kl}\right]_{ij} \in M_{nm}(\mathcal{S})^+$  and we conclude that  $\varphi^{(kn)}(X) \geq 0$ . In particular,  $\mathcal{S}^d \ni f \mapsto [L_{u_{ij}}(f)] \in M_n(\mathcal{T})$  is CP.

( $\Leftarrow$ ) Conversely, suppose the map  $\mathcal{S}^d \ni f \mapsto [L_{u_{ij}}(f)] \in M_n(\mathcal{T})$  is CP. Let  $k, m \geq 1$ ,  $\varphi \in \text{UCP}(\mathcal{S}, M_k)$  and  $\psi \in \text{UCP}(\mathcal{T}, M_m)$ . We need to show that  $(\varphi \otimes \psi)^{(n)}([u_{ij}]) \in M_{nm}^+$ . As above, there exist  $\varphi_{rs} \in \mathcal{S}^d$ ,  $\psi_{uv} \in \mathcal{T}^d$  for  $1 \leq r, s \leq k$  and  $1 \leq u, v \leq m$  such that  $\varphi(s) = [\varphi_{rs}(s)]$  and  $\psi(t) = [\psi_{uv}(t)]$  for all  $s \in \mathcal{S}, t \in \mathcal{T}$ . Also, since  $\varphi$  and  $\psi$  are UCP, we have  $[\varphi_{rs}] \in M_k(\mathcal{S}^d)^+$  and  $[\psi_{uv}] \in M_m(\mathcal{T}^d)^+$ . Suppose  $u_{ij} = \sum_p s_p^{ij} \otimes t_p^{ij}$ . Then

$$\begin{aligned} (\varphi \otimes \psi)^{(n)}([u_{ij}]) &= \left[ \sum_p \varphi(s_p^{ij}) \otimes \psi(t_p^{ij}) \right]_{ij} \\ &= \left[ \left[ \left[ \sum_p \varphi_{rs}(s_p^{ij}) \psi_{uv}(t_p^{ij}) \right]_{uv} \right]_{rs} \right]_{ij} \\ &= \left[ \left[ \left[ \psi_{uv} \left( \sum_p \varphi_{rs}(s_p^{ij}) t_p^{ij} \right) \right]_{uv} \right]_{rs} \right]_{ij} \\ &= \left[ \left[ \psi_{uv}(L_{u_{ij}}(\varphi_{rs})) \right]_{uv} \right]_{rs} \Big]_{ij} \\ &= \left[ \left[ \psi_{uv}(L_{u_{ij}}(\varphi_{rs})) \right]_{rs} \right]_{ij} \\ &= [\psi_{uv}]^{(kn)} \left[ \left[ (L_{u_{ij}}(\varphi_{rs})) \right]_{rs} \right]_{ij} \\ &= [\psi_{uv}]^{(kn)} \circ \tau \left( [L_{u_{ij}}]^{(k)}([\varphi_{rs}] \right), \end{aligned}$$

where  $\tau$  is the canonical complete order isomorphism  $\tau : M_k \otimes M_n \otimes_{\min} \mathcal{T} \simeq M_n \otimes M_k \otimes_{\min} \mathcal{T}$ .

Thus, with all the data that we have at our disposal, we immediately conclude that  $(\varphi \otimes \psi)^{(n)}([u_{ij}]) \geq 0$  and hence  $[u_{ij}] \in M_n(\mathcal{S} \otimes_{\min} \mathcal{T})^+$ .  $\square$

**Lemma 3.4.14** 1. Let  $\mathcal{S}$  and  $\mathcal{T}$  be operator systems and  $u \in \mathcal{D}_1^{\max}$ . Then, there exist  $n \geq 1$ ,  $[p_{ij}] \in M_n(\mathcal{S})^+$  and  $[q_{ij}] \in M_n(\mathcal{T})^+$  such that  $u = \sum_{i,j=1}^n p_{ij} \otimes q_{ij}$ .  
2. Let  $\mathcal{F}$  be a finite dimensional operator system and  $\{v_1, \dots, v_n\}$  be a basis of  $\mathcal{F}$  with a dual basis  $\{\delta_1, \dots, \delta_n\}$ . Then  $\sum_i \delta_i \otimes v_i \in (\mathcal{F}^d \otimes_{\min} \mathcal{F})^+$ .

*Proof* (1) By definition, there exist  $n, m \geq 1$ ,  $X \in M_{1, nm}$ ,  $P = [p_{ij}] \in M_n(\mathcal{S})^+$  and  $Q = [q_{rs}] \in M_m(\mathcal{T})^+$  such that  $u = X(P \otimes Q)X^*$ . We first note that, adding suitable zeros to  $X$  and  $P$  or  $Q$  according as  $n$  or  $m$  is smaller among them, we can assume that  $m = n$ . Thus, we have  $X = [x_{11}, x_{12}, \dots, x_{1n}, x_{21}, \dots, x_{nn}] \in M_{1, n^2}$ ,  $P = [p_{ij}] \in M_n(\mathcal{S})^+$  and  $Q = [q_{rs}] \in M_n(\mathcal{T})^+$  such that  $u = X(P \otimes Q)X^*$ . Consider  $\tilde{P} \in M_{n^2}(\mathcal{S})$  and  $\tilde{Q} \in M_{n^2}(\mathcal{T})$  given by  $\tilde{P}_{(i,r),(j,s)} = x_{ir}p_{ij}\bar{x}_{js}$  and  $\tilde{Q}_{(i,r),(j,s)} = q_{rs}$ ,  $1 \leq i, j, r, s \leq n$ . Clearly  $\tilde{P} \in M_{n^2}(\mathcal{S})^+$  and  $\tilde{Q} \in M_{n^2}(\mathcal{T})^+$  as under the identification  $M_{n^2}(\mathcal{S}) \simeq M_n(\mathcal{S}) \otimes M_n$ , we have  $\tilde{P} = \tilde{X}(P \otimes J_n)\tilde{X}^*$ , where  $\tilde{X} = \text{diag}(x_{11}, \dots, x_{1n}, x_{21}, \dots, x_{nn}) \in M_{n^2}$ ,  $J_n \in M_n$  is the positive semi-definite matrix with entries  $(J_n)_{ij} = 1$  for all  $1 \leq i, j \leq n$ ; and, under the identification  $M_{n^2}(\mathcal{T}) \simeq M_n \otimes M_n(\mathcal{T})$ , we have  $\tilde{Q} = J_n \otimes Q$ . Finally

$$\begin{aligned} \sum_{(i,r),(j,s)} \tilde{P}_{(i,r),(j,s)} \otimes \tilde{Q}_{(i,r),(j,s)} &= \sum_{i,j,r,s} x_{ir}p_{ij}\bar{x}_{js} \otimes q_{rs} = \sum_{i,j,r,s} x_{ir}(p_{ij} \otimes q_{rs})\bar{x}_{js} \\ &= X(P \otimes Q)X^* = u. \end{aligned}$$

(2) Let  $u = \sum_i \delta_i \otimes v_i$ . By Lemma 3.4.13, we just need to show that the map

$$(\mathcal{F}^d)^d \ni \hat{x} \mapsto L_u(\hat{x}) = \sum_i \hat{x}(\delta_i)v_i = \sum_i \delta_i(x)v_i = x \in \mathcal{F}$$

is CP, which is precisely the complete order isomorphism in Remark 3.4.11.  $\square$

*Proof of Theorem 3.4.7* ( $\Rightarrow$ ) By functoriality of the tensor products  $\otimes_{\min}$  and  $\otimes_{\max}$ , and by nuclearity of matrix algebras, for any operator system  $\mathcal{T} \subset B(K)$ , we have CP maps

$$\mathcal{S} \otimes_{\min} \mathcal{T} \xrightarrow{\varphi_\lambda \otimes id_{\mathcal{T}}} M_{n_\lambda} \otimes_{\min} \mathcal{T} = M_{n_\lambda} \otimes_{\max} \mathcal{T} \xrightarrow{\psi_\lambda \otimes id_{\mathcal{T}}} \mathcal{S} \otimes_{\max} \mathcal{T}.$$

In particular, their composition  $(\psi_\lambda \circ \varphi_\lambda) \otimes id_{\mathcal{T}} : \mathcal{S} \otimes_{\min} \mathcal{T} \rightarrow \mathcal{S} \otimes_{\max} \mathcal{T}$  is CP. Then, by the characterization of the norm on an operator system as given in Remark 3.4.8, one sees that the norm  $\|\cdot\|_{\mathcal{S} \otimes_{\max} \mathcal{T}}$  induced on  $\mathcal{S} \otimes \mathcal{T}$  by the operator system  $\mathcal{S} \otimes_{\max} \mathcal{T}$  is a sub-cross norm and thus  $(\psi_\lambda \circ \varphi_\lambda) \otimes id_{\mathcal{T}}(z)$  converges to  $z$  for all  $z \in \mathcal{S} \otimes_{\min} \mathcal{T}$ . In particular,  $id : \mathcal{S} \otimes_{\min} \mathcal{T} \rightarrow \mathcal{S} \otimes_{\max} \mathcal{T}$  is CP and we obtain  $\mathcal{S} \otimes_{\min} \mathcal{T} = \mathcal{S} \otimes_{\max} \mathcal{T}$ .

Conversely, suppose  $\mathcal{S} \otimes_{\min} \mathcal{T} = \mathcal{S} \otimes_{\max} \mathcal{T}$  for all operator systems  $\mathcal{T}$ . Let  $\mathcal{F} \subset \mathcal{S}$  be a finite dimensional operator sub-system ( $1_{\mathcal{F}} = 1_{\mathcal{S}}$ ). Using the fact that the tensor products  $\otimes_{\min}$  and  $\otimes_{\text{sp}}$  coincide, we have

$$\mathcal{F}^d \otimes_{\min} \mathcal{F} \subset \mathcal{F}^d \otimes_{\min} \mathcal{S} = \mathcal{F}^d \otimes_{\max} \mathcal{S}.$$

In particular, for a basis  $\{v_i\}$  of  $\mathcal{F}$  with dual basis  $\{\delta_i\}$ , we see, by Lemma 3.4.14(2), that  $\sum_i \delta_i \otimes v_i \in (\mathcal{F}^d \otimes_{\max} \mathcal{S})^+ = \mathcal{C}_1^{\max}(\mathcal{F}^d \otimes \mathcal{S})$ .

By Theorem 3.3.14, fix an  $f \in \mathcal{F}^d$  that plays the role of an Archimedean matrix order unit for  $\mathcal{F}^d$ . Since  $rf$  is also an Archimedean matrix order unit for any  $r > 0$ ,

we can assume that  $f(e_S) = 1$ . So, for all  $\delta > 0$ ,

$$\delta(f \otimes e_S) + \sum_i \delta_i \otimes v_i \in \mathcal{D}_1^{\max},$$

which implies, by Lemma 3.4.14(1), that for each  $\delta > 0$  there exist  $n \geq 1$ ,  $[f_{ij}] \in M_n(\mathcal{F}^d)^+$  and  $[p_{ij}] \in M_n(\mathcal{S})^+$  such that

$$\delta(f \otimes e_S) + \sum_i \delta_i \otimes v_i = \sum_{ij} f_{ij} \otimes p_{ij}.$$

Recall that  $[f_{ij}] \in M_n(\mathcal{F}^d)^+$  if and only if the map  $\mathcal{F} \ni v \mapsto [f_{ij}(v)] \in M_n$  is CP.

Claim We can choose  $[f_{ij}]$  in such way that the corresponding map  $\Phi$  is UCP.

*Proof of claim* Let  $Q = [q_{ij}] \in M_n$  be the support projection of the positive semi-definite matrix  $[f_{ij}(e_S)] \in M_n$ , i.e.,  $Q : \mathbb{C}^n \rightarrow \mathbb{C}^n$  is the orthogonal projection onto the subspace  $[f_{ij}(e_S)]\mathbb{C}^n \subset \mathbb{C}^n$ . Note that  $Y := [f_{ij}(e_S)]$  is invertible in  $QM_nQ$ . Let  $Y^{-1}$  denote the inverse of  $[f_{ij}(e_S)]$  in  $QM_nQ$ . Also, if  $p = \text{rank } Q$ , let  $U^*QU = \text{diag}(I_p, 0)$  be the diagonalization of  $Q$ , where  $U$  is a unitary matrix. Let  $X = [x_{11}, x_{12}, \dots, x_{1n}, x_{21}, \dots, x_{nm}] \in M_{1,n^2}$ , be given by  $X_{ij} = \delta_{i,j}$ . Then,

$$\begin{aligned} \sum_{ij} f_{ij} \otimes p_{ij} &= X([f_{ij}] \otimes [p_{ij}])X^* \\ &= X \left( Y^{1/2}U \begin{pmatrix} I_p \\ 0 \end{pmatrix} \otimes I_n \right) \cdot \left[ (I_p, 0)U^*Y^{-1/2}[f_{ij}]Y^{-1/2}U \right. \\ &\quad \left. \times \begin{pmatrix} I_p \\ 0 \end{pmatrix} \otimes [p_{ij}] \right] \cdot \left( Y^{1/2}U \begin{pmatrix} I_p \\ 0 \end{pmatrix} \otimes I_n \right)^* X^* \end{aligned}$$

and

$$(I_p, 0)U^*Y^{-1/2}[f_{ij}(e_S)]Y^{-1/2}U \begin{pmatrix} I_p \\ 0 \end{pmatrix} = I_p.$$

Hence the claim.

So, by Arveson's extension Theorem, there exists a UCP map  $\tilde{\Phi} : \mathcal{S} \rightarrow M_n$  such that  $\tilde{\Phi}|_{\mathcal{S}} = \Phi$ .

Now, consider the linear map  $\Psi : M_n \rightarrow \mathcal{S}$  sending  $E_{ij} \mapsto p_{ij}$  for all  $1 \leq i, j \leq n$ . By (Choi's) Theorem 1.2.4, this map is CP. Then, for  $v \in \mathcal{F}$ , we have

$$\Psi \circ \Phi(v) = \Psi([f_{ij}(v)]) = \sum_{ij} f_{ij}(v) p_{ij}.$$

On the other hand, via the canonical identification between  $\mathcal{F}^d \otimes \mathcal{S}$  and the space of linear transformations from  $\mathcal{F}^d$  into  $\mathcal{S}$ , we have  $\sum_{ij} f_{ij}(v) p_{ij} = (\sum_{ij} f_{ij} \otimes p_{ij})(v) = \delta(f \otimes e_{\mathcal{S}})(v) + (\sum_i \delta_i \otimes v_i)(v) = \delta f(v) e_{\mathcal{S}} + v$  for all  $v \in \mathcal{F}$ . Also, since  $f \in (\mathcal{F}^d)^+$ , we have  $f : \mathcal{F} \rightarrow \mathbb{C}$  is UCP (by our choice); so,  $\|f\|_{cb} = 1$  and  $|f(v)| \leq \|v\|$  for all  $v \in \mathcal{F}$ . Consider the directed set

$$\Lambda = \{(\mathcal{F}, \delta) : \mathcal{F} \subset \mathcal{S}, \text{ operator sub-system with } \dim \mathcal{F} < \infty, \delta > 0\}$$

with respect to the partial order  $\leq$  given by

$$(\mathcal{F}_1, \delta_1) \leq (\mathcal{F}_2, \delta_2) \Leftrightarrow \mathcal{F}_1 \subseteq \mathcal{F}_2 \text{ and } \delta_1 > \delta_2.$$

Thus, for each  $\lambda = (\mathcal{F}, \delta) \in \Lambda$ , there exist  $n_\lambda \geq 1$ ,  $\varphi_\lambda \in UCP(\mathcal{S}, M_{n_\lambda})$  and  $\psi'_\lambda \in CP(M_{n_\lambda}, \mathcal{S})$  satisfying  $\|\psi'_\lambda \circ \varphi_\lambda(v) - v\| \leq \delta \|v\|$  for all  $v \in \mathcal{F}$ . In particular,  $\|\psi'_\lambda \circ \varphi_\lambda(v) - v\| \rightarrow 0$  for every  $v \in \mathcal{S}$ . Since each  $\varphi_\lambda$  is unital,  $\psi'_\lambda(I_{n_\lambda})$  converges to  $e_{\mathcal{S}}$ . Fix a state  $\omega_\lambda$  on  $M_{n_\lambda}$  and set

$$\psi_\lambda(A) = \frac{1}{\|\psi'_\lambda\|} \psi'_\lambda(A) + \omega_\lambda(A)(e_{\mathcal{S}} - \frac{1}{\|\psi'_\lambda\|} \psi'_\lambda(I_{n_\lambda})).$$

Now,  $\psi_\lambda \in UCP(M_{n_\lambda}, \mathcal{S})$  for all  $\lambda \in \Lambda$  and we still have  $\|\psi_\lambda \circ \varphi_\lambda(v) - v\| \rightarrow 0$  for every  $v \in \mathcal{S}$ .  $\square$

### An Example of a Nuclear Operator System that is Not a $C^*$ -Algebra [6]

Let  $\mathcal{K}_0 = \overline{\text{span}}\{E_{i,j} : (i, j) \neq (1, 1)\} \subset B(\ell^2)$ , where  $E_{i,j}$  are the standard matrix units. Consider

$$\mathcal{S}_0 := \{\lambda I + T : \lambda \in \mathbb{C}, T \in \mathcal{K}_0\} \subset B(\ell^2),$$

the operator system spanned by the identity operator and  $\mathcal{K}_0$ . In [6], it has been proved that  $\mathcal{S}_0$  is a (min, max)-nuclear operator system and is not unittally completely order isomorphic to any unital  $C^*$ -algebra.

## 3.5 Graph Operator Systems

Graphs, especially the confusability graph, play an important role in Shannon's information theory. In the work of [2] on quantum capacity, they associate an operator system with a graph and show that many of Shannon's concepts have quantum interpretations in terms of these *graph operator systems*. Many concepts that we wish to deal with become much more transparent in this setting.

Given a (finite) graph  $G$  with vertices  $V = \{1, 2, \dots, n\}$  and edge set  $\mathcal{E} \subset V \times V$  (edges are not ordered; thus,  $(i, j) \in \mathcal{E} \Rightarrow (j, i) \in \mathcal{E}$ ), we have an operator system  $\mathcal{S}_G$  given by

$$\mathcal{S}_G := \text{span} \{ \{E_{ij} : (i, j) \in \mathcal{E}\} \cup \{E_{ii} : 1 \leq i \leq n\} \} \subset M_n.$$

Note that  $\mathcal{S}_G$  does not depend upon the number of bonds between two edges and the loops, if any. Thus, we assume that the graphs that we consider are simple, i.e., they have no loops or parallel edges.

### 3.5.1 Dual of a Graph Operator System

As a vector space, the operator system dual of a graph operator system  $\mathcal{S}_G$  can be identified with a vector subspace of  $(M_n)^d$  as

$$\mathcal{S}_G^d = \text{span} \{ \{ \delta_{ij} : (i, j) \in \mathcal{E} \} \cup \{ \delta_{ii} : 1 \leq i \leq n \} \} \subset (M_n)^d.$$

But it is not clear what the matrix order should be on this subspace. We will show that it is **not** the induced order. That is, while this is a natural vector space inclusion as operator systems,

$$\mathcal{S}_G^d \not\subset M_n^d !$$

*Example 3.5.1* If we consider the graph  $G$  with vertex set  $V = \{1, 2, \dots, n\}$  and edge set  $\mathcal{E} = \{(1, 2), (2, 3), (3, 4), \dots, (n-1, n), (n, n-1), \dots, (3, 2), (2, 1)\}$ , then

$$\mathcal{S}_G = \{\text{tridiagonal matrices}\} = \{[a_{ij}] \in M_n : a_{ij} = 0 \text{ for } |i - j| > 1\}.$$

**Observation** For the above graph  $G$ , given  $f = \sum_{1 \leq i, j \leq n, |i-j| \leq 1} b_{ij} \delta_{ij} \in \mathcal{S}_G^d$ , setting  $b_{ij} = 0$  for  $|i - j| > 1$ , one would like to know the conditions on the tridiagonal matrix  $B = [b_{ij}]$  such that  $f \in (\mathcal{S}_G^d)^+$ . It turns out that  $f \in (\mathcal{S}_G^d)^+$  if and only if the matrix  $B$  is *partially positive*, i.e., we can choose the off-tridiagonal entries of  $B$  to get a positive semi-definite matrix:

Indeed, if  $f \in (\mathcal{S}_G^d)^+$ , then  $f : \mathcal{S}_G \rightarrow \mathbb{C}$  is CP if and only if (by Arveson's extension theorem) it extends to a CP map  $\tilde{f} : M_n \rightarrow \mathbb{C}$ . Suppose  $\tilde{f} = \sum_{i,j} \tilde{b}_{ij} \delta_{ij}$  for some  $\tilde{b}_{ij} \in \mathbb{C}$ . Then, we have  $[\tilde{b}_{ij}] = \tilde{f}^{(n)}([E_{ij}]) \in (M_n)^+$  and, since  $\tilde{f}|_{\mathcal{S}_G} = f$ , we also see that  $\tilde{b}_{ij} = b_{ij}$  for all  $|i - j| \leq 1$ .

Conversely, if  $\tilde{B} = [\tilde{b}_{ij}]$  is a positive semi-definite matrix with  $\tilde{b}_{ij} = b_{ij}$  for all  $|i - j| \leq 1$ , then, by Theorem 1.2.4, the Schur multiplication map  $S_{\tilde{B}} : M_n \ni [x_{ij}] \mapsto [\tilde{b}_{ij} x_{ij}] \in M_n$  is CP. By Theorem 1.2.4 again, the map  $M_n \ni [x_{ij}] \mapsto$



$\sum_{i,j} x_{ij} \in \mathbb{C}$  is also CP; hence, the map  $M_n \ni [x_{ij}] \mapsto \sum_{i,j} \tilde{b}_{ij} x_{ij} \in \mathbb{C}$  is CP, i.e.,  $\tilde{f} := \sum_{i,j} \tilde{b}_{ij} \delta_{ij} \in (M_n^d)^+$ , and, since  $\tilde{f}|_{\mathcal{S}_G} = f$ , we have  $f \in (\mathcal{S}_G^d)^+$ .

In general, a tridiagonal matrix  $B$  can be extended to a positive semi-definite matrix if and only if each  $2 \times 2$  submatrix of  $B$  is positive semi-definite [12]. Thus, we have a very clear picture in this case of which linear functionals are positive.  $\square$

The situation for a general graph is as follows:

Given a graph  $G$  then a functional  $f : \mathcal{S}_G \rightarrow \mathbb{C}$  has the form  $f = \sum_{(i,j) \in \mathcal{E} \text{ or } i=j} b_{i,j} \delta_{i,j}$ . Set  $B = \sum_{(i,j) \in \mathcal{E} \text{ or } i=j} b_{i,j} E_{i,j}$ . Then  $f \in (\mathcal{S}_G^d)^+$  if and only if there exists an  $M \in \mathcal{S}_G^\perp$  such that  $B + M \in M_n^+$ .

This result can be put into the language of *partially defined matrices*. Notice that when we have  $f : \mathcal{S}_G \rightarrow \mathbb{C}$  and we try to form the density matrix of this functional,  $(f(E_{i,j}))$ , then because only some of the matrix units  $E_{i,j}$  belong to the space  $\mathcal{S}_G$  we only have some of the entries of our matrix specified. This is what is meant by a partially defined matrix, i.e., a matrix where only some entries are given and the rest are viewed as free variables. Choosing the matrix  $M$  above is tantamount to choosing values for the unspecified entries. In the language of partially defined matrices, this is called *completing the matrix*.

Thus, what we have shown is that each functional gives rise to a partially defined density matrix and that the positive functionals on  $\mathcal{S}_G$  are precisely those whose density matrices can be completed to positive semidefinite matrices.

## 3.6 Three More Operator System Tensor Products

There are at least three more operator system tensor products that are important in the operator algebras community and are likely to have some importance for quantum considerations as well. In particular, the one that we call the commuting tensor product is important for the study of the Tsirelson conjectures [7, 16].

Moreover, there are many important properties of operator systems that are equivalent to the behaviour of the operator system with respect to these tensor products. We give a very condensed summary of this theory below.

### 3.6.1 The Commuting Tensor Product $\otimes_c$ .

Let  $\mathcal{S}$  and  $\mathcal{T}$  be operator systems and  $\varphi : \mathcal{S} \rightarrow B(H)$ ,  $\psi : \mathcal{T} \rightarrow B(H)$  be UCP maps with commuting ranges, i.e.,  $\varphi(s)\psi(t) = \psi(t)\varphi(s)$  for all  $s \in \mathcal{S}$ ,  $t \in \mathcal{T}$ . Define  $\varphi \odot \psi : \mathcal{S} \otimes \mathcal{T} \rightarrow B(H)$  by  $(\varphi \odot \psi)(s \otimes t) = \varphi(s)\psi(t)$ ,  $s \in \mathcal{S}$ ,  $t \in \mathcal{T}$ .

Consider

$$C_n^c = \{[u_{ij}] \in M_n(\mathcal{S} \otimes \mathcal{T}) : (\varphi \odot \psi)^{(n)}([u_{ij}]) \in B(H^{(n)})^+, \text{ Hilbert spaces } H, \\ \varphi \in \text{UCP}(\mathcal{S}, B(H)), \psi \in \text{UCP}(\mathcal{T}, B(H)) \text{ with commuting ranges}\}.$$

**Theorem 3.6.1** *With above set up,*

1.  $\{C_n^c\}$  is an operator tensor on  $\mathcal{S} \otimes \mathcal{T}$  and the consequent operator system is denoted by  $\mathcal{S} \otimes_c \mathcal{T}$ .
2.  $\otimes_c$  is functorial and symmetric.

*Remark 3.6.2* Explicit examples showing non-associativity of the tensor product  $\otimes_c$  are not known yet.

### 3.6.2 The Tensor Products $\otimes_{el}$ and $\otimes_{er}$ .

Let  $\mathcal{S} \subset B(H)$  and  $\mathcal{T} \subset B(K)$  be operator systems. Then the inclusion  $\mathcal{S} \otimes \mathcal{T} \subset B(H) \otimes_{\max} \mathcal{T}$  induces an operator system on  $\mathcal{S} \otimes \mathcal{T}$ , which is referred as *enveloping on the left* and is denoted by  $\mathcal{S} \otimes_{el} \mathcal{T}$ . Likewise, the inclusion  $\mathcal{S} \otimes \mathcal{T} \subset \mathcal{S} \otimes_{\max} B(K)$  induces an operator system on  $\mathcal{S} \otimes \mathcal{T}$  called *enveloping on the right* and is denoted by  $\mathcal{S} \otimes_{er} \mathcal{T}$ .

It requires some work to establish that the operator system tensor products  $\otimes_{el}$  and  $\otimes_{er}$  of  $\mathcal{S}$  and  $\mathcal{T}$  do not depend (upto complete order isomorphisms) on the embeddings  $\mathcal{S} \subset B(H)$  and  $\mathcal{T} \subset B(K)$ .

- $\otimes_{el}$  and  $\otimes_{er}$  are both functorial. However, it is not clear whether they are associative or not?
- $\mathcal{S} \otimes_{el} \mathcal{T} \simeq \mathcal{T} \otimes_{er} \mathcal{S}$  as operator systems.

### 3.6.3 Lattice of Operator System Tensor Products

We have the following lattice structure among the above five operator system tensor products:

$$\min \leq el, er \leq c \leq \max.$$

Given functorial operator system tensor products  $\alpha$  and  $\beta$ , an operator system  $\mathcal{S}$  is said to be  $(\alpha, \beta)$ -nuclear provided  $\mathcal{S} \otimes_{\alpha} \mathcal{T} \simeq \mathcal{S} \otimes_{\beta} \mathcal{T}$  as operator spaces for all operator systems  $\mathcal{T}$ .

Recall that a  $C^*$ -algebra is nuclear if and only if it satisfies completely positive approximation property (CPAP). We saw in Theorem 3.4.7, that this generalizes to

the context of operator system as well, i.e., an operator system  $\mathcal{S}$  is (min, max)-nuclear if and only if it satisfies completely positive factorization property (CPFP). In general,  $(\alpha, \beta)$ -nuclearity of operator systems have some further analogous structural characterizations.

## 3.7 Some Characterizations of Operator System Tensor Products

### 3.7.1 Exact Operator Systems

Analogous to the notion of exactness for  $C^*$ -algebras and operator systems, there is a notion of exactness for operator systems as well—see [9].

**Theorem 3.7.1 ([9, Theorem 5.7])** *An operator system  $\mathcal{S}$  is 1-exact if and only if it is (min, el)-nuclear.*

For an operator system  $\mathcal{S}$ , we consider its Banach space dual  $\mathcal{S}^*$  and endow it with a matrix ordering as we did for  $\mathcal{S}^d$  above. We repeat the process to endow its double dual  $\mathcal{S}^{**}$  also with a matrix ordering. It is not very difficult to see that the canonical embedding  $\mathcal{S} \subset \mathcal{S}^{**}$  is a complete order isomorphism onto its image. Also, it is a fact—[9, Proposition 6.2]—that  $\hat{e}_{\mathcal{S}}$  is an Archimedean matrix order unit for  $\mathcal{S}^{**}$ .

### 3.7.2 Weak Expectation Property (WEP)

**Lemma 3.7.2 ([9, Lemma 6.3])** *Let  $\mathcal{S}$  be an operator system. Then the following are equivalent:*

1. *There exists an inclusion  $\mathcal{S} \subset B(H)$  such that the canonical embedding  $\iota : \mathcal{S} \rightarrow \mathcal{S}^{**}$  extends to a CP map  $\tilde{\iota} : B(H) \rightarrow \mathcal{S}^{**}$ .*
2. *For every operator system inclusion  $\mathcal{S} \subset \mathcal{T}$ , the map  $\iota : \mathcal{S} \rightarrow \mathcal{S}^{**}$  extends to a CP map  $\tilde{\iota} : \mathcal{T} \rightarrow \mathcal{S}^{**}$ .*
3. *The canonical embedding  $\iota : \mathcal{S} \subset \mathcal{S}^{**}$  factors through an injective operator system by UCP maps, i.e., there is an injective operator system  $\mathcal{T}$  and UCP maps  $\varphi_1 : \mathcal{S} \rightarrow \mathcal{T}$  and  $\varphi_2 : \mathcal{T} \rightarrow \mathcal{S}^{**}$  such that  $\iota = \varphi_2 \circ \varphi_1$ .*

**Definition 3.7.3** An operator system  $\mathcal{S}$  is said to have weak expectation property (WEP) if it satisfies any of the equivalent conditions above.

**Theorem 3.7.4 ([5, 9])** *Let  $\mathcal{S}$  be an operator system. Then  $\mathcal{S}$  possesses WEP if and only if it is (el, max)-nuclear.*

**Theorem 3.7.5 ([9, Theorem 6.11])** *Let  $\mathcal{S}$  be a finite dimensional operator system. Then the following are equivalent:*

1.  $\mathcal{S}$  possesses WEP;
2.  $\mathcal{S}$  is (el, max)-nuclear;
3.  $\mathcal{S}$  is (min, max)-nuclear;
4.  $\mathcal{S}$  is completely order isomorphic to a  $C^*$ -algebra; and
5.  $\mathcal{S} \otimes_{\text{el}} \mathcal{S}^* = \mathcal{S} \otimes_{\text{max}} \mathcal{S}^*$ .

### 3.7.3 Operator System Local Lifting Property (OSLLP)

Let  $H$  be an infinite dimensional Hilbert space and  $\mathcal{S}$  an operator system. Let  $u \in \text{UCP}(\mathcal{S}, Q(H))$ , where  $Q(H)$  is the Calkin algebra  $Q(H) = B(H)/K(H)$ . The operator system  $\mathcal{S}$  is said to have OSLLP if, for each such  $u$ , every finite dimensional operator sub-system  $\mathcal{F} \subset \mathcal{S}$  admits a lifting  $\tilde{u} \in \text{UCP}(\mathcal{F}, B(H))$  such that  $\pi \circ \tilde{u} = u|_{\mathcal{F}}$ , where  $\pi : B(H) \rightarrow Q(H)$  is the canonical quotient map.

**Theorem 3.7.6 ([9, Theorems 8.1, 8.5])** *Let  $\mathcal{S}$  be an operator system. Then the following are equivalent:*

1.  $\mathcal{S}$  possesses OSSLP;
2.  $\mathcal{S} \otimes_{\text{min}} B(H) = \mathcal{S} \otimes_{\text{max}} B(H)$  for every Hilbert space  $H$ ; and
3.  $\mathcal{S}$  is (min, er)-nuclear.

### 3.7.4 Double Commutant Expectation Property (DCEP)

An operator system  $\mathcal{S}$  is said to have DCEP if for every completely order embedding  $\varphi : \mathcal{S} \rightarrow B(H)$  there exists a completely positive mapping  $E : B(H) \rightarrow \varphi(\mathcal{S})''$  fixing  $\mathcal{S}$ , i.e., satisfying  $E \circ \varphi = \varphi$ .

**Theorem 3.7.7 ([9])** *An operator system  $\mathcal{S}$  possesses DCEP if and only if it is (el, c)-nuclear.*

*Remark 3.7.8* In particular, since  $\text{el} \leq \text{c}$ , an operator system  $\mathcal{S}$  is 1-exact and possesses DCEP if and only if it is (min, c)-nuclear. Of course, it will be desirable to have a better characterization for (min, c)-nuclearity.

**Definition 3.7.9** An unordered graph  $G = (V, \mathcal{E})$  is said to be a chordal graph if every cycle in  $G$  of length greater than 3 has a chord, or, equivalently, if  $G$  has no minimal cycle of length  $\geq 4$ .

**Theorem 3.7.10 ([8])** *If  $G$  is a chordal graph, then  $\mathcal{S}_G$  is (min, c)-nuclear.*

- Question 3.7.11** 1. For which graphs  $G$  are the operator systems  $\mathcal{S}_G$  (min, c)-nuclear?
2. Consider the graph  $G$  consisting of a quadrilateral. Clearly  $G$  is not chordal. Is the graph operator system  $\mathcal{S}_G$  (min, c)-nuclear?
3. Is every graph operator system (min, c)-nuclear?
4. For arbitrary operator system tensor products  $\alpha$  and  $\beta$ , which graphs give  $(\alpha, \beta)$ -nuclear graph operator systems?
5. For a graph  $G$ , study the above problems for the dual operator system  $\mathcal{S}_G^d$  as well.
6. Obtain characterizations for  $(\alpha, \beta)$ -nuclearity of operator systems for the remaining cases.
7. If a graph operator system  $\mathcal{S}_G$  is  $(\alpha, \beta)$ -nuclear, identify the tensor products  $\eta$  and  $\xi$  (if any) such that  $\mathcal{S}_G^d$  is  $(\eta, \xi)$ -nuclear.

### 3.8 Operator System Tensor Products and the Conjectures of Kirchberg and Tsirelson

#### 3.8.1 Special Operator Sub-systems of the Free Group $C^*$ -Algebras

Let  $\mathbb{F}_n$  be the free group on  $n$  generators, say,  $\{g_1, g_2, \dots, g_n\}$ . For any Hilbert space  $H$ , any choice of  $n$  unitaries  $\{U_1, U_2, \dots, U_n\}$  in  $B(H)$  gives a (unitary) representation  $\pi : \mathbb{F}_n \rightarrow B(H)$  of  $\mathbb{F}_n$  sending  $g_i$  to  $U_i$  for all  $1 \leq i \leq n$ . Recall, the full group  $C^*$ -algebra  $C^*(\mathbb{F}_n)$  is the closure of the group algebra  $\mathbb{C}[\mathbb{F}_n]$  in the norm obtained by taking supremum over all (unitary) representations of the group  $\mathbb{F}_n$ . Let

$$\mathcal{S}_n = \text{span}\{1, g_1, \dots, g_n, g_1^*, \dots, g_n^*\} \subset C^*(\mathbb{F}_n).$$

Clearly,  $\mathcal{S}_n$  is a  $(2n + 1)$ -dimensional operator system.

#### 3.8.2 Kirchberg's Conjecture

A famous conjecture of Kirchberg states that the full group  $C^*$ -algebra  $C^*(\mathbb{F}_\infty)$  has WEP—[11]. It attracts immense importance from its equivalence with some other important conjectures in the world of Operator Algebras and now thanks to the work of [7] we now know that Tsirelson's attempts at determining the possible sets of density matrices for quantum outcomes is also related.

In fact, it is now known that if Tsirelson's conjectures are true then necessarily Kirchberg's conjecture is true. For a physicists perspective on these issues see [4].

**Theorem 3.8.1 ([11])** *The following statements are equivalent:*

1. *Connes' Embedding Theorem: Every  $II_1$ -factor with separable predual can be embedded as a subfactor in to the free ultraproduct of the hyperfinite  $II_1$ -factor.*
2.  $C^*(\mathbb{F}_n) \otimes_{\text{sp}} C^*(\mathbb{F}_n) = C^*(\mathbb{F}_n) \otimes_{C^*-\max} C^*(\mathbb{F}_n)$  for all  $n \geq 1$ .
3.  $C^*(\mathbb{F}_n)$  has WEP for all  $n \geq 1$ .
4.  $C^*(\mathbb{F}_\infty)$  has WEP. (Kirchberg's conjecture)
5. Every  $C^*$ -algebra is a quotient of a  $C^*$ -algebra with WEP.

To the above list, the techniques of operator system tensor products, has contributed the following ("seemingly simpler") equivalent statements:

**Theorem 3.8.2 ([9])** *The following statements are equivalent:*

1.  $C^*(\mathbb{F}_\infty)$  has WEP.
2.  $\mathcal{S}_n$  is (el, c)-nuclear for all  $n \geq 1$ .
3.  $\mathcal{S}_n \otimes_{\min} \mathcal{S}_n = \mathcal{S}_n \otimes_c \mathcal{S}_n$  for all  $n \geq 1$ .
4. Every (min, er)-nuclear operator system is (el, c)-nuclear.
5. Every operator system possessing OSLLP possesses DCEP.

In the above list of equivalences, the equivalence (1)  $\leftrightarrow$  (2) in Theorem 3.8.1 is the deepest link and was proved first by Kirchberg in [10]. An essential part of the proof of this equivalence involved the following deep theorem due to Kirchberg:

**Theorem 3.8.3 ([10])**  $C^*(\mathbb{F}_n) \otimes_{\text{sp}} B(H) = C^*(\mathbb{F}_n) \otimes_{C^*-\max} B(H)$  for all  $n \geq 1$  and for all Hilbert spaces  $H$ .

- Quite surprisingly, making use of the notion of *quotient of an operator system*, a relatively much easier proof of Kirchberg's Theorem has been obtained in [3].

### 3.8.3 Quotient of an Operator System

The idea of quotient of an operator system comes from the requirement that given operator systems  $\mathcal{S}$  and  $\mathcal{T}$ , and a UCP  $\varphi : \mathcal{S} \rightarrow \mathcal{T}$ , we would like to have a quotient operator system  $\mathcal{S}/\ker\varphi$  such that the canonical quotient map  $q : \mathcal{S} \rightarrow \mathcal{S}/\ker\varphi$  is UCP and so is the factor map  $\tilde{\varphi} : \mathcal{S}/\ker\varphi \rightarrow \mathcal{T}$ .

It also gives a way to explain the duals of graph operator systems, they are actually quotients of the matrix algebra.

**Definition 3.8.4 (Quotient Map)** Let  $\mathcal{S}$  and  $\mathcal{T}$  be operator systems. Then a UCP map  $\varphi : \mathcal{S} \rightarrow \mathcal{T}$  is said to be a quotient map if  $\varphi$  is surjective and the canonical factor map  $\tilde{\varphi} : \mathcal{S}/\ker\varphi \rightarrow \mathcal{T}$  is a complete order isomorphism. In other words,  $\mathcal{T}$  is a quotient of  $\mathcal{S}$ .

*Example 3.8.5* Let  $\mathcal{T}_{n+1} = \{\text{tridiagonal } (n+1) \times (n+1) \text{ matrices}\} \subset M_{n+1}$  and

$$\mathcal{K}_{n+1} = \{\text{trace 0 diagonal } (n+1) \times (n+1) \text{ matrices}\} \subset M_{n+1}.$$

Consider  $\varphi : \mathcal{T}_{n+1} \rightarrow \mathcal{S}_n$  given by  $\varphi(E_{ii}) = \frac{1}{n}1$ ,  $\varphi(E_{i,i+1}) = \frac{1}{n}g_i$ ,  $\varphi(E_{i+1,i}) = \frac{1}{n}g_i^*$ . Clearly  $\varphi$  is onto and UCP. Also,  $\ker \varphi = \mathcal{K}_{n+1}$ . It is a fact that  $\varphi$  is a quotient map in above sense. In particular,  $\mathcal{S}_n$  is completely order isomorphic to  $\mathcal{T}_{n+1}/\mathcal{K}_{n+1}$ .

## References

1. M.-D. Choi, E. Effros, Injectivity and operator spaces. *J. Funct. Anal.* **24**, 156–209 (1977)
2. R. Duan, S. Severini, A. Winter, Zero-error communication via quantum channels, non-commutative graphs, and a quantum Lovász number. *IEEE Trans. Inf. Theory* **59**, 1164–1174 (2013)
3. D. Farenick, V.I. Paulsen, Operator system quotients of matrix algebras and their tensor products. *Math. Scand.* **111**, 210–243 (2012)
4. T. Fritz, Operator system structures on the unital direct sum of  $C^*$ -algebras. arXiv:1011.1247 (2010)
5. K.H. Han, On maximal tensor products and quotient maps of operator systems. *J. Math. Anal. Appl.*, **384**(2), 375–386 (2011)
6. K.H. Han, V. Paulsen, An approximation theorem for nuclear operator systems. *J. Funct. Anal.* **261**, 999–1009 (2011)
7. M. Junge, M. Navascues, C. Palazuelos, D. Perez-Garcia, V.B. Scholtz, R.F. Werner, Connes' embedding problem and Tsirelson's problem. *J. Math. Phys.* **52**, 012102 (2011)
8. A. Kavruk, V. Paulsen, I. Todorov, M. Tomforde, Tensor products of operator systems. *J. Funct. Anal.* **261**, 267–299 (2011)
9. A. Kavruk, V. Paulsen, I. Todorov, M. Tomforde, Quotients, exactness and nuclearity in the operator system category. *Adv. Math.* **235**, 321–360 (2013)
10. E. Kirchberg, On non-semisplit extensions, tensor products and exactness of group  $C^*$ -algebras. *Invent. Math.* **112**, 449–489 (1993)
11. E. Kirchberg, Commutants of unitaries in UHF algebras and functorial properties of exactness. *J. Reine Angew. Math.* **452**, 39–77 (1994)
12. V. Paulsen, S.C. Power, R. Smith, Schur products and matrix completions. *J. Funct. Anal.* **85**, 151–178 (1989)
13. V.I. Paulsen, F. Shultz, Complete positivity of the map from a basis to its dual basis. *J. Math. Phys.* **54**, 072201 (2013)
14. V.I. Paulsen, I.G. Todorov, M. Tomforde, Operator system structures on ordered spaces. *Proc. Lond. Math. Soc.*, **102**(1), 25–49 (2011)
15. B.S. Tsirelson, Quantum generalizations of Bell's inequality. *Lett. Math. Phys.* **4**(4), 93–100 (1980)
16. B.S. Tsirelson, Some results and problems on quantum Bell-type inequalities. *Hadronic J. Suppl.* **8**(4), 329–345 (1993)

# Chapter 4

## Quantum Information Theory

**Abstract** In this chapter, based on the lectures by Andreas Winter, we survey four different areas in quantum information theory in which ideas from the theory of operator systems and operator algebras play a natural role. The first is the problem of zero-error communication over quantum channels, which makes use of concepts from operator systems and quantum error correction. Second, we study the strong subadditivity property of quantum entropy and the central role played by the structure theorem of matrix algebras in understanding its equality case. Next, we describe different norms on quantum states and the corresponding induced norms on quantum channels. Finally, we look at matrix-valued random variables, prove Hoeffding-type tail bounds and describe the applications of such matrix tail bounds in the quantum information setting.

### 4.1 Zero-Error Communication Via Quantum Channels

A quantum channel  $\mathcal{T}$  is a completely positive, trace-preserving (CPTP) map from the states of one system ( $A$ ) to another ( $B$ ). Specifically,  $\mathcal{T} : \mathcal{B}(\mathcal{H}_A) \rightarrow \mathcal{B}(\mathcal{H}_B)$  is a CPTP map from the set of bounded linear operators on Hilbert space  $\mathcal{H}_A$  to operators in  $\mathcal{H}_B$ . In this section we focus on the problem of zero-error communication using quantum channels. We begin with a brief review of preliminaries including the idea of purification and the Choi-Jamiolkowski isomorphism.

**Definition 4.1.1 (Purification)** Given any positive semi-definite operator  $\rho \geq 0$  in  $\mathcal{B}(\mathcal{H}_A)$ , suppose there exists a vector  $|v\rangle \in \mathcal{H}_A \otimes \mathcal{H}_{A'}$ , where  $\mathcal{H}_{A'}$  is simply an auxiliary Hilbert space, such that  $\text{Tr}_{A'}[|v\rangle\langle v|] = \rho$ . The vector  $|v\rangle$  in the extended Hilbert space is said to be a **purification** of the operator  $\rho$ .

When  $\text{Tr}[\rho] = 1$ , that is, when  $\rho$  is a valid quantum state, then the corresponding vector  $|v\rangle$  is a pure state of the extended Hilbert space, satisfying  $\langle v|v\rangle = 1$ .

In order to obtain a purification of a state  $\rho$ , it suffices to have the dimensions of the auxiliary space  $\mathcal{H}'$  to be equal to the rank of  $\rho$ . To see this, suppose  $\rho$  has a



spectral decomposition  $\rho = \sum_i r_i |e_i\rangle\langle e_i|$ , then a purification of  $\rho$  is simply given by

$$|v\rangle = \sum_i \sqrt{r_i} |e_i\rangle \otimes |e_i\rangle.$$

The purification of a given state  $\rho$  is not unique. Suppose there exists another purification  $|w\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$ , via a different extension of the Hilbert space  $\mathcal{H}_A$ , such that  $\text{Tr}_B[|w\rangle\langle w|] = \rho$ . Then, there exists a unique isometry  $U : \mathcal{H}_{A'} \rightarrow \mathcal{H}_B$  such that<sup>1</sup>

$$|w\rangle = (\mathbb{I} \otimes U)|v\rangle.$$

This isometry can in fact be obtained from the Choi uniqueness theorem (Theorem 3.1.9).

We next define the Choi-Jamiolkowski matrix corresponding to a quantum channel  $\mathcal{T} : \mathcal{B}(\mathcal{H}_A) \rightarrow \mathcal{B}(\mathcal{H}_B)$ . This provides an alternate way to obtain the Stinespring dilation of the CPTP map  $\mathcal{T}$ , discussed earlier in Theorem 1.1.8.

**Definition 4.1.2 (Choi-Jamiolkowski Matrix)** Let  $\{|i\rangle\}$  denote an orthonormal basis for  $\mathcal{H}_A$ . Consider the (non-normalized) maximally entangled state:  $|\Phi_{AA'}\rangle = \sum_i |ii\rangle \in \mathcal{H}_A \otimes \mathcal{H}_{A'}$ , with  $\mathcal{H}_{A'}$  chosen to be isomorphic to  $\mathcal{H}$ . The Choi-Jamiolkowski matrix corresponding to a CPTP map  $\mathcal{T} : \mathcal{B}(\mathcal{H}_A) \rightarrow \mathcal{B}(\mathcal{H}_B)$  is then defined as

$$J_{AB} := (\mathbb{I} \otimes \mathcal{T})\Phi, \quad (4.1)$$

where, the operator  $\Phi = |\Phi\rangle\langle\Phi| \in \mathcal{B}(\mathcal{H}_A \otimes \mathcal{H}_{A'})$  is simply

$$\Phi := \sum_{i,j} |i\rangle\langle j| \otimes |i\rangle\langle j| = \sum_{i,j} |ii\rangle\langle jj|.$$

Complete positivity of  $\mathcal{T}$  implies that  $J \geq 0$ , and the trace-preserving condition on  $\mathcal{T}$  implies,

$$\text{Tr}_B[J] = I_A = \text{Tr}_{A'}(\Phi). \quad (4.2)$$

Pick a purification  $|G_{ABC}\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_C$  of the matrix  $J$ . Thus,  $|G_{ABC}\rangle\langle G_{ABC}|$  is a rank-one, positive operator satisfying  $\text{Tr}_C[|G_{ABC}\rangle\langle G_{ABC}|] = J_{AB}$ .  $\mathcal{H}_C$  is any auxiliary Hilbert space whose dimension is  $\dim(\mathcal{H}_C) \geq \text{rank}(J)$ . Then, Eq. (4.2) implies that

$$\text{Tr}_{BC}[|G_{ABC}\rangle\langle G_{ABC}|] = I_A = \text{Tr}[\Phi_{AA'}].$$

---

<sup>1</sup>A remark on notation: throughout this chapter we use  $I$  to denote the identity operator and  $\mathbb{I}$  to denote the identity map, for example,  $\mathbb{I}_A : \mathcal{B}(\mathcal{H}_A) \rightarrow \mathcal{B}(\mathcal{H}_A)$ .

Therefore, there exists an isometry  $U : \mathcal{H}_{A'} \rightarrow \mathcal{H}_B \otimes \mathcal{H}_C$  such that

$$|G_{ABC}\rangle = (I \otimes U)|\Phi_{AA'}\rangle.$$

Since the Choi matrix  $J_{AB}$  corresponding to the map  $\mathcal{T}$  is unique, we have for any  $X \in \mathcal{B}(\mathcal{H}_A)$ ,

$$\mathcal{T}(X) = \text{Tr}_C[UXU^\dagger].$$

The isometry thus corresponds to the Stinespring dilation of the map  $\mathcal{T}$ . Furthermore, we also obtain the Choi-Kraus decomposition (Theorem 1.2.1) of the map  $\mathcal{T}$  by noting that the isometry can be rewritten as  $U = \sum_i E_i \otimes |v_i\rangle$ , where  $\{|v_i\rangle\}$  is an orthonormal basis for  $\mathcal{H}_C$ . Thus,

$$\mathcal{T}(X) = \text{Tr}_C[UXU^\dagger] = \sum_i E_i X E_i^\dagger, \quad \forall X \in \mathcal{B}(\mathcal{H}_A).$$

The non-uniqueness of the Kraus representation is captured by the non-uniqueness of the choice of basis  $\{|v_i\rangle\}$ .

In physical terms, this approach to the Choi-Kraus decomposition offers an important insight that CP maps can in fact be used to represent noisy interactions in physical systems. Examples of such noise processes include sending a photon through a lossy optic fibre or a spin in a random magnetic field. Any physical noise affecting a system  $A$  is typically thought of as resulting from unwanted interaction with an *environment* which is represented by the system  $C$  here. The total evolution of the system + environment is always *unitary* (a restriction of the isometry  $U$ ) and the noise results from the act of performing the partial trace which physically corresponds to the fact that we do not have access to complete information about the environment.

### 4.1.1 Conditions for Zero-Error Quantum Communication

In the context of quantum communication, a quantum channel described by the CPTP map  $\mathcal{T} : \mathcal{B}(\mathcal{H}_A) \rightarrow \mathcal{B}(\mathcal{H}_B)$  represents a *process*: it takes as input, states  $\rho \in \mathcal{B}(\mathcal{H}_A)$  and produces corresponding states  $\mathcal{T}(\rho) \in \mathcal{B}(\mathcal{H}_B)$  as output. It could model an information transmission process that transmits some set of input signals from one location to another, or, it could model a data storage scenario in which some information is input into a noisy memory at one time to be retrieved later.

A classical channel  $\mathcal{N}$ , in the Shannon formulation, is simply characterized by a kernel or a probability transition function  $N(Y|X)$ .  $\{N(y|x) \geq 0\}$  are the conditional probabilities of obtaining output  $y \in Y$  given input  $x \in X$ , so that,  $\sum_y N(y|x) = 1$ .  $X$  and  $Y$  are often called the input and output *alphabets*

respectively. The probabilities  $\{N(y|x)\}$  thus completely describe the classical channel  $\mathcal{N}$ .

The quantum channel formalism includes a description of such classical channels as well. For example, let  $\{|x\rangle\}$  denote an orthonormal basis for  $\mathcal{H}_A$  and  $\{|y\rangle\}$  denote an orthonormal basis for  $\mathcal{H}_B$ , where the labels  $x$  and  $y$  are drawn from the alphabets  $X$  and  $Y$  respectively. Then, corresponding to the channel  $\mathcal{N} \equiv \{N(y|x)\}$ , we can construct the following map on states  $\rho \in \mathcal{B}(\mathcal{H}_A)$ :

$$\mathcal{T}(\rho) = \sum_{x,y} N(y|x) |y\rangle\langle x|(\rho)|x\rangle\langle y|.$$

It is easy to see that  $\mathcal{T} : \mathcal{B}(\mathcal{H}_A) \rightarrow \mathcal{B}(\mathcal{H}_B)$  is a CPTP map with Kraus operators  $E_{xy} = \sqrt{N(y|x)} |y\rangle\langle x|$ , which maps diagonal matrices to diagonal matrices. Any non-diagonal matrix in  $\mathcal{B}(\mathcal{H}_A)$  is also mapped on to a matrix that is diagonal in the  $\{|y\rangle\}$ -basis. Classical channels are thus a special case of quantum channels.

Apart from the action of the CPTP map, a quantum communication protocol also includes an *encoding* map at the input side and a *decoding* map at the output. Given a set of messages  $\{m = 1, 2, \dots, q\}$ , the encoding map assigns a quantum state  $\rho_m \in \mathcal{B}(\mathcal{H}_A)$  to each message  $m$ . The decoding map has to identify the message  $m$  corresponding to the output  $\mathcal{T}(\rho_m)$  of the channel  $\mathcal{T}$ . In other words, the decoding process has to extract classical information from the output quantum state; this is done via a *quantum measurement*. Recall from the discussion in Sect. 2.1, that the outcome  $M$  of a measurement of state  $\mathcal{T}(\rho_m)$  is a random variable distributed according to some classical probability distribution. Here, we are interested in **zero-error communication**, where the outcome  $M$  is equal to the original message  $m$  with probability 1.

Zero-error transmission via general quantum channels was originally studied in [2, 21] and more recently in [5, 6, 8]. In this section we first review some of this earlier work, highlighting the role of operator systems in the study of zero-error communication. In the next section, we focus on the recent work of Duan et al. [9] where a quantum version of the Lovász  $\vartheta$ -function is introduced in the context of studying the zero-error capacity of quantum channels.

Firstly, note that the requirement of zero-error communication imposes the following constraint on the output states  $\{\mathcal{T}(\rho_m)\}$ .

**Exercise 4.1.3** *There exists a quantum measurement  $\mathcal{M}$  in  $\mathcal{H}_B$  such that the outcome  $M$  corresponding to a measurement of state  $\mathcal{T}(\rho_m)$  is equal to  $m$  with probability 1, if and only if the ranges of the states  $\{\mathcal{T}(\rho_m)\}$  are mutually orthogonal.*

Since the states  $\{\mathcal{T}(\rho_m)\}$  are positive semi-definite operators, the fact their ranges are mutually orthogonal implies the following condition:

$$\text{Tr}[\mathcal{T}(\rho_m)\mathcal{T}(\rho_{m'})] = 0, \quad \forall m \neq m'. \quad (4.3)$$

Suppose we choose a particular Kraus representation for  $\mathcal{T}$ , so that  $\mathcal{T}(\rho) = \sum_i E_i \rho E_i^\dagger$ . It then follows from Eq. (4.3) that,

$$\sum_{i,j} \text{Tr}[E_i \rho_m E_i^\dagger E_j \rho_m E_j^\dagger] = 0, \quad \forall m \neq m'.$$

Note that the orthogonality condition on the ranges of the output states further implies that the input states  $\{\rho_m\}$  can be chosen to be rank-one operators ( $\{\rho_m = |\psi_m\rangle\langle\psi_m|\}$ ) without loss of generality. Therefore, the condition for zero-error communication becomes

$$\begin{aligned} & |\langle\psi_m|E_i^\dagger E_j|\psi_{m'}\rangle|^2 = 0, \quad \forall m \neq m'. \\ & \Rightarrow \langle\psi_m|E_i^\dagger E_j|\psi_{m'}\rangle = 0, \quad \forall m \neq m', \quad \forall i, j. \\ & \Rightarrow \text{Tr}[|\psi_{m'}\rangle\langle\psi_{m'}|E_i^\dagger E_j] = 0, \quad \forall m \neq m', \quad \forall i, j. \end{aligned} \quad (4.4)$$

We have thus obtained the following condition for zero-error communication using the quantum channel  $\mathcal{T}$ .

**Lemma 4.1.4 (Condition for Zero-Error Communication)** *Given a channel  $\mathcal{T}$  with a choice of Kraus operators  $\{E_i\}$ , zero-error communication via  $\mathcal{T}$  is possible if and only if the input states  $\{|\psi_m\rangle \in \mathcal{H}_A\}$  to the channel satisfy the following:  $\forall m \neq m'$ , the operators  $|\psi_{m'}\rangle\langle\psi_m| \in \mathcal{B}(\mathcal{H}_A)$  must be orthogonal to the span*

$$S := \text{span}\{E_i^\dagger E_j, i, j\}, \quad (4.5)$$

with orthogonality defined in terms of the Hilbert-Schmidt inner product.<sup>2</sup>

Note that  $S \subset \mathcal{B}(\mathcal{H}_A)$ , and,  $S = S^\dagger$ . Further, since the channel  $\mathcal{T}$  is trace-preserving,  $\sum_i E_i^\dagger E_i = I$ , so that  $S \ni I$ . This implies that  $S$  is an **Operator System**, as defined in Definition 3.3.6. Since all Kraus representations for  $\mathcal{T}$  give rise to the same subset  $S$ , the above condition is unaffected by the non-uniqueness of the Kraus representation.

The **Complementary Channel**  $\hat{\mathcal{T}}$  and its dual  $\hat{\mathcal{T}}^*$  corresponding to a channel  $\mathcal{T}$ , are defined as follows.

**Definition 4.1.5 (Complementary Channel)** Suppose the channel  $\mathcal{T}$  is given by  $\mathcal{T}(\rho) = \text{Tr}_C[V\rho V^\dagger]$ , where  $V : \mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_C$  is the Stinespring isometry. Then, the Complementary Channel  $\hat{\mathcal{T}} : \mathcal{B}(\mathcal{H}_A) \rightarrow \mathcal{B}(\mathcal{H}_C)$  is defined as:

$$\hat{\mathcal{T}}(\rho) = \text{Tr}_B[V\rho V^\dagger].$$

<sup>2</sup>The Hilbert-Schmidt inner product between two operators  $A, B$  is simply the inner product defined by the trace, namely,  $\langle A, B \rangle = \text{Tr}[A^\dagger B]$ .

The **dual** map  $\hat{\mathcal{T}}^*$  is defined via  $\text{Tr}[\rho\hat{\mathcal{T}}^*(X)] = \text{Tr}[\hat{\mathcal{T}}(\rho)X^\dagger]$ . Then, the following result was shown in [9].

**Observation 4.1.6** Given a channel  $\mathcal{T}$  with a complementary channel  $\hat{\mathcal{T}}$ ,

$$S = \hat{\mathcal{T}}^*(\mathcal{B}(\mathcal{H}_C)), \quad \hat{\mathcal{T}}^* : \mathcal{B}(\mathcal{H}_C) \rightarrow \mathcal{B}(\mathcal{H}_A),$$

where  $S$  is the operator system defined in Eq. (4.5), and,  $\hat{\mathcal{T}}^*$  is the dual to the map  $\hat{\mathcal{T}}$ .

Furthermore, it turns out that every operator system can be realized in this manner [5, 8]:

**Observation 4.1.7** Given an operator system  $S \subset \mathcal{B}(\mathcal{H}_A)$ , there exists a CPTP map  $\mathcal{T}$  with a choice of Kraus operators  $\{E_i\}$  such that  $S = \text{span}\{E_i^\dagger E_j, i, j\}$ .

### 4.1.2 Zero-Error Capacity and Lovasz $\vartheta$ Function

Using the condition for zero-error communication, we next quantify the maximum number of messages  $m$  that can be transmitted reliably through the channel  $\mathcal{T}$ . We begin with the notion of a *quantum independence number* originally defined in [9].

**Definition 4.1.8 (Independence Number of S)** Given an operator system  $S = \text{span}\{E_i^\dagger E_j\}$ , the independence number  $\alpha(S)$  is defined as the maximum value of  $q$ , such that

$$\exists \{|\psi_1\rangle, |\psi_2\rangle, \dots, |\psi_q\rangle\} : \forall m \neq m', |\psi_m\rangle\langle\psi_{m'}| \perp S. \quad (4.6)$$

To understand better the motivation for this definition, consider the example of the classical channel again. If  $\mathcal{T}$  is classical, the Kraus operators of the channel are  $\{E_{xy} = \sqrt{N(y|x)} |y\rangle\langle x|\}$ . The operator system  $S$  in this case is given by

$$S = \text{span}\{E_{x'y'}^\dagger E_{xy}\} = \text{span}\{\sqrt{N(y|x)}\sqrt{N(y'|x')}|x\rangle\langle y|y'\rangle\langle x'|\}.$$

Note that,  $E_{x'y'}^\dagger E_{xy} \neq 0$  iff  $x = x'$  (so that  $y = y'$ ) or  $N(y|x)N(y|x') \neq 0$ , whenever  $x \neq x'$ . The latter condition naturally leads to the notion of the *Confusability Graph* associated with a classical channel, and its *Independence Number*.

**Definition 4.1.9 (Confusability Graph, Independence Number)**

- (i) The confusability graph of a classical channel  $N$  is the graph  $G$  with vertices  $x \in X$  and edges  $x \sim x'$  iff  $\exists y \in Y$  such that  $N(y|x)N(y|x') \neq 0$ .
- (ii) A set of vertices  $X_0 \subset X$  such that no pair of vertices in the set is has an edge between them is said to be an independent set. The maximum size of an

independent set of vertices  $X_0$  in  $G$  is called the Independence Number  $\alpha(G)$  of the graph  $G$ .

The name comes from the fact that the edges  $x \sim x'$  of the graph correspond to *confusable* inputs, that is, inputs  $x, x'$  that map to the same output  $y$ . Thus, the operator system  $S$  corresponding to this classical quantum channel  $\mathcal{T}$  carries information about the structure of the underlying graph  $G$ :

$$\text{span}\{|x\rangle\langle x'| : x = x' \text{ or } x \sim x'\} \quad (4.7)$$

More generally, using this definition, every graph  $G$  gives rise to an operator system  $S$ . Furthermore, if two such operator systems are isomorphic, then the underlying graphs are also isomorphic in a combinatorial sense. Thus, for the special case of an operator system  $S$  coming from a classical channel as in Eq. (4.7),

$$\alpha(S) = \alpha(G).$$

Definition 4.1.8 of the independence number is therefore a generalization of the notion of the independence number of a graph.

Given a graph  $G$ , estimating its independence number  $\alpha(G)$  is known to be an NP-complete problem. Similarly, it was shown that [2] estimating  $\alpha(S)$  for an operator system  $S$  arising from a quantum channel  $\mathcal{T}$  is a QMA-complete problem. QMA (Quantum Merlin-Arthur) is the class of problems that can be solved by a quantum polynomial time algorithm given a quantum witness; it is the quantum generalization of probabilistic version of NP. Rather than estimating  $\alpha(S)$ , in what follows we seek to find upper bounds on  $\alpha(S)$ .

We first rewrite the condition in Eq. (4.6) in terms of positive semi-definite operators. Note that  $|\psi_m\rangle\langle\psi_{m'}| \perp S$  implies that

$$M = \sum_{m \neq m'} |\psi_m\rangle\langle\psi_{m'}| \perp S.$$

Since  $S \ni I$ , the states  $\{|\psi_m\rangle, m = 1, 2, \dots, q\}$  satisfying Eq. (4.6) are mutually orthogonal. Also, the operator  $M + \sum_m |\psi_m\rangle\langle\psi_m|$  is positive semi-definite. Therefore,

$$0 \leq M + \sum_{m=1}^q |\psi_m\rangle\langle\psi_m| = M + I,$$

where the number  $q$  is simply  $q = \|M + \sum_m |\psi_m\rangle\langle\psi_m|\|$ . Recalling that  $\alpha(S)$  is simply the maximum value of  $q$ , we have,

$$\alpha(S) \leq \max_{\substack{M \in \mathcal{B}(\mathcal{H}_A): M \perp S, \\ I + M \geq 0}} \|M + I\| = \vartheta(S). \quad (4.8)$$

The quantity on the RHS, defines a **quantum  $\vartheta$ -function**, as a straightforward generalization of a well known classical quantity. If the operator system  $S$  arises from a graph  $G$  as in Eq. (4.7), then,  $\vartheta(S) = \vartheta(G)$ , where,  $\vartheta(G)$  is the Lovász number of the graph  $G$  [20]. It was shown by Lovász that  $\vartheta(G)$  is an upper bound for the independence number  $\alpha(G)$  and is in fact a semidefinite program [30]. In other words, for the case that  $S$  arises from a graph  $G$ , the optimization problem that evaluates  $\vartheta(S)$  has a well behaved objective function and the optimization constraints are either linear ( $M \perp S$ ) or semi-definite ( $I + M \geq 0$ ). However for a general operator system  $S$ , the optimization problem that evaluates  $\vartheta(S)$  is not a semidefinite program (SDP) in general.

In order to define the **zero-error capacity** of a channel, we move to the asymptotic setting, where we consider several copies ( $n$ ) of the channel in the limit of  $n \rightarrow \infty$ . The operator system corresponding to the  $n$ -fold tensor product of a channel is simply the  $n$ -fold tensor product of the operator systems associated with the original channel. Then, the capacity is formally defined as follows.

**Definition 4.1.10 (Zero-Error Capacity)** The zero-error capacity ( $C_0(S)$ ) of a channel with associated operator system  $S$  is the maximum number of bits that can be transmitted reliably per channel use, in the asymptotic limit.

$$C_0(S) := \lim_{n \rightarrow \infty} \frac{1}{n} \alpha(S^{\otimes n}).$$

The capacity is even harder to compute than the independence number. In fact, it is not even known if  $C_0(S)$  is in general a computable quantity in the sense of Turing! For classical channels, where the operator system  $S$  arises from the confusability graph  $G$ , Lovász obtained an upper bound on  $C_0(S)$ . In this case, the Lovász number satisfies,

$$\vartheta(G_1 \times G_2) = \vartheta(G_1)\vartheta(G_2),$$

which immediately implies that  $C_0(S) \leq \vartheta(G)$ . Till date, this remains the best upper bound on the zero-error capacity of a classical channel.

To gain familiarity with the quantum  $\vartheta$ -function, we evaluate it for two simple examples of quantum channels.

*Example 4.1.11* Consider the channel whose Kraus operators span the entire space  $\mathcal{B}(\mathcal{H}_A)$ . The corresponding operator system is given by  $S = \mathcal{B}(\mathcal{H}_A)$ . Then, the only  $M \geq 0$  satisfying  $M \perp S$  is in fact  $M = 0$ . Therefore, by the definition in Eq. (4.8),  $\vartheta(S) \equiv \vartheta(\mathcal{B}(\mathcal{H}_A)) = 1$ .

Next we consider the case where  $S$  is a multiple of the identity.

*Example 4.1.12* Suppose  $S = \mathbb{C}I$ . Then, the operators  $M \perp S$  have to satisfy  $\text{Tr}[M] = 0$ . Without loss of generality, we may assume  $M = \text{diag}(\lambda_1, \dots, \lambda_n)$  with  $\sum_i \lambda_i = 1$ . Suppose the eigenvalues are ordered as follows:  $\lambda_1 \geq \lambda_2 \geq \dots \lambda_n$ . Furthermore, the positive semi-definiteness constraint on  $(I + M)$  implies  $\lambda_i \geq -1$ .

Therefore,

$$\begin{aligned} \|I + M\| &= 1 + \lambda_1 \leq n. \\ \Rightarrow \vartheta(S) &= \vartheta(\mathbb{C}I) = n = \dim(\mathcal{H}_A). \end{aligned} \quad (4.9)$$

We note the following interesting property of the  $\vartheta$ -function.

**Lemma 4.1.13** *Given two operator systems  $S_1$  and  $S_2$ ,*

$$\vartheta(S_1 \otimes S_2) \geq \vartheta(S_1)\vartheta(S_2). \quad (4.10)$$

*Proof* Suppose the operator  $M_1 \perp S_1$  achieves  $\vartheta(S_1)$  and  $M_2 \perp S_2$  achieves  $\vartheta(S_2)$ . Then, define,

$$I + M := (I + M_1) \otimes (I + M_2).$$

By definition,  $M \perp S_1 \otimes S_2$ . Since the norm is multiplicative under tensor product,

$$\vartheta(S_1 \otimes S_2) \geq \|I + M\| = (\|I + M_1\|)(\|I + M_2\|) = \vartheta(S_1)\vartheta(S_2).$$

□

To see that the inequality in Lemma 4.1.13 can be strict, consider the case where  $S = I_n \otimes \mathcal{B}(\mathbb{C}^n)$ . Then, it is a simple exercise to evaluate  $\vartheta(S)$ .

**Exercise 4.1.14** *Show that*

$$\vartheta(I_n \otimes \mathcal{B}(\mathbb{C}^n)) = n^2.$$

*Hint: Recall dense coding (Sect. 2.2.2)!*

The product of the individual  $\vartheta$ -functions evaluated in Examples 4.1.11 and 4.1.12 is much smaller:

$$\vartheta(I_n)\vartheta(\mathcal{B}(\mathbb{C}^n)) = n < \vartheta(I_n \otimes \mathcal{B}(\mathbb{C}^n)).$$

Thus we have a simple example of the *non-multiplicativity* of the quantum  $\vartheta$ -function.

This non-multiplicativity motivates the definition of a modified  $\vartheta$ -function which can be thought of as a completion of  $\vartheta(S)$ .

**Definition 4.1.15 (Quantum Lovász  $\vartheta$ -Function)** For any operator system  $S$ , define the quantum Lovász function as follows.

$$\tilde{\vartheta}(S) = \sup_{\mathcal{H}_C} \vartheta(S \otimes \mathcal{B}(\mathcal{H}_C)). \quad (4.11)$$



Note how the definition is reminiscent of norm-completion. To clarify the operational significance of this modified  $\vartheta$ -function, we define a related independence number in a modified communication scenario.

**Definition 4.1.16 (Entanglement-Assisted Independence Number)** For any operator system  $S \subset \mathcal{B}(\mathcal{H}_A)$ , the quantity  $\tilde{\alpha}(S)$  is the maximum value of  $q$  such that  $\exists$  Hilbert spaces  $\mathcal{H}_{A_0}, \mathcal{H}_C$  and isometries  $\{V_1, V_2, \dots, V_q\} : \mathcal{H}_{A_0} \rightarrow \mathcal{H}_A \otimes \mathcal{H}_C$ , such that,

$$\forall m \neq m', V_m \rho V_m^\dagger \perp S \otimes \mathcal{B}(\mathcal{H}_C), \forall \rho \in \mathcal{S}(\mathcal{H}_{A_0}).$$

Physically,  $\tilde{\alpha}(S)$  corresponds to the maximum number of messages that can be transmitted reliably in an *entanglement-assisted* communication problem, which allows for some entanglement to be shared beforehand between the sender and the receiver.

It is easy to see that  $\alpha(S) \leq \tilde{\alpha}(S)$ , since the non-multiplicativity of the  $\vartheta$ -function implies  $\vartheta(S) \leq \tilde{\vartheta}(S)$ . Furthermore,  $\tilde{\vartheta}(S)$  is an upper bound for  $\tilde{\alpha}(S)$ , just as  $\vartheta(S)$  is to  $\alpha(S)$ .

**Exercise 4.1.17** Given an operator system  $S$ , the  $\tilde{\vartheta}$ -function is an upper bound for the entanglement-assisted independence number:

$$\tilde{\alpha}(S) \leq \tilde{\vartheta}(S).$$

The following simple observations are left as exercises.

**Exercise 4.1.18 (Larger Operator Systems Have a Smaller Independence Number)** Given two operator systems  $S_1, S_2$  such that  $S_1 \subset S_2$ ,  $\alpha(S_1) \geq \alpha(S_2)$ . This also holds for the  $\tilde{\alpha}$ ,  $\vartheta$  and  $\tilde{\vartheta}$  functions.

**Exercise 4.1.19 (Upper Bounds)** Show that for an operator system  $S \subset \mathcal{B}(\mathcal{H}_A)$ , (a)  $\vartheta(S) \leq \dim(\mathcal{H}_A)$ , and, (b)  $\tilde{\vartheta}(S) \leq (\dim(\mathcal{H}_A))^2$ . Equality holds in both cases when  $S = \mathbb{C}I_{\dim(\mathcal{H}_A)}$ .

In fact, when  $S = \mathbb{C}I_2$ ,  $\tilde{\alpha}(S) = (\dim(\mathcal{H}_A))^2$ , which is easily proved using the idea of superdense coding (Sect. 2.2.2).

*Example 4.1.20 ( $\tilde{\alpha}(S)$  for a Qubit)* For example, consider the simplest case of a qubit, for which  $\dim(\mathcal{H}_A) = 2$  and  $S = \mathbb{C}I_2$ . Since this is the entanglement-assisted communication scenario, suppose the sender and receiver share the maximally entangled state

$$|\psi\rangle_{AC} = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle).$$

The sender modifies the part of the state that belongs to her subsystem via conjugation by a unitary  $V_m \in \{I_2, X, Y, Z\}$ . These four operators constitute a basis for the space of  $2 \times 2$  matrices, and were discussed earlier in the construction of Shor's code. It is easy to check that the states  $\{(V_m \otimes I)|\psi\rangle, m = 1, \dots, 4\}$  are

mutually orthogonal. Thus, once the sender sends across these modified states, the receiver can perfectly distinguish them, implying that  $\tilde{\alpha}(\mathbb{C}I_2) = 4$ .

Comparing with Definition 4.1.16, we see that  $\mathcal{H}_{A_0} = \mathcal{H}_A$  in this case, so that the isometries  $V_m$  become unitaries and the input state is  $\rho = \text{Tr}_C[|\psi\rangle\langle\psi|]$ .

Following [9], we now study the  $\tilde{\vartheta}$ -function in greater detail and show that it is indeed a true generalization of the classical Lovász number. Note that the non-multiplicativity of the  $\vartheta$ -function carries over to the  $\tilde{\vartheta}$ -function. Given two operator systems  $S_1$  and  $S_2$ ,

$$\tilde{\vartheta}(S_1 \otimes S_2) \geq \tilde{\vartheta}(S_1)\tilde{\vartheta}(S_2).$$

However, unlike  $\vartheta(S)$ ,  $\tilde{\vartheta}(S)$  can be computed efficiently via a SDP, analogous to the classical Lovász number  $\vartheta(G)$ .

**Theorem 4.1.21** *For any operator system  $S$ ,  $\tilde{\vartheta}(S)$  is a semidefinite program.*

*Proof* We first show that  $\vartheta(S \otimes \mathcal{B}(\mathcal{H}_C))$  is in fact a semidefinite program. For  $S \subset \mathcal{B}(\mathcal{H}_A)$ , recall that,

$$\vartheta(S \otimes \mathcal{B}(\mathcal{H}_C)) = \max_{|\psi\rangle \in \mathcal{H}_A \otimes \mathcal{H}_C} \langle \psi | I \otimes M | \psi \rangle, \quad (4.12)$$

subject to  $\| |\psi\rangle \| = 1$ ,  $I + M \geq 0$  and  $M \perp S \otimes \mathcal{B}(\mathcal{H}_C)$ . The last constraint is equivalent to the constraint that  $M \in S^\perp \otimes \mathcal{B}(\mathcal{H}_C)$ . The objective function in Eq. (4.12) is a multi-linear function and the constraint on the norm of  $|\psi\rangle$  is a non-linear one. To recast this as a SDP, we use the following trick.

Assume  $\dim(\mathcal{H}_C) = \dim(\mathcal{H}_A)$ . Then,  $\exists \rho \in \mathcal{S}(\mathcal{H}_C)$ , such that for  $|\Phi\rangle = \sum_i |i\rangle|i\rangle$ ,

$$|\psi\rangle = (I \otimes \sqrt{\rho})|\Phi\rangle.$$

Inserting this in Eq. (4.12), the objective function becomes,

$$\langle \Phi | I \otimes \rho + M' | \Phi \rangle, \quad M' = (I \otimes \sqrt{\rho})M(I \otimes \sqrt{\rho}) \in S \otimes \mathcal{B}(\mathcal{H}_C), \quad (4.13)$$

with the constraints  $I \otimes \rho + M' \geq 0$ ,  $\rho \geq 0$  and  $\text{Tr}[\rho] = 1$ . This defines a semidefinite program SDP.

Since  $\tilde{\vartheta}(S)$  involves a further maximization over  $\mathcal{H}_C$  (Eq. (4.11)), we can assume without loss of generality that  $\dim(\mathcal{H}_C) \geq \dim(\mathcal{H}_A)$ . If  $\dim(\mathcal{H}_C) > \dim(\mathcal{H}_A)$ , we can still identify a subspace of  $\mathcal{H}_C$  where we can construct the state  $|\Phi\rangle$  and set up the same optimization problem as in Eq. (4.13). Therefore,

$$\begin{aligned} \tilde{\vartheta}(S) &= \max_{\rho, M'} \langle \Phi | I \otimes \rho + M' | \Phi \rangle, \\ \text{such that : } & I \otimes \rho + M' \geq 0, \quad M' \perp S \otimes \mathcal{B}(\mathcal{H}_C), \\ & \rho \geq 0, \quad \text{Tr}[\rho] = 1. \end{aligned} \quad (4.14)$$

This is indeed an SDP since the objective function is linear and the constraints are either semi-definite or linear.  $\square$

It was further shown by Duan et al. [9] that the **dual**<sup>3</sup> to the optimization problem in Eq. (4.14) is also an SDP of the following form:

$$\begin{aligned} \min_{Y \in S \otimes \mathcal{B}(\mathcal{H}_C)} \quad & \lambda, \\ \text{such that : } \quad & \lambda I \geq \text{Tr}_A[Y], \quad Y \geq |\Phi\rangle\langle\Phi|. \end{aligned} \quad (4.15)$$

Furthermore, strong duality holds here, implying that the two optimization problems in Eqs. (4.14) and (4.15) are equivalent:  $\max\langle\Phi|I \otimes \rho + M'| \Phi\rangle \equiv \min \lambda$ .

The form of the dual in Eq. (4.15) also implies that

$$\tilde{\vartheta}(S_1 \otimes S_2) \leq \tilde{\vartheta}(S_1)\tilde{\vartheta}(S_2).$$

Therefore we see from the SDP formulation that the  $\tilde{\vartheta}$ -function is in fact *multiplicative*:

$$\tilde{\vartheta}(S_1 \otimes S_2) = \tilde{\vartheta}(S_1)\tilde{\vartheta}(S_2).$$

We conclude this section with a few open questions. One important question is regarding the largest dimension of a non-trivial operator system  $S$ . Consider for example  $S \subset M_n$  such that  $S^\perp = \text{diag}(n-1, -1, \dots, -1)$ . It trivially follows that  $\tilde{\vartheta}(S) = n = \vartheta(S)$ . However,  $\alpha(S) = 1$  since there does not exist a rank-1 operator in  $S^\perp$ . It can be shown that  $\tilde{\alpha}(S) = 2$ . This example already shows that there could exist a large gap between the independence numbers  $\alpha(S), \tilde{\alpha}(S)$  and their upper bounds  $\vartheta(S), \tilde{\vartheta}(S)$ . An important open question is therefore that of finding the **entanglement-assisted zero-error capacity** of the operator system  $S$ , which is defined as follows:

$$\mathcal{C}_{0E}(S) := \lim_{k \rightarrow \infty} \frac{1}{k} \log \tilde{\alpha}(S^{\otimes k}).$$

From the values of  $\tilde{\alpha}(S)$  and  $\tilde{\vartheta}(S)$ , it is clear that  $1 \leq \mathcal{C}_{0E} \leq \log n$ . However, even the simple question of whether  $\tilde{\alpha}(S \otimes S) = 4$  or  $\tilde{\alpha} > 4$  remains unanswered for any value of  $n$ . Another line of investigation is to explore further the connection between graphs and operator systems that emerges naturally in this study of zero-error quantum communication.

---

<sup>3</sup>The dual of a convex program [3, 4] is obtained by introducing a variable for each constraint, and a constraint for every co-efficient in the objective function. Thus, the objective function and the constraints get interchanged in the dual problem.

## 4.2 Strong Subadditivity and Its Equality Case

The von Neumann entropy of a state  $\rho \in \mathcal{S}(\mathcal{H}_A)$  of a finite-dimensional Hilbert space  $\mathcal{H}_A$  is defined as

$$S(\rho) := -\text{Tr}[\rho \log \rho] = -\sum_i \lambda_i \log \lambda_i, \quad (4.16)$$

where  $\text{spec}(\rho) = \{\lambda_1, \dots, \lambda_{|A|}\}$ . It is easy to see that  $S(\rho) \geq 0$ , with equality iff  $\rho = |\psi\rangle\langle\psi|$ , a pure state. By analogy with classical Shannon entropy, it is also possible to define joint and conditional entropies for composite quantum systems. The von Neumann entropy of a joint state  $\rho_{AB} \in \mathcal{S}(\mathcal{H}_A \otimes \mathcal{H}_B)$  (where  $\mathcal{H}_A$  and  $\mathcal{H}_B$  are finite-dimensional Hilbert spaces) is thus defined as  $S(\rho_{AB}) = -\text{Tr}(\rho_{AB} \log \rho_{AB})$ .

In this section we will focus on some important inequalities regarding the entropies of states of composite system, and understand the structure of the states that saturate them.

**Definition 4.2.1 (Subadditivity)** For any joint state of a bipartite system  $\rho_{AB} \in \mathcal{S}(\mathcal{H}_A \otimes \mathcal{H}_B)$  with the *reduced states* given by the partial traces  $\rho_A = \text{Tr}_B[\rho_{AB}]$  and  $\rho_B = \text{Tr}_A[\rho_{AB}]$ ,

$$S(\rho_A) + S(\rho_B) - S(\rho_{AB}) \geq 0. \quad (4.17)$$

The quantity on the LHS is in fact the **Quantum Mutual Information**  $I(A : B)$  between systems  $A$  and  $B$ , so that the inequality can also be restated as the positivity of the mutual information.

$$I(A : B) := S(\rho_A) + S(\rho_B) - S(\rho_{AB}) \geq 0 \quad (4.18)$$

Alternately, the subadditivity inequality can also be expressed in terms of the **Relative Entropy** between the joint state  $\rho_{AB}$  and the product state  $\rho_A \otimes \rho_B$ .

**Definition 4.2.2 (Relative Entropy)** The quantum relative entropy between any two states  $\rho$  and  $\sigma$  ( $\rho, \sigma \in \mathcal{S}(\mathcal{H}_A)$ ) is defined as

$$D(\rho \parallel \sigma) := \text{Tr}[\rho(\log \rho - \log \sigma)], \quad (4.19)$$

when  $\text{Range}(\rho) \subset \text{Range}(\sigma)$ , and  $\infty$  otherwise [29].

Note that this definition generalizes the Kullback-Liebler divergence [16] of two probability distributions, just as the von Neumann entropy generalizes the Shannon entropy.

**Exercise 4.2.3** Show that  $I(A : B) = D(\rho_{AB} \parallel \rho_A \otimes \rho_B)$ .

**Solution**

$$\begin{aligned}
 I(A : B) &= S(\rho_A) + S(\rho_B) - S(\rho_{AB}) \\
 &= -\text{Tr}_A[\rho_A \log \rho_A] - \text{Tr}_B[\rho_B \log \rho_B] + \text{Tr}[\rho_{AB} \log \rho_{AB}] \\
 &= \text{Tr}[\rho_{AB} \log \rho_{AB}] - \text{Tr}[\rho_{AB}(\log \rho_A + \log \rho_B)] \\
 &= \text{Tr}[\rho_{AB} \log \rho_{AB}] - \text{Tr}[\rho_{AB} \log(\rho_A \otimes \rho_B)] \\
 &= D(\rho_{AB} \parallel \rho_A \otimes \rho_B), \tag{4.20}
 \end{aligned}$$

where the last step follows from the observation that  $\log(\rho_A \otimes \rho_B) = \log \rho_A + \log \rho_B$ . ■

Thus, subadditivity is simply the statement that the relative entropy  $D(\rho_{AB} \parallel \rho_A \otimes \rho_B) \geq 0$  is positive. Positivity of relative entropy is easy to prove, using the observation that  $D(\rho \parallel \sigma) \geq \text{Tr}[(\rho - \sigma)^2]$ .

**Exercise 4.2.4** Show that  $D(\rho \parallel \sigma) \geq \text{Tr}[(\rho - \sigma)^2] \geq 0$ , with equality iff  $\rho = \sigma$ .

Thus the subadditivity inequality becomes an inequality iff  $\rho_{AB} = \rho_A \otimes \rho_B$ , in other words, if and only if the joint state  $\rho_{AB}$  of the system is in fact a product state.

The subadditivity inequality for two quantum systems can be extended to three systems. This result is often known as **Strong Subadditivity**, and is one of the most important and useful results in quantum information theory.

**Theorem 4.2.5 (Strong Subadditivity [17])** Any tripartite state  $\rho_{ABC} \in \mathcal{S}(\mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_C)$ , with reduced density operators  $\rho_{AB} = \text{Tr}_C[\rho_{ABC}]$ ,  $\rho_{BC} = \text{Tr}_A[\rho_{ABC}]$ , and  $\rho_B = \text{Tr}_C[\rho_{ABC}]$ , satisfies,

$$S(\rho_{AB}) + S(\rho_{BC}) - S(\rho_B) - S(\rho_{ABC}) \geq 0 \tag{4.21}$$

*Proof* To prove the strong subadditivity property, it will again prove useful to interpret the LHS as a mutual information, in particular, the **Conditional Mutual Information**  $I(A : C | B)$ —the mutual information between systems  $A$  and  $C$  given  $B$ —which is defined as

$$\begin{aligned}
 I(A : C | B) &= I(A : BC) - I(A : B) \\
 &= S(\rho_A) + S(\rho_{BC}) - S(\rho_{ABC}) - S(\rho_A) - S(\rho_B) + S(\rho_{AB}) \\
 &= S(\rho_{AB}) + S(\rho_{BC}) - S(\rho_B) - S(\rho_{ABC}) \tag{4.22}
 \end{aligned}$$

Thus strong subadditivity is proved once we establish the positivity of the conditional mutual information. This in turn is proved by first expressing  $I(A : C | B)$  in

terms of relative entropies.

$$\begin{aligned} I(A : C|B) &= D(\rho_{ABC} \parallel \rho_A \otimes \rho_{BC}) - D(\rho_{AB} \parallel \rho_A \otimes \rho_B) \\ &= D(\rho_{ABC} \parallel \rho_A \otimes \rho_{BC}) \\ &\quad - D(\text{Tr}_C[\rho_{ABC}] \parallel \text{Tr}_C[\rho_A \otimes \rho_{BC}]) \end{aligned} \quad (4.23)$$

The final step is to then use Uhlmann's theorem [28] (proved earlier by Lindblad [18] for the finite-dimensional case) on the monotonicity of relative entropy.

**Theorem 4.2.6 (Monotonicity of Relative Entropy Under CPTP Maps)** *For all states  $\rho$  and  $\sigma$  on a space  $\mathcal{H}$ , and quantum operations  $\mathcal{T} : \mathcal{L}(\mathcal{H}) \rightarrow \mathcal{L}(\mathcal{K})$ ,*

$$D(\phi \parallel \sigma) \geq D(\mathcal{T}(\phi) \parallel \mathcal{T}(\sigma)). \quad (4.24)$$

To obtain strong subadditivity, one simply has to make the correspondences  $\phi \equiv \rho_{ABC}$ ,  $\sigma \equiv \rho_A \otimes \rho_{BC}$ , and  $\mathcal{T} \equiv \text{Tr}_C$  (recall that the partial trace is indeed a CPTP map). Then, Uhlmann's theorem implies

$$D(\rho_{ABC} \parallel \rho_A \otimes \rho_{BC}) - D(\text{Tr}_C[\rho_{ABC}] \parallel \text{Tr}_C[\rho_A \otimes \rho_{BC}]) \geq 0, \quad (4.25)$$

which in turn implies strong subadditivity through Eqs. (4.23) and (4.22). In fact, it was shown by Ruskai [24] that the contractive property of the quantum relative entropy under CPTP maps and the positivity of conditional mutual information are equivalent statements.  $\square$

Finally, we note that the corresponding inequalities in the classical setting, namely, the subadditivity and the strong subadditivity of the Shannon entropy are easier to prove, since they follow almost directly from the concavity of the log function.

### 4.2.1 Monotonicity of Relative Entropy: Petz Theorem

The above discussion on strong subadditivity implies that the question of finding the conditions under which equality holds for Eq. (4.21), translates to that of finding the conditions under which Eq. (4.24) describing the monotonicity of the relative entropy (under the partial trace operation) is saturated. Note that there is a trivial case of such an equality, namely, if there exists a quantum operation  $\hat{\mathcal{T}}$  which maps  $\mathcal{T}(\phi)$  to  $\phi$  and  $\mathcal{T}(\sigma)$  to  $\sigma$ . It was shown by Petz [22, 23] that this is in fact the only case of such an equality.

**Theorem 4.2.7 (Saturating Uhlmann's Theorem)** *For states  $\phi, \sigma \in \mathcal{H}$ , and the CPTP map  $\mathcal{T} : \mathcal{L}(\mathcal{H}) \rightarrow \mathcal{L}(\mathcal{K})$ ,  $D(\phi \parallel \sigma) = D(\mathcal{T}(\phi) \parallel \mathcal{T}(\sigma))$  if and only if*

$$(\hat{\mathcal{T}} \circ \mathcal{T})(\phi) = \phi, \quad (4.26)$$

for the map  $\hat{\mathcal{T}} : \mathcal{L}(\mathcal{K}) \rightarrow \mathcal{L}(\mathcal{H})$  given by

$$X \rightarrow \sqrt{\sigma} \mathcal{T}^\dagger ([\mathcal{T}(\sigma)]^{-1/2} X [\mathcal{T}(\sigma)]^{-1/2}) \sqrt{\sigma}, \quad \forall X \in \mathcal{L}(\mathcal{H}) \quad (4.27)$$

where  $\mathcal{T}^\dagger$  is the adjoint map<sup>4</sup> of  $\mathcal{T}$ , and the map  $\mathcal{T}$  is assumed to be such that  $\mathcal{T}(\sigma) > 0$ .

Note that  $(\hat{\mathcal{T}} \circ \mathcal{T})(\sigma) = \sigma$ . In other words, the CPTP map  $\hat{\mathcal{T}}$  is defined to be the partial inverse of  $\mathcal{T}$  on the state  $\sigma$ . Petz's theorem states that such a map saturates monotonicity if and only if  $\hat{\mathcal{T}}$  is also a partial inverse for  $\mathcal{T}$  on  $\phi$ .

To apply the Petz theorem for strong subadditivity, the relevant map  $\mathcal{T}$  is the partial trace over system  $C$ . Thus,  $\mathcal{T} \equiv \text{Tr}_C : \mathcal{L}(ABC) \rightarrow \mathcal{L}(AB)$ . The adjoint map is therefore  $\mathcal{T}^\dagger : \mathcal{L}(AB) \rightarrow \mathcal{L}(ABC)$ , with  $\mathcal{T}^\dagger(X) = X \otimes I_C$ . The states  $\phi$  and  $\sigma$  are respectively,  $\phi = \rho_{ABC}$  and  $\sigma = \rho_A \otimes \rho_{BC}$ . Thus the Petz map defined in Eq. (4.27) is given by

$$\hat{\mathcal{T}} \equiv \mathbb{I}_A \otimes \mathcal{T}''_{B \rightarrow BC} \quad (4.28)$$

with the map  $\mathcal{T}'' : \mathcal{L}(B) \rightarrow \mathcal{L}(BC)$  given by,

$$\mathcal{T}''(X) = \sqrt{\rho_{BC}} (\rho_B^{-1/2} X \rho_B^{-1/2} \otimes I_C) \sqrt{\rho_{BC}}, \quad \forall X \in \mathcal{L}(\mathcal{H}_B). \quad (4.29)$$

It is easy to verify that  $(\hat{\mathcal{T}} \circ \mathcal{T})(\sigma) = \hat{\mathcal{T}}(\rho_A \otimes \rho_B) = \rho_A \otimes \rho_{BC}$ . Thus, Petz's theorem states that the inequality in Eq. (4.25) is saturated if and only if the joint state  $\rho_{ABC}$  is such that

$$(\hat{\mathcal{T}} \circ \mathcal{T})(\rho_{ABC}) = \rho_{ABC} \quad \Rightarrow \quad \hat{\mathcal{T}}(\rho_{AB}) \equiv (\mathbb{I}_A \otimes \mathcal{T}'')(\rho_{AB}) = \rho_{ABC}, \quad (4.30)$$

for the map  $\mathcal{T}''$  defined in Eq. (4.29).

To summarize, Petz's theorem explicitly gives a condition on the states that saturate Uhlmann's theorem on the monotonicity of the quantum relative entropy under the action of a CPTP map  $\mathcal{T}$ . This is done by constructing a map  $\hat{\mathcal{T}}$  corresponding to the map  $\mathcal{T}$  such that  $(\hat{\mathcal{T}} \circ \mathcal{T})$  leaves both the states in the argument of the relative entropy function invariant. This in turn gives us a handle on states that saturate strong subadditivity. Specifically, a tripartite state  $\rho_{ABC}$  saturates the strong subadditivity inequality in Eq. (4.21) if and only if it satisfies Eq. (4.30).

*Remark* Note that the condition  $(\mathbb{I}_A \otimes \mathcal{T}'')(\rho_{AB}) = \rho_{ABC}$  can in fact be thought to characterize a *short Quantum Markov Chain* (see [11] for a more rigorous definition). Given three classical random variables  $A, B, C$ , they form a short Markov chain  $A \rightarrow B \rightarrow C$  iff  $P_{AC|B}(a, c|b) = P_{A|B}(a|b)P_{C|B}(c|b)$ , that is, the random variable  $C$  is independent of  $A$ . Analogously, Eq. (4.30) implies that the

---

<sup>4</sup>For any CPTP map  $\mathcal{T}$  in a Kraus representation  $\{T_i\}$ , the adjoint map is another CPTP map with Kraus operators  $\{T_i^\dagger\}$ .

tripartite state  $\rho_{ABC}$  is such that the state of system  $C$  depends only on the state of  $B$  (via the map  $\mathcal{T}''$ ) and is independent of system  $A$ . In fact, the classical conditional mutual information  $I(A : C|B)$  vanishes iff  $A \rightarrow B \rightarrow C$  form a Markov chain in the order. Petz's Theorem is simply a quantum analogue of this statement!

### 4.2.2 Structure of States that Saturate Strong Subadditivity

The structure of the tripartite states that satisfy Eq. (4.30), can in fact be characterized more explicitly as follows.

**Theorem 4.2.8** *A state  $\rho_{ABC} \in \mathcal{L}(\mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_C)$  satisfies strong subadditivity (Eq. (4.21)) with equality if and only if there exists a decomposition of subsystem  $B$  as*

$$\mathcal{H}_B = \bigoplus_j \mathcal{H}_{b_j^L} \otimes \mathcal{H}_{b_j^R}, \quad (4.31)$$

into a direct sum of tensor products, such that,

$$\rho_{ABC} = \bigoplus_j q_j \rho_{Ab_j^L}^{(j)} \otimes \rho_{b_j^R C}^{(j)}, \quad (4.32)$$

with states  $\rho_{Ab_j^L}^{(j)} \in \mathcal{H}_A \otimes \mathcal{H}_{b_j^L}$  and  $\rho_{b_j^R C}^{(j)} \in \mathcal{H}_{b_j^R} \otimes \mathcal{H}_C$ , and a probability distribution  $\{q_j\}$ .

This result due to Hayden et al. [11] provides an explicit characterization of tripartite states that form a short quantum Markov chain. Equation (4.32) states that for a given  $j$  in the orthogonal sum decomposition of the Hilbert space  $\mathcal{H}_B$ , the state on  $\mathcal{H}_A$  is independent of the state in  $\mathcal{H}_C$ .

*Proof* The sufficiency of the condition is immediately obvious—it is easy to check that states  $\rho_{ABC}$  of the form given in Eq. (4.32) indeed saturate the inequality in Eq. (4.21).

To prove that such a structure is necessary, first note that Eq. (4.30), can be simplified to a condition on  $\rho_{AB}$  alone:

$$\begin{aligned} (\mathbb{I}_A \otimes \mathcal{T}'')(\rho_{AB}) &= \rho_{ABC} \\ \Rightarrow (\mathbb{I}_A \otimes \text{Tr}_C \circ \mathcal{T}'')(\rho_{AB}) &= \rho_{AB}. \end{aligned} \quad (4.33)$$

Define  $\Phi \equiv \text{Tr}_C \circ \mathcal{T}''$ , a quantum operation on  $\mathcal{S}(\mathcal{H}_B)$ . Then, states  $\rho_{AB}$  satisfy a fixed point like equation:  $(\mathbb{I}_A \otimes \Phi)\rho_{AB} = \rho_{AB}$ . For any operator  $M$  on  $\mathcal{H}_A$  with  $0 \leq M \leq I$ , define,

$$\sigma = \frac{1}{p} \text{Tr}_A(\rho_{AB}(M \otimes I)), \quad p = \text{Tr}(\rho_{AB}(M \otimes I)). \quad (4.34)$$



Then,

$$(\mathbb{I}_A \otimes \Phi)\rho_{AB} = \rho_{AB} \iff \Phi(\sigma) = \sigma. \quad (4.35)$$

Varying the operator  $M$  thus gives a family of fixed points  $\mathbf{M} = \{\sigma\}$  for  $\Phi$ .

The next step is to make use of the structure theorem for the (finite-dimensional) matrix algebra of the fixed points. More specialized to the situation on hand, is a result due to Koashi and Imoto 4.2.11 that makes use of the structure theorem to characterize the structure of a family of invariant states. See Sect. 4.2.3 for a formal statement and proof of this result.

As discussed in Theorem 4.2.11 below, the family of invariant states  $\mathbf{M} \equiv \{\sigma\}$  induces a decomposition on the Hilbert space  $\mathcal{H}_B$ :

$$\mathcal{H}_B = \bigoplus_j \mathcal{H}_{b_j^L} \otimes \mathcal{H}_{b_j^R}, \quad (4.36)$$

such that every state  $\sigma \in \mathbf{M}$  can be written as

$$\sigma = \bigoplus_j q_j(\sigma) \rho_j(\sigma) \otimes \omega_j, \quad (4.37)$$

with states  $\rho_j(\sigma) \in \mathcal{H}_{b_j^L}$  that depend on  $\sigma$  and  $\omega_j \in \mathcal{H}_{b_j^R}$  that are constant for all  $\sigma$ . The equivalence in Eq. (4.35) implies that  $\rho_{AB} \in \text{Span}\{\xi \otimes \sigma, \xi \in \mathcal{S}(\mathcal{H}_A)\}$ . This in turn implies the following structure for  $\rho_{AB}$ :

$$\rho_{AB} = \bigoplus_j q_j \rho_{Ab_j^L} \otimes \omega_{b_j^R}. \quad (4.38)$$

Finally, Theorem 4.2.11 also gives the following decomposition for the map  $\Phi$  (see Eq. (4.47)):

$$\Phi = \bigoplus_j \mathbb{I}_{\mathcal{H}_{b_j^L}} \otimes \Phi_j. \quad (4.39)$$

Since  $\Phi = \text{Tr}_C \circ \mathcal{T}''$ , this implies

$$\begin{aligned} \rho_{ABC} &= (\mathbb{I}_A \otimes \mathcal{T}'')(\rho_{AB}) \\ &= \bigoplus_j q_j \rho_{Ab_j^L} \otimes \mathcal{T}''(\omega_j) \\ &= \bigoplus_j q_j \rho_{Ab_j^L} \otimes \rho_{b_j^R C} \quad (\rho_{b_j^R C} = \mathcal{T}''(\omega_j)), \end{aligned} \quad (4.40)$$

as desired. □

This explicit characterization of the states saturating strong subadditivity offers an interesting insight into the physical properties of these states. In particular, it gives a neat condition for the tripartite state  $\rho_{ABC}$  to be separable along the  $A-C$  cut.

**Corollary 4.2.9** *For a state  $\rho_{ABC}$  satisfying strong subadditivity with equality, that is,*

$$I(A : C | B) = S(\rho_{AB}) + S(\rho_{BC}) - S(\rho_{ABC}) - S(\rho_B) = 0, \quad (4.41)$$

*the marginal state  $\rho_{AC}$  is separable. Conversely, for every separable state  $\rho_{AC}$  there exists an extension  $\rho_{ABC}$  such that  $I(A : C | B) = 0$ .*

### 4.2.3 An Operator Algebraic Proof of the Koashi-Imoto Theorem

Recall first the structure theorem for matrix algebras which lies at the heart of the Koashi-Imoto characterization. For notational consistency, we state the theorem for the finite-dimensional Hilbert space  $\mathcal{H}_B$ .

**Lemma 4.2.10** *Let  $\mathcal{A}$  be a  $*$ -subalgebra of  $\mathcal{B}(\mathcal{H}_B)$ , with a finite dimensional  $\mathcal{H}_B$ . Then, there exists a direct sum decomposition*

$$\mathcal{H} = \bigoplus_j \mathcal{H}_{B_j} = \bigoplus_j \mathcal{H}_{b_j^L} \otimes \mathcal{H}_{b_j^R}, \quad (4.42)$$

*such that*

$$\mathcal{A} = \bigoplus_j \mathcal{B}(\mathcal{H}_{b_j^L}) \otimes I_{b_j^R}. \quad (4.43)$$

*For  $X \in \mathcal{B}(\mathcal{H})$ , any completely positive and unital projection  $P^*$  of  $\mathcal{B}(\mathcal{H})$  into  $\mathcal{A}$  is of the form*

$$P^*(X) = \bigoplus_j \text{Tr}_{b_j^R}(\Pi_j X \Pi_j (I_{b_j^L} \otimes \omega_j)) \otimes I_{b_j^R}, \quad (4.44)$$

*with  $\Pi_j$  being projections onto the subspaces  $\mathcal{H}_{b_j^L} \otimes \mathcal{H}_{b_j^R}$ , and  $\omega_j$  being states on  $\mathcal{H}_{b_j^R}$ .*

*Proof* See [26]. □

**Theorem 4.2.11 (Koashi-Imoto [15])** *Given a family of states  $\mathbf{M} = \{\sigma\}$  on a finite-dimensional Hilbert space  $\mathcal{H}_B$ , there exists a decomposition of  $\mathcal{H}_B$  as*

$$\mathcal{H}_B = \bigoplus_j \mathcal{H}_{B_j} = \bigoplus_j \mathcal{H}_{b_j^L} \otimes \mathcal{H}_{b_j^R}, \quad (4.45)$$

into direct sum of tensor products, such that:

(a) The states  $\{\sigma\}$  decompose as

$$\sigma = \bigoplus_j q_j \rho_j(\sigma) \otimes \omega_j, \quad (4.46)$$

where  $\rho_j$  are states on  $\mathcal{H}_{b_j^L}$ ,  $\omega_j$  are states on  $\mathcal{H}_{b_j^R}$  which are constant for all  $\sigma$  and  $q_j$  is a probability distribution over  $j$ .

(b) Every  $\Phi$  which leaves the set  $\{\sigma\}$  invariant, is of the form

$$\Phi|_{\mathcal{H}_{B_j}} = \mathbb{I}_{b_j^L} \otimes \Phi_j, \quad (4.47)$$

where the  $\Phi_j$  are CPTP maps on  $\mathcal{H}_{b_j^R}$  such that  $\Phi_j(\omega_j) = \omega_j$ .

Note that Statement (a) of the theorem was first proved by Lindblad [19], and the proof here closely follows his approach.

*Proof* First, consider the fixed points of the dual (or adjoint) map of  $\Phi$ . Since  $\Phi^*$  is a CP unital map, the fixed points form an algebra, unlike the fixed points of the CPTP map  $\Phi$  which are just a set of states. Furthermore, the following Lemma holds.

**Lemma 4.2.12** *If  $\Phi^*(X) = \sum_i B_i^* X B_i$ , then,  $\mathcal{I}_\Phi = \{X : \Phi^*(X) = X\}$  is a \*-subalgebra and is in fact equal to the commutant of the Kraus operators of  $\Phi^*$ :*

$$\mathcal{I}_\Phi = \{B_i, B_i^*\}' \quad (4.48)$$

*Proof* Using the Kraus representation, it is easy to see that  $\Phi^*(X^*) = (\Phi^*(X))^* = X^*$ . For  $X \in \mathcal{I}_\Phi$ , direct computation gives:

$$\sum_i [X, B_i]^* [X, B_i] = \mathcal{F}^*(X^* X) - X^* X = 0. \quad (4.49)$$

The last equality follows from the fact that the family of invariant states  $\mathbf{M} = \{\sigma\}$  contains a faithful state. Since the LHS is a sum of positive terms, all of them must be 0, so that  $[X, B_i] = 0$  for all  $i$ . Similarly,  $[X, B_i^*] = 0$  for all  $i$ , which immediately implies  $\mathcal{I}_\Phi \subset \{B_i, B_i^*\}'$ . The converse relation that  $\{B_i, B_i^*\}' \subset \mathcal{I}_\Phi$  is trivial.  $\square$

Iterating the map  $\Phi^*$  asymptotically many times gives a projection of the full algebra onto the subalgebra of the fixed points.<sup>5</sup> That is,

$$\mathcal{P}^* = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N (\Phi^*)^n, \quad (4.50)$$

<sup>5</sup>This is essentially a *mean ergodic theorem* for the dual of a quantum operation. See [11] for a proof.

is such that  $\mathcal{P}^*(\sigma) = \sigma$ , for  $\sigma$  defined in Eq. (4.34). And the adjoint of  $\mathcal{P}^*$ ,

$$\mathcal{P} = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N \Phi^n, \quad (4.51)$$

is also a map that leaves the family of states  $\mathbf{M} \equiv \{\sigma\}$  invariant. Using the decomposition for  $\mathcal{P}^*$  from the structure theorem above (Eq. (4.44)), the map  $\mathcal{P}$  has a similar decomposition, as follows. For all  $\xi \in \mathcal{B}(\mathcal{H}_B)$ ,

$$\mathcal{P}(\xi) = \bigoplus_j \text{Tr}_{b_j^R} (\Pi_j \xi \Pi_j) \otimes \omega_j, \quad (4.52)$$

where  $\Pi_j$  are projectors onto the subspaces  $\mathcal{H}_{B_j}$  and  $\omega_j \in \mathcal{S}(\mathcal{H}_{b_j^R})$ . Since  $\mathcal{P}_0(\sigma) = \sigma$  for all  $\sigma \in \mathbf{M}$ , we obtain the desired form of the states:

$$\sigma = \mathcal{P}_0^*(\sigma) = \bigoplus_j q_j \rho_j(\sigma) \otimes \omega_j. \quad (4.53)$$

To characterize the properties of a general  $\Phi$  that leaves a set of states  $\mathbf{M} \equiv \{\sigma\}$  invariant, we look at the set of such maps  $\Phi \equiv \{\Phi\}$  that leave  $\mathbf{M}$  invariant. Taking a suitable convex combination of the maps  $\{\Phi\}$  such that the corresponding invariant algebra  $\mathcal{I}$  is the smallest gives the most stringent decomposition of  $\mathcal{H}_B$ . Define  $\mathcal{I}_0 = \bigcap_{\mathcal{F} \in \mathbf{F}} \mathcal{I}_{\mathcal{F}}$  which is a  $*$ -subalgebra itself. Because all dimensions are finite, this is actually a finite intersection:  $\mathcal{I}_0 = \mathcal{I}_{\mathcal{F}_1} \cap \dots \cap \mathcal{I}_{\mathcal{F}_M}$ . Consider for example,

$$\mathcal{F}_0 = \frac{1}{M} \sum_{\mu=1}^M \mathcal{F}_{\mu}. \quad (4.54)$$

From Lemma 4.2.10, there exists the following decomposition of  $\mathcal{I}_0$ :

$$\mathcal{I}_0 = \mathcal{B}(\mathcal{H}_{b_j^L}) \otimes I_{b_j^R}, \quad (4.55)$$

Since  $\mathcal{I}_0 \subset \mathcal{I}_{\Phi}$ ,  $\Phi^*|_{\mathcal{I}_0} = \mathbb{I}_{\mathcal{I}_0}$ . Explicitly, for  $\rho \in \mathcal{S}(\mathcal{H}_{b_j^L})$ ,

$$\Phi^*(\rho \otimes I_{\mathcal{H}_{b_j^R}}) = \rho \otimes I_{\mathcal{H}_{b_j^R}}. \quad (4.56)$$

Now, consider  $\mu \in \mathcal{B}(\mathcal{H}_{b_j^R})$  such that  $0 \leq \mu \leq I$ . Then,

$$0 \leq \Phi^*(\rho \otimes \mu) \leq \Phi^*(\rho \otimes I_{\mathcal{H}_{b_j^L}}) = \rho \otimes I_{\mathcal{H}_{b_j^R}} = I_{\mathcal{H}_{b_j^L}} \otimes I_{\mathcal{H}_{b_j^R}} \quad (4.57)$$

This implies that  $\Phi^*$  maps  $\mathcal{B}(\mathcal{H}_{b_j^L} \otimes \mathcal{H}_{b_j^R})$  into itself for all  $j$ , and hence the same applies to  $\Phi$ . Applying Eq. (4.57) to a rank-one projector  $|\psi\rangle\langle\psi| \in \mathcal{S}(\mathcal{H}_{b_j^L})$ , we get,

$$\Phi^*(|\psi\rangle\langle\psi| \otimes \mu) = |\psi\rangle\langle\psi| \otimes \mu', \quad (4.58)$$

with  $\mu'$  depending linearly on  $\mu$ , and independent of  $|\psi\rangle\langle\psi|$ . Thus, for states  $\rho \in \mathcal{S}(\mathcal{H}_{b_j^L})$  and  $\sigma \in \mathcal{S}(\mathcal{H}_{b_j^R})$ ,

$$\Phi(\rho \otimes \sigma) = \rho \otimes \Phi_j(\sigma). \quad (4.59)$$

Applying this to the invariant states  $\sigma$  gives the invariance of  $\omega_j$  under  $\Phi_j$ .  $\square$

### 4.3 Norms on Quantum States and Channels

We motivate this section by recalling the mathematical formalism of quantum error correction discussed in detail in Sect. 3.2. Given a *noise channel* (a CPTP map)  $\mathcal{T} : \mathcal{B}(\mathcal{H}_A) \rightarrow \mathcal{B}(\mathcal{H}_B)$ , the objective of error correction is to (a) identify a subspace  $\mathcal{H}_{A_0} \subset \mathcal{H}_A$ , which translates into an embedding  $\mathcal{I}_0$  of matrices on  $\mathcal{H}_{A_0}$  completing them as matrices on  $\mathcal{H}_A$ , and, (b) find a CPTP decoding map  $\mathcal{D} : \mathcal{B}(\mathcal{H}_A) \rightarrow \mathcal{B}(\mathcal{H}_{A_0})$ . Perfect error correction demands that any state  $\rho \in \mathcal{B}(\mathcal{H}_{A_0})$  should therefore transform as

$$\mathcal{D}(\mathcal{T}(\mathcal{I}_0(\rho))) = \rho, \quad \forall \rho \in \mathcal{B}(\mathcal{H}_{A_0}).$$

Mathematically, this gives rise to a closed, algebraic theory. But in practice, the condition of perfect error correction is rather hard to achieve. The best we can hope for is that the overall map on  $\rho$  (including noise and decoding) is something close to an identity map. In order to describe such an almost perfect error correction, we need an appropriate choice of norms and metrics on states and channels. In particular, for some choice of embedding  $\mathcal{I}_0$  and decoding map  $\mathcal{D}$ , we should be able to quantify how close  $\mathcal{D}(\mathcal{T}(\mathcal{I}_0(\rho)))$  is to the state  $\rho$ , as well as, the closeness of the map  $\mathcal{D} \circ \mathcal{T} \circ \mathcal{I}_0$  to the identity map  $\mathbb{I}_{A_0}$  on  $\mathcal{H}_{A_0}$ .

We begin with a few standard definitions.

**Definition 4.3.1 (Operator Norm)** For any operator  $X$  on a finite dimensional Hilbert space, the operator norm is defined as

$$\|X\| = \text{largest singular value}(X) = \|X\|_\infty \quad (4.60)$$

This is a limiting case ( $p \rightarrow \infty$ ) of the non-commutative  $L^p$  **norms**:

$$\|X\|_p = (\text{Tr}[|X|^p])^{1/p}, \quad |X| = \sqrt{X^\dagger X}. \quad (4.61)$$

We also recall the following well known inequalities. For any  $X \in \mathcal{B}(\mathcal{H}_A)$ , and  $p \leq q$ ,

$$\|X\|_\infty \leq \|X\|_p \leq \|X\|_q \leq \|X\|_1 \leq \sqrt{\dim(\mathcal{H}_A)} \|X\|_2.$$

The norm of choice to quantify distance between quantum states is the  $L^1$ -norm, which is often referred to as the trace-norm.

**Definition 4.3.2 (Distance Between States)** The distance between any two states  $\rho, \sigma \in \mathcal{S}(\mathcal{H}_A)$  is quantified via the  $L^1$ -norm (trace-norm) as  $\|\rho - \sigma\|_1$ .

The most compelling reason for this choice of definition is that it captures exactly how well the states  $\rho, \sigma$  can be distinguished in an experimental setting. To see this clearly, we briefly discuss the problem of **quantum state discrimination**.

State discrimination is an information-theoretic game which can be described as follows. Assume that some experimental procedure prepares, with equal probability, one of two states  $\rho$  or  $\sigma$ . The goal of the game is to minimize the error in guessing which of these two states was prepared, based on a suitable measurement. Since the goal is to distinguish between two states, it suffices for the measurement process to have just two outcomes. Without loss of generality, we can characterize the measurement apparatus by a pair of positive operators  $\{M, I - M\}$ , with  $0 \leq M \leq I$ . Even if a more complex measuring apparatus were to be used, the outcomes will have to be grouped together so that one set corresponds to  $\rho$  and the other to  $\sigma$ . Say,  $M$  corresponds to guessing  $\rho$ . Then, the error probability is given by,

$$p_{\text{err}} = \frac{1}{2} \text{Tr}[\rho(I - M)] + \frac{1}{2} \text{Tr}[\sigma M]. \quad (4.62)$$

Clearly,  $p_{\text{err}} \leq \frac{1}{2}$ , where the upper bound corresponds to making a random guess without really performing a measurement (that is,  $M = I$ ). Defining the bias  $\beta$  to be  $\beta = 1 - 2p_{\text{err}}$ , we have,

$$\beta = \text{Tr}[(\rho - \sigma)M]. \quad (4.63)$$

The optimal measurement is the one that maximizes  $\eta$ . It is easy to see that the optimal  $M$  is simply the projection onto the subspace where  $(\rho - \sigma) \geq 0$ . Therefore,

$$\beta_{\text{max}} = \frac{1}{2} \|\rho - \sigma\|_1. \quad (4.64)$$

This proves that the minimum error probability is given by

$$\min[p_{\text{err}}] = \frac{1}{2} (1 - \|\rho - \sigma\|_1), \quad (4.65)$$

thus providing an operational meaning to the trace-norm [12]. In mathematical terms, we have simply realized the fact that the trace-norm is the dual of the operator norm:

$$\|\xi\|_1 = \max_{\|X\| \leq 1} |\text{Tr}[\xi X]|. \quad (4.66)$$

This is easily seen once we can rewrite the bias as follows:

$$\beta = \frac{1}{2} \text{Tr}[(\rho - \sigma)(2M - I)]. \quad (4.67)$$

Since  $(\rho - \sigma)$  is Hermitian, to maximize  $\beta$ , it suffices to maximize over all Hermitian operators of norm less than unity.

Note that  $\|\rho - \sigma\|_1$  is an upper bound on the Kolmogorov distance (the  $L^1$  distance) between the measurement statistics corresponding to a given measurement on  $\rho$  and  $\sigma$ . The optimizing measurement that achieves equality is in fact the one that minimizes  $p_{\text{err}}$  in the state discrimination problem.

We move on to distances between channels, or more generally, linear maps  $\mathcal{L} : \mathcal{B}(\mathcal{H}_A) \rightarrow \mathcal{B}(\mathcal{H}_B)$ . The trace-norm on states naturally induces the following norm on maps.

$$\|\mathcal{L}\|_{1 \rightarrow 1} = \max_{\|\xi\|_1} \|\mathcal{L}(\xi)\|_1. \quad (4.68)$$

The so-called **diamond norm** for quantum channels which is used often in the quantum information literature is defined as the completion of this induced norm.

$$\sup_C \|\mathcal{L} \otimes \mathcal{I}_C\|_{1 \rightarrow 1} := \|\mathcal{L}\|_{\diamond}. \quad (4.69)$$

It follows from the duality between the trace-norm and the operator norm, the  $\diamond$ -norm is the dual of the CB-norm (Definition 1.1.2).

The  $\diamond$ -norm also has an operational significance similar to that of the trace-norm. Consider  $\mathcal{L} = \mathcal{T}_1 - \mathcal{T}_2$ , where  $\mathcal{T}_1 : \mathcal{B}(\mathcal{H}_A) \rightarrow \mathcal{B}(\mathcal{H}_B)$  and  $\mathcal{T}_2 : \mathcal{B}(\mathcal{H}_A) \rightarrow \mathcal{B}(\mathcal{H}_B)$  are both CPTP maps. We seek to evaluate  $\|\mathcal{L}\|_{1 \rightarrow 1} \equiv \|\mathcal{T}_1 - \mathcal{T}_2\|_{1 \rightarrow 1}$ . Noting that it suffices to maximize over operators of unit-trace rather than operators of unit-norm, we have,

$$\|\mathcal{T}_1 - \mathcal{T}_2\|_{1 \rightarrow 1} = \max_{\rho \in \mathcal{S}(\mathcal{H}_A)} \|\mathcal{T}_1(\rho) - \mathcal{T}_2(\rho)\|_1. \quad (4.70)$$

This quantity is simply twice the maximum bias of distinguishing between the channels  $\mathcal{T}_1$  and  $\mathcal{T}_2$ , with respect to a restricted set of strategies, namely, the “prepare and measure the output” strategies.

Similarly, when we consider the  $\diamond$ -norm  $\|\mathcal{T}_1 - \mathcal{T}_2\|_{\diamond}$ , it again suffices to minimize over all states in the extended Hilbert space  $\mathcal{H}_A \otimes \mathcal{H}_C$ , including of course

the entangled states. Therefore,

$$\| \mathcal{T}_1 - \mathcal{T}_2 \|_{\diamond} = \max_{\rho \in \mathcal{S}(\mathcal{H}_A \otimes \mathcal{H}_{A'})} \| (\mathcal{T}_1 \otimes \mathcal{I}_{A'}) (\rho) - (\mathcal{T}_2 \otimes \mathcal{I}_{A'}) (\rho) \|_1. \quad (4.71)$$

This is also twice the maximum bias of distinguishing  $\mathcal{T}_1$  and  $\mathcal{T}_2$ , but allowing for all possible strategies.

Thus, both the simple induced norm ( $\| \cdot \|_1$ ) and the  $\diamond$ -norm have a nice operational interpretation in the quantum information theoretic setting. The mathematical properties of these norms have interesting physical consequences in this operational setting, and vice-versa. For example, if we take Eq.(4.65) as the definition of the trace-norm, it immediately follows that the norm must be contractive under CPTP maps. If not, we could always perform the measurement after the action of a CPTP map and distinguish the states with lesser probability of error. Similarly, it follows from the operational interpretation that the  $1 \rightarrow 1$ -norm or the  $\diamond$ -norm cannot increase when the channel is preceded or succeeded by the action of another CPTP map.

We now present an example where the simple induced norm ( $\| \cdot \|_{1 \rightarrow 1}$ ) and its norm completion (the  $\diamond$ -norm) behave rather differently. Let  $\mathcal{T}_1 : M_d \rightarrow M_d$  and  $\mathcal{T}_2 : M_d \rightarrow M_d$  be two CPTP maps acting on the space of  $d \times d$  matrices. We define these maps by their action on the following state  $\rho \in M_d \otimes M_d$ :

$$\rho = |\phi\rangle\langle\phi| = \frac{1}{d} \sum_{i,j=1}^d |ii\rangle\langle jj|.$$

Note that  $|\phi\rangle = \frac{1}{\sqrt{d}} \sum_i |ii\rangle$  is the maximally entangled state in  $M_d \otimes M_d$ . Under the action of  $\mathcal{T}_1$  and  $\mathcal{T}_2$ ,  $\rho$  transforms as:

$$\begin{aligned} (\mathcal{T}_1 \otimes \mathbb{I})\rho &:= \alpha = \frac{1 - F}{d(d-1)}, \\ (\mathcal{T}_2 \otimes \mathbb{I})\rho &:= \sigma = \frac{1 + F}{d(d+1)}, \end{aligned} \quad (4.72)$$

where,  $F = \sum_{i,j} |ij\rangle\langle ji|$  is the so-called SWAP operator—the unitary operator that interchanges the two Hilbert spaces.  $F$  has eigenvalues  $\pm 1$ , and so,  $1 + F$  and  $1 - F$  are projections onto two mutually orthogonal subspaces. Thus, upto suitable normalization,  $\alpha$  is the projector onto the antisymmetric subspace and  $\sigma$  is the projector onto the symmetric subspace of  $M_d \otimes M_d$ .

It suffices to define the maps  $\mathcal{T}_1, \mathcal{T}_2$  via Eq.(4.72), since  $\alpha$  and  $\sigma$  are simply the Choi-Jamiolkowski matrices corresponding to  $\mathcal{T}_1$  and  $\mathcal{T}_2$  respectively. Since we have already identified a state  $\rho$  which is mapped onto orthogonal subspaces by  $\mathcal{T}_1 \otimes \mathbb{I}$  and  $\mathcal{T}_2 \otimes \mathbb{I}$ , the  $\diamond$ -norm is easily computed.

$$\| \mathcal{T}_2 - \mathcal{T}_1 \|_{\diamond} = \| \sigma - \alpha \|_1 = 2. \quad (4.73)$$



This follows from the fact that trace-norm between any two density matrices cannot be larger than 2, and the maximum value of 2 is attained iff the states are orthogonal.

On the other hand, to compute  $\| \mathcal{T}_1 - \mathcal{T}_2 \|_{1 \rightarrow 1}$  we still need to perform a maximization over states:

$$\| \mathcal{T}_1 - \mathcal{T}_2 \|_{1 \rightarrow 1} = \max_{\rho} \| (\mathcal{T}_1 - \mathcal{T}_2)\rho \|_1 .$$

Since the trace-norm is convex, the maximum will be attained for an extreme point on the set of states, namely a pure state. In fact, it suffices to evaluate  $\| (\mathcal{T}_1 - \mathcal{T}_2)|0\rangle\langle 0| \|_1$  for some arbitrary fixed state  $|0\rangle\langle 0|$ , because of the following observation. The states  $\alpha, \sigma$  have the property that they are left invariant under conjugation by any unitary  $U \otimes U \in \mathcal{U}(M_d \otimes M_d)$ :

$$(U \otimes U)\alpha(U \otimes U)^\dagger = \alpha,$$

and similarly for  $\sigma$ . The action of the map  $\mathcal{T}_1$  on any state  $X \in M_d$  is obtained from the corresponding Choi-Jamiolkowski matrix  $\alpha$  as follows:

$$\mathcal{T}_1(X) = d \text{Tr}_2[(I \otimes X^T)\alpha],$$

where the partial trace is over the second system. Thus, the invariance of  $\alpha$  under conjugation by unitaries translates into a covariance property for the channels, and so the norm  $\| (\mathcal{T}_1 - \mathcal{T}_2)\rho \|_1$  is the same for  $\rho$  and  $U\rho U^\dagger$ . Therefore, we have,

$$\| \mathcal{T}_1 - \mathcal{T}_2 \|_{1 \rightarrow 1} = \| (\mathcal{T}_1 - \mathcal{T}_2)|0\rangle\langle 0| \|_1 = \frac{4}{d+1}. \quad (4.74)$$

The value of the RHS follows from the fact that the states  $\mathcal{T}_i(|0\rangle\langle 0|)$ ,  $i = 1, 2$  for some arbitrary rank-1 projector  $|0\rangle\langle 0|$ , are close to the maximally mixed state  $\frac{I}{d}$ .

This example clearly highlights that the naive  $\| \cdot \|_{1 \rightarrow 1}$ -norm can in fact be smaller than the  $\diamond$ -norm by a factor that scales as the dimension of the system. Operationally, this difference between the two norms has an important consequence. Equation (4.73) implies that the two channels in this example can in fact be distinguished perfectly, provided the experimenter has access to the right equipment. In particular, the experimenter needs to be able to create the maximally entangled state  $|\phi\rangle = \frac{1}{\sqrt{d}} \sum_i |ii\rangle$ , store it and in the end be able to perform measurements on the composite system. In short, the experimenter needs access to a quantum computer! On the other hand, if the experimenter is restricted to performing local measurements, the distinguishability, which is now characterized by  $\| \mathcal{T}_1 - \mathcal{T}_2 \|_{1 \rightarrow 1}$ , is rather small.

This observation also has significance in a quantum cryptographic setting. Suppose we consider the problem of distinguishing the states  $\alpha, \sigma \in M_d \otimes M_d$ , when the experimenter is restricted to performing only local operations and using classical communication. This restriction defines a set of operations called **LOCC**, which is in fact a convex subset of  $\mathcal{B}(\mathcal{H})$  whose elements lie between  $[-I, I]$ . Just

as the trace-norm quantifies the distinguishability allowing for arbitrary possible measurements, the distinguishability of the states subject to this restricted set<sup>6</sup> can be characterized by the so-called LOCC-norm. Denoting this norm as  $\|\cdot\|_{\text{LOCC}}$ , it was shown that [7] for the states  $\alpha, \sigma$  defined in the above example,

$$\|\alpha - \sigma\|_{\text{LOCC}} = \frac{4}{d+1}.$$

It is thus possible to encode information in two perfectly distinguishable states  $\alpha$  and  $\sigma$ , which are however close to indistinguishable under a restricted set of operations (like LOCC). This observation leads to a quantum cryptographic scheme called **quantum data-hiding** [7].

## 4.4 Matrix-Valued Random Variables

We begin with a brief introduction to the standard theory of large deviations in classical probability theory. Recall that for some set  $\{X_1, X_2, \dots, X_n\}$  of independent, identically distributed (i.i.d.) random variables taking values in  $\mathbb{R}$ , the **law of large numbers** states that

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n X_i = \mathbb{E}[X] \equiv \mu,$$

with probability 1. In statistics and other applications, we are often interested in how significantly the empirical mean on the LHS deviates from the expectation value  $\mu$ . Here, we focus on large deviations. In particular, for finite  $n$ , we are interested in bounding the probability that the empirical mean is larger than some  $\alpha > \mu$ . All known bounds of this form require some additional knowledge about the distribution of the  $X$ .

The simplest and most commonly encountered large deviation setting is one where the  $X_i$ 's are bounded, so that all higher moments exist. Specifically, assuming that  $X_i \in [0, 1]$ , the following bound is known to be asymptotically optimal.

$$\Pr \left\{ \frac{1}{n} \sum_{i=1}^n X_i > \alpha \right\} \leq e^{-nD(\alpha \parallel \mu)}, \quad (\alpha > \mu) \quad (4.75)$$

---

<sup>6</sup>Any convex, centrally symmetric subset  $\mathbb{M} \subset [-\mathcal{I}, \mathcal{I}] \subset \mathcal{B}(\mathcal{H})$  with the property  $-\mathbb{M} = \mathbb{M}$ , gives rise to a norm, the dual of which is a measure of the distinguishability.

where,

$$D(\alpha \parallel \mu) = \alpha \ln \frac{\alpha}{\mu} + (1 - \alpha) \ln \frac{1 - \alpha}{1 - \mu}$$

is the relative entropy between two binary variables. Note that  $D(\alpha \parallel \mu) \geq 0$  for  $\alpha > \mu$  with equality holding iff  $\alpha = \mu$ . The upper bound in Eq. (4.75) can further be simplified to

$$\Pr \left\{ \frac{1}{n} \sum_{i=1}^n X_i > \alpha \right\} \leq e^{-2n(\alpha - \mu)^2}.$$

We briefly sketch the proof of the large deviation bound in Eq. (4.75).

*Proof* Introducing a real parameter  $t > 0$ , it is easy to see that,

$$\Pr \left\{ \frac{1}{n} \sum_{i=1}^n X_i > \alpha \right\} = \Pr \left\{ e^{t \sum_i X_i} > e^{nt\alpha} \right\}$$

Note that we could have chosen any monotonic function of the two quantities  $\frac{1}{n} \sum_{i=1}^n X_i$  and  $\alpha$  on the RHS. It turns out that the choice of the exponential function is indeed asymptotically optimal.

The next step is to use Markov's inequality which states that the probability that a positive-valued random variable  $Y$  is greater than some positive constant  $a$  is upper bound by  $\mathbb{E}[Y]/a$ . This implies,

$$\Pr \left\{ \frac{1}{n} \sum_{i=1}^n X_i > \alpha \right\} \leq \frac{\mathbb{E}[e^{t \sum_i X_i}]}{e^{nt\alpha}} = (\mathbb{E}[e^{tX - t\alpha}])^n,$$

where the second equality follows from the fact that the  $X_i$  are independent and identically distributed.

From the convexity of the exponential function, it follows that for fixed  $\mu$  and  $X \in [0, 1]$ , the expectation on the RHS is maximized when  $X$  is a Bernoulli variable with  $\Pr(X = 0) = 1 - \mu$ ,  $\Pr(X = 1) = \mu$ . Therefore,

$$\Pr \left\{ \frac{1}{n} \sum_{i=1}^n X_i > \alpha \right\} \leq ((1 - \mu)e^{-t\alpha} + \mu e^{t(1-\alpha)})^n.$$

Optimizing the RHS over the parameter  $t$  yields the desired bound in Eq. (4.75).  $\square$

The same large deviation bound holds for real-valued vector variables, upto a dimensional constant that comes from the union bound, when we seek to quantify the deviation of each co-ordinate from its mean value. We would like to obtain similar tail bounds for matrices.

### 4.4.1 Matrix Tail Bounds

Consider a set of matrices  $X_1, X_2, \dots, X_n, \dots \in \mathcal{B}(\mathcal{H})$  that are i.i.d., where each  $X_i \in [0, I]$  and  $\mathcal{H}$  is a finite dimensional Hilbert space. Note that this setting is rather different from that of random matrix theory, where the entries of the matrix are chosen in an i.i.d fashion. Here, we do not care about the distribution over the entries of the matrices, rather each  $X_i$  is chosen independently from the same distribution. Here again, in the large- $n$  limit, the empirical mean converges to the expectation value  $\mathbb{E}[X] = M \in [0, I]$  with probability 1.

$$\frac{1}{n} \sum_{i=1}^n X_i \rightarrow \mathbb{E}[X] \equiv M.$$

The corresponding large deviation problem seeks to bound the following probability:

$$\Pr \left\{ \frac{1}{n} \sum_{i=1}^n X_i \not\leq A \right\},$$

given a positive matrix  $A > M$ . In order to obtain such a bound, we first need a matrix version of Markov's inequality. The following lemma is left as an exercise; it simply follows from the proof of the standard (classical) Markov's inequality.

**Lemma 4.4.1 (Matrix Markov Inequality)** *Let  $X \geq 0$  be a positive semi-definite matrix random variable with  $\mathbb{E}(X) = M$ . Then, for some positive definite matrix  $A > 0$ ,*

$$\Pr \{X \not\leq A\} \leq \text{Tr}[MA^{-1}]. \quad (4.76)$$

We will henceforth assume that  $\mathbb{E}[X] = M \geq \mu I$ , for  $\mu > 0$ . This is a natural assumption to make in sampling problems. If the probability of a certain event is too small, then we will have to sample a very large number of times to get an estimate of this probability. Assuming that the expectation value is larger than a certain minimum value excludes such rare events. In our setting, this assumption implies that the mean does not have very small eigenvalues. Then, defining the operator  $Y = \mu M^{-1/2} X M^{-1/2}$ , we see that  $Y \in [0, I]$ , and,  $\mathbb{E}[Y] = \mu I$ . Further,

$$\Pr \{X \not\leq A\} = \Pr \{Y \not\leq A'\}, \quad A' = \mu M^{-1/2} A M^{-1/2}.$$

Thus, without loss of generality, we have an operator  $A'$  with the property that it is strictly larger than the expectation value  $\mathbb{E}[X]$ . All of the eigenvalues of  $A'$  are strictly larger than  $\mu$ , so that,  $A' \geq \alpha I > \mu I$ .

In this setting, assuming  $\mathbb{E}[X] = M = \mu I$ , the following large deviation bound was proved in [1].

$$\Pr \left\{ \frac{1}{n} \sum_{i=1}^n X_i \not\leq \alpha I \right\} \leq d e^{-nD(\alpha \parallel \mu)}, \quad \text{for } \alpha > \mu, \quad (4.77)$$

$$\Pr \left\{ \frac{1}{n} \sum_{i=1}^n X_i \not\geq \alpha I \right\} \leq d e^{-nD(\alpha \parallel \mu)}, \quad \text{for } \alpha < \mu.$$

Before proceeding to the proof, we note some useful facts about the relative entropy function.

*Remark 4.4.2* The relative entropy  $D(\alpha \parallel \mu)$  satisfies:

- When  $(\alpha - \mu)$  is fixed,

$$D(\alpha \parallel \mu) \geq 2(\alpha - \mu)^2.$$

- When  $\mu$  is small, for  $\alpha = (1 + \epsilon)\mu$ , a stronger bound holds:

$$D(\alpha \parallel \mu) \geq c\mu\epsilon^2,$$

for some constant  $c$ .

*Proof* We closely follow the arguments in [1] where the matrix tail bounds were originally proved. First, note that,

$$\Pr \left\{ \frac{1}{n} \sum_{i=1}^n X_i \not\leq \alpha I \right\} \leq \Pr \left\{ e^{t \sum_{i=1}^n X_i} \not\leq e^{nt\alpha} I \right\}. \quad (4.78)$$

In the simplified setting we consider, these two probabilities are in fact equal. However, equality does not hold in general, since the function  $x \rightarrow e^x$  is not a matrix monotone. The inverse function  $x \rightarrow \ln x$  is indeed a matrix monotone. This follows from two simple facts: first, for matrices  $X, Y$ ,  $0 \leq X \leq Y$  implies  $Y^{-1} \leq X^{-1}$ . Using this, along with the well known integral representation of the log function:

$$\ln x = \int_{t=1}^x \frac{dt}{t} = \int_{t=0}^{\infty} dt \left( \frac{1}{t+1} - \frac{1}{x+t} \right),$$

it is easy to prove that  $\ln x$  is a matrix monotone.

Invoking the Markov bound (Eq. (4.76)) in Eq. (4.78), we have,

$$\Pr \left\{ \frac{1}{n} \sum_{i=1}^n X_i \not\leq \alpha I \right\} \leq \text{Tr}[\mathbb{E}[e^{t \sum_{i=1}^n X_i}] e^{-nt\alpha}].$$

Note that we cannot make use of the independence argument at this stage because the sum on the exponent does not factor into a product of exponents in this case. The operators  $X_i$  could be non-commuting in general. Instead, we use the Golden-Thompson inequality.

**Lemma 4.4.3 (Golden-Thompson Inequality)**

$$\text{Tr}[e^{A+B}] \leq \text{Tr}[e^A e^B]. \quad (4.79)$$

Using this, we have,

$$\begin{aligned} \Pr \left\{ \frac{1}{n} \sum_{i=1}^n X_i \not\leq \alpha I \right\} &\leq \text{Tr}[\mathbb{E}[e^{tX_n}] \mathbb{E}[e^{t \sum_{i=1}^{n-1} X_i}]] e^{-nt\alpha} \\ &\leq \|\mathbb{E}[e^{tX_n}]\| \text{Tr}[\mathbb{E}[e^{t \sum_{i=1}^{n-1} X_i}]] e^{-nt\alpha}, \end{aligned} \quad (4.80)$$

where we have upper bounded the operator  $e^{tX_n}$  by its norm in the final step. Repeating these steps iteratively, we have,

$$\Pr \left\{ \frac{1}{n} \sum_{i=1}^n X_i \not\leq \alpha I \right\} \leq d (\|\mathbb{E}[e^{tX}]\| e^{-t\alpha})^n.$$

Finally, noting that  $\mathbb{E}[e^{tX}]$  is maximized when  $X$  is Bernoulli distributed over  $\{0, \mathbb{I}\}$ , and then optimizing over  $t$ , we get,

$$\Pr \left\{ \frac{1}{n} \sum_{i=1}^n X_i \not\leq \alpha I \right\} \leq d ([1 - \mu + \mu e^t] e^{-t\alpha})^n \leq d e^{-nD(\alpha\|\mu)}. \quad (4.81)$$

□

A stronger tail bound is obtained by avoiding the Golden-Thompson inequality and using Lieb concavity instead [27]. In particular, in place of Eq. (4.80), we have,

$$\Pr \left\{ \frac{1}{n} \sum_{i=1}^n X_i \not\leq \alpha I \right\} \leq e^{-nt\alpha} \text{Tr}[(\mathbb{E}[e^{tX}])^n].$$

This bound is clearly better than the bound in Eq. (4.81) when the operator  $e^{tX}$  has one dominating eigenvalue and the other eigenvalues are quite small. We refer to [27] for further details of this approach.

Such matrix tail bounds were originally obtained by Ahlswede and Winter in the context of a specific quantum information theoretic problem [1]. Here we will focus on two other applications, namely in destroying correlations [10] and in quantum state merging.

### 4.4.2 Destroying Correlations

An important question in quantum information theory as well as statistical physics is to quantify the amount of correlations in a bipartite quantum system. One approach to quantifying the correlations in a bipartite state  $\rho_{AB} \in \mathcal{S}(\mathcal{H}_A \otimes \mathcal{H}_B)$ , proposed in [10], is to understand the process by which  $\rho_{AB}$  can be transformed to an uncorrelated (product) state of the form  $\rho_A \otimes \rho_B$ . The fundamental idea is to shift from characterizing the states to characterizing processes  $\mathcal{T} : \rho_{AB} \rightarrow \rho_A \otimes \rho_B$ , in particular, to quantify the amount of randomness that has to be introduced to effect such a transformation.

For sufficiently large systems  $\mathcal{H}_A, \mathcal{H}_B$ , there always exists a *global* unitary operation, that is, a unitary operator on  $\mathcal{H}_A \otimes \mathcal{H}_B$ , that maps  $\rho_{AB} \rightarrow \rho_A \otimes \rho_B$  deterministically. On the other hand, if our physical model allows only for *local* unitary conjugations of the type  $U_A \otimes I_B$  and mixtures thereof, a natural question to ask is, how much randomness is required for such a transformation.<sup>7</sup> Note that a single local unitary conjugation cannot change the amount of correlations in a bipartite state. However, if we insert a finite amount of randomness by constructing maps involving probabilistic mixtures of local unitaries, a bipartite correlated state can indeed get mapped to a product state.

We thus consider CP maps of the following type:

$$\mathcal{T} : X \rightarrow \sum_i p_i (U_i \otimes I) X (U_i \otimes I)^\dagger,$$

where  $\{U_i\}$  are drawn from the set of unitary matrices on  $\mathcal{H}_A$ . To see a concrete example of such a map, consider the maximally two-qubit state:

$$\rho_{\text{ent}} = \frac{1}{2}(|00\rangle + |11\rangle)(\langle 00| + \langle 11|).$$

Then, choosing the unitaries  $U_0 = I$ ,  $U_1 = X$ ,  $U_2 = Y$  and  $U_3 = Z$ , where  $X, Y, Z$  are the Pauli matrices defined in Sec 3.2.2., and choosing the probabilities  $p_i = \frac{1}{4}$ , we get,

$$\mathcal{T}(\rho_{\text{ent}}) = \sum_i p_i U_i \rho U_i^\dagger = \frac{I}{2} \otimes \frac{I}{2}.$$

We thus have a simple example of transforming a maximally entangled state into a product state via a probabilistic mixture of unitaries.

In order to quantify the amount of noise or randomness involved in the process, one approach is to simply use the difference in the (quantum) entropies between

---

<sup>7</sup>We could equally well have considered unitaries of the type  $I_A \otimes U_B$ , and the same results would hold.

the final and initial states. However, it might be more meaningful to consider the classical entropy of the probability distribution  $\{p_i\}$ , which really captures the thermodynamic cost associated with the process. Physically, the erasing of correlations is in fact a consequence of erasing the knowledge of the probability distribution  $\{p_i\}$  associated with the unitaries  $\{U_i\}$ . *Landauer's principle* states that there is an energy cost associated with erasing information, and in the asymptotic limit, this cost is proportional to the entropy of the distribution  $\{p_i\}$ .

**Definition 4.4.4** Given  $n$  copies of a bipartite state  $\rho_{AB}$ , the family  $(p_i, U_i)_{i=1}^N$  of probabilities  $p_i$  and unitaries  $U_i \in \mathcal{U}(\mathcal{H}_A)$  is said to be  $\epsilon$ -randomizing for  $\rho_{AB}^{\otimes n}$  if the associated map  $\mathcal{T}$  is such that

$$\left\| \sum_{i=1}^N p_i (U_i \otimes I) \rho_{AB}^{\otimes n} (U_i^\dagger \otimes I) - \tilde{\rho}_A \otimes \rho_B^{\otimes n} \right\| \leq \epsilon, \tag{4.82}$$

where  $\tilde{\rho}_A = \sum_i p_i U_i \rho_A^{\otimes n} U_i^\dagger \in \mathcal{H}_A^{\otimes n}$ .

We are interested in quantifying the size of the smallest such  $\epsilon$ -randomizing family  $(p_i, U_i)$  for the state  $\rho_{AB}^{\otimes n}$ , which we denote as  $N(n, \epsilon)$ .

**Definition 4.4.5**  $N(n, \epsilon)$  is defined to be the smallest  $N$  such that  $\exists$  an  $\epsilon$ -randomizing family  $(p_i, U_i)_{i=1}^N$  for  $\rho_{AB}^{\otimes n}$ .

Specifically, we are interested in the asymptotic behavior of  $\frac{\log N(n, \epsilon)}{n}$  in the limit  $\epsilon \rightarrow 0$  and  $n \rightarrow \infty$ , as way to quantify the correlation of the bipartite state  $\rho_{AB}$ . Following the work of Groisman et al. [10], we first prove the following lower bound.

**Theorem 4.4.6** Given  $n$  copies of a bipartite state  $\rho_{AB} \in \mathcal{S}(\mathcal{H}_A \otimes \mathcal{H}_B)$ , any  $\epsilon$ -randomizing set  $\{U_i\}_{i=1}^{N(n, \epsilon)}$  must have at least  $N(n, \epsilon)$  unitaries, where  $N(n, \epsilon)$  satisfies,

$$\liminf_{\substack{n \rightarrow \infty \\ \epsilon \rightarrow 0}} \frac{1}{n} \log N(n, \epsilon) \geq I(A : B)_\rho, \tag{4.83}$$

where,  $I(A : B)_\rho = S(\rho_A) + S(\rho_B) - S(\rho_{AB})$  is the mutual information of the state  $\rho_{AB}$  (see Eq. (4.18)).

*Proof* Let  $(p_i, U_i)_{i=1}^{N(n, \epsilon)}$  be an  $\epsilon$ -randomizing family for  $\rho_{AB}^{\otimes n}$ . The corresponding CP map  $\mathcal{T}$  is given by

$$\mathcal{T}(\rho_{AB}^{\otimes n}) = \sum_{i=1}^N p_i (U_i \otimes I) \rho_{AB}^{\otimes n} (U_i^\dagger \otimes I).$$



Then, it can be shown that the von Neumann entropy of the transformed state satisfies,

$$S(\mathcal{T}(\rho_{AB}^{\otimes n})) \leq S(\rho_{AB}^{\otimes n}) + H(p_1, \dots, p_N), \quad (4.84)$$

where  $H(p_1, \dots, p_N)$  is the Shannon entropy of the distribution  $\{p_i\}$ . The inequality above follows from the concavity of the von Neumann entropy, in particular,

$$S\left(\sum_i p_i \rho_i\right) \leq \sum_i p_i S(\rho_i) + H(p_1, p_2, \dots, p_N).$$

From the definition in Eq. (4.82), the final state  $\mathcal{T}(\rho_{AB}^{\otimes n})$  is close to the product state  $\tilde{\rho}_A \otimes \rho_B^{\otimes n}$ . This implies, via the Fannes inequality,

$$S(\tilde{\rho}_A \otimes \rho_B^{\otimes n}) \leq S(\mathcal{T}(\rho_{AB}^{\otimes n})) + \epsilon \log[(d_A)^n (d_B)^n],$$

where  $d_A = \dim(\mathcal{H}_A)$  and  $d_B = \dim(\mathcal{H}_B)$ . Further, from the definition of  $\tilde{\rho}_A$ , we have,

$$S(\tilde{\rho}_A \otimes \rho_B^{\otimes n}) = nS(\rho_A) + nS(\rho_B).$$

These observations along with the inequality in Eq. (4.84) imply,

$$nS(\rho_A) + nS(\rho_B) - n\epsilon \log(d_A d_B) \leq nS(\rho_{AB}) + \log N(n, \epsilon).$$

Thus we have the following lower bound:

$$\frac{\log N(n, \epsilon)}{n} \geq I(A : B)_\rho - O(\epsilon), \quad (4.85)$$

where  $I(A : B)_\rho = S(\rho_A) + S(\rho_B) - S(\rho_{AB})$  is the mutual information of the state  $\rho_{AB}$ . This in turn implies the asymptotic lower bound,

$$\liminf_{\substack{n \rightarrow \infty \\ \epsilon \rightarrow 0}} \frac{1}{n} \log N(n, \epsilon) \geq I(A : B). \quad (4.86)$$

□

To show that the mutual information is also an upper bound, we make use of the matrix sampling bound proved in the previous section. We will also need to invoke the typicality principle, which we recall here.

**Definition 4.4.7 (Typicality Principle)** Given  $n$  copies of a state  $\rho_{AB}$ , for all  $\epsilon$  and large enough  $n$ , there exists a truncated state  $\rho_{\hat{A}\hat{B}}^{(n)} \in \mathcal{S}(\hat{\mathcal{H}}_A^{\otimes n} \otimes \hat{\mathcal{H}}_B^{\otimes n})$  with the

following properties:

- (i)  $\|\rho_{\hat{A}\hat{B}}^{(n)} - \rho_{AB}^{\otimes n}\|_1 \leq \epsilon$ ,
- (ii)  $\text{Range}(\rho_{\hat{A}\hat{B}}^{(n)}) \subset \hat{\mathcal{H}}_A \otimes \hat{\mathcal{H}}_B$ , where,  $\hat{\mathcal{H}}_A \subset \mathcal{H}_A^{\otimes n}$  and  $\hat{\mathcal{H}}_B \subset \mathcal{H}_B^{\otimes n}$  with the dimensions of the truncated spaces satisfying,

$$d_{\hat{A}} \equiv \dim(\hat{\mathcal{H}}_A) \leq 2^{nS(\rho_A) \pm \epsilon}, \quad d_{\hat{B}} \equiv \dim(\hat{\mathcal{H}}_B) \leq 2^{nS(\rho_B) \pm \epsilon}.$$

- (iii) Asymptotic Equipartition Property:  $\rho_{AB}^{(n)} \approx 2^{-nS(\rho_{AB}) \pm \epsilon}$  on its range. Also,

$$\rho_A^{(n)} \approx 2^{-nS(\rho_A) \pm \epsilon} I_{\hat{A}}, \quad \rho_B^{(n)} \approx 2^{-nS(\rho_B) \pm \epsilon} I_{\hat{B}}.$$

In other words, in the truncated spaces, the reduced states have an almost flat spectrum. The global state also has a near flat spectrum on its range space.

It now suffices to find a map that can destroy the correlations in the truncated state  $\rho_{\hat{A}\hat{B}}^{(n)}$ . We now prove that there always exists such a map which is a probabilistic mixture of  $N(n, \epsilon)$  unitaries, where  $N(n, \epsilon)$  satisfies the following upper bound:

$$\limsup_{\substack{n \rightarrow \infty \\ \epsilon \rightarrow 0}} \frac{1}{n} \log N(n, \epsilon) \leq I(A : B). \quad (4.87)$$

*Proof* The proof is to essentially construct an ensemble from which the random unitaries that make up the map can be picked. First, choose a probability distribution  $\mu$  over the set of all unitaries  $\mathcal{U}(\mathcal{H}_{\hat{A}})$  acting on  $\mathcal{H}_{\hat{A}}$  with the property that

$$\sigma \rightarrow \int d\mu(U) U \sigma U^\dagger = \frac{I_{\hat{A}}}{d_{\hat{A}}}, \quad \forall \sigma \in \mathcal{S}(\hat{\mathcal{H}}_A). \quad (4.88)$$

There are many measures  $\mu$  that satisfy this property. Mathematically, the most useful measure is the *Haar measure*, which is the unique unitary invariant probability measure. This makes the above integral manifestly unitary invariant. The smallest support of such a measure is  $(d_{\hat{A}})^2$ . Another choice of measure can be realized by considering the discrete Weyl-Hiesenberg group. This is the group of unitaries generated by the cyclic shift operator and the diagonal matrix with the  $n^{\text{th}}$  roots of unity along the diagonal. Either choice of measure would work for our purpose.

Next, we draw unitaries  $U_1, U_2, \dots, U_N$ , independently at random from  $\mu$ . Define the operators

$$Y_i = (U_i \otimes I) \rho_{AB}^{(n)} (U_i^\dagger \otimes I).$$

Note that the expectation values of these operators satisfy

$$\mathbb{E}[Y_i] = \frac{1}{d_{\hat{A}}} I_{\hat{A}} \otimes \rho_{\hat{B}}^{(n)}, \quad \forall i.$$

The  $Y_i$ s are not upper bounded by  $I$  but by some exponential factor as stated in the typicality principle. Therefore, we rescale the  $Y_i$ s as follows.

$$X_i = 2^{n(S(\rho_{AB})-\epsilon)} Y_i \in [0, I].$$

Then, the expectation values of the  $X_i$ s satisfy,

$$\begin{aligned} \mathbb{E}[X_i] &= 2^{n(S(\rho_{AB})-\epsilon)} \frac{1}{d_{\hat{A}}} I_{\hat{A}} \otimes \rho_{\hat{B}}^{(n)} \\ &\geq 2^{-n(I(A:B)-3\epsilon)} I_{\hat{A}\hat{B}} \equiv v I_{\hat{A}\hat{B}}, \end{aligned} \quad (4.89)$$

where the second inequality comes from the typicality principle.

Once we have this lower bound on the expectation values, we can use the matrix tail bounds in Eq. (4.77) as follows. Let  $M$  denote the expectation value  $\mathbb{E}[Y_i]$ . Then,

$$\begin{aligned} \Pr \left\{ \frac{1}{N} \sum_{i=1}^N Y_i \not\leq (1 + \epsilon)M \right\} &\leq 2 d_{\hat{A}} d_{\hat{B}} e^{-ND((1+\epsilon)v\|v)} \\ &\leq 2(d_A)^n (d_B)^n e^{-c\epsilon^2 vN}. \end{aligned} \quad (4.90)$$

We can similarly get a lower bound by using the other tail bound in Eq. (4.77) for  $(1 - \epsilon)M$ . If the bound on the RHS is less than one, then, there exist unitaries  $U_1, U_2, \dots, U_N$  with the property

$$(1 - \epsilon)M \leq \frac{1}{N} \sum_{i=1}^N (U_i \otimes I) \rho_{AB}^{(n)} (U_i^\dagger \otimes I) \leq (1 + \epsilon)M. \quad (4.91)$$

Taking the trace-norm on both sides, we see that the set  $\{U_1, U_2, \dots, U_N\}$  is indeed  $\epsilon$ -randomizing as defined in Eq. (4.82). Note that the statement we have proved in Eq. (4.91) is infact a stronger one!

Thus, in order to achieve the  $\epsilon$ -randomizing property, the bound in Eq. (4.90) shows that it suffices to have

$$N \geq \frac{1}{c\epsilon} 2^{-nI(A:B)+3\epsilon} \log[(d_A)^n (d_B)^n] \approx 2^{nI(A:B)+\delta},$$

in the limit of large  $n$ . □