ENTANGLEMENT AND NON-LOCALITY
PMATH 990/QIC 890

SAMUEL J. HARRIS AND SATISH K. PANDEY

Abstract. These are the lecture notes from the course Dr. Paulsen taught at University of Waterloo in Winter 2016. All inner products in these notes are linear in second variable and conjugate linear in first variable. If you find any mistake/typo in these notes, please email us.

Contents

1. Outline

Course Outline:

- Mathematical Background
  - Hilbert space
  - Some basic matrix theory
  - Direct sums of Vector spaces and partitioned matrices
  - Tensor products of Hilbert spaces
- Basics of the Quantum Viewpoint
  - Measurement systems and pure states
  - Quantum distinguishability
  - Ensembles and Mixed States
  - Von Neumann's density matrix
- Theory of CP maps
  - Choi-Krauss reperesentation
  - Non-uniquesness of Choi-Kraus
  - Douglas Factorization theorem
  - An application to Quantum Error Correction
- Zero Error Capacity
  - The classical binary case
  - Graph theory
  - Zero error capacity for quantum channels
  - Entanglement assisted zero-error capacity

Date: April 4, 2016.

  - Introduction to operator systems
  - Distinguishability of outputs
- Quantum Correlations
  - Classical vs Quantum conditional probabilities
  - State purification and POVMs vs PVMs
  - Introduction to C*-algebras
  - Conjectures of Connes and Tsirelson
- Non-Local Games
  - Finite input-output games
  - Games based on Graphs
  - More C*-algebras
  - Values of Games

2. Mathematical Background

We will write $\mathbb{C}^p = \{(\lambda_1, ..., \lambda_p) : \lambda_i \in \mathbb{C}\}$ (so p-tuples), with the canonical basis $e_i = |i\rangle$ for $1 \leq i \leq p$, or $e_i = \delta_i$, where e_i denotes the vector with 1 in the i-th position and 0 in every other position. We write $M_{n,p}$ for $n \times p$ matrices. We usually write an element of $M_{n,p}$ as $A = (a_{ij})$ for $1 \leq i \leq n$ and $1 \leq j \leq p$. The canonical basis for $M_{n,p}$ is given by the matrix units $E_{i,j} = |j\rangle\langle i| = e_j^* e_i$ (which is 1 in the (i,j)-entry and 0 everywhere else).

We will always identify $M_{n,p} \simeq \mathcal{L}(\mathbb{C}^p, \mathbb{C}^n)$ via

$$A \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_p \end{pmatrix} = \begin{pmatrix} \sum_{j=1}^p a_{1j}\lambda_j \\ \vdots \\ \sum_{j=1}^p a_{mj}\lambda_j \end{pmatrix}.$$

Given $A \in M_{n,p}$, we define $A^t \in M_{p,n}$ by $A^t = (b_{ij})$ where $b_{ij} = a_{ji}$. We denote $A^* = (\overline{a_{ji}})$.

Definition 2.1. If V, W are vector spaces, then we define $V \oplus W = \{(v, w) : v \in V, w \in W\}$, which is a vector space with coordinate-wise addition and scalar multiplication on each coordinate.

One can check that if $\{v_i : i \in I\}$ is a basis for V and $\{w_j : j \in J\}$ is a basis for W, then the set $\{(v_i, 0) : i \in I\} \cup \{(0, w_j) : j \in J\}$ is a basis for $V \oplus W$. In particular, we have $\dim(V \oplus W) = \dim(V) + \dim(W)$. Using this we obtain the isomorphism $\mathbb{C}^p \oplus \mathbb{C}^q \simeq \mathbb{C}^{p+q}$ via

$$((\lambda_1, ..., \lambda_p), (\mu_1, ..., \mu_q)) \mapsto (\lambda_1, ..., \lambda_p, \mu_1, ..., \mu_q).$$

We use the following identification:

$$\mathcal{L}(\mathbb{C}^{p_1} \oplus \mathbb{C}^{p_2}, \mathbb{C}^{n_1} \oplus \mathbb{C}^{n_2}) \simeq \mathcal{L}(\mathbb{C}^{p_1+p_2}, \mathbb{C}^{n_1+n_2}) \simeq M_{n_1+n_2,p_1+p_2},$$

and we express this identification with block matrices. In particular, if $A \in M_{n_1+n_2,p_1+p_2}$, then we can write $A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$ where $A_{ij} \in M_{n_i,p_j}$.

Recall that if V, W, Z are vector spaces, then $B : V \times W \to Z$ is *bilinear* if B is linear in each coordinate. The *tensor product* of V and W is defined by

$$V \otimes W = \text{span } \{v \otimes w : v \in V,\, w \in W\},$$

such that whenever $v, v_1, v_2 \in V$, $w, w_1, w_2 \in W$ and $\lambda \in \mathbb{C}$, we have $(v_1 + v_2) \otimes w = v_1 \otimes w + v_2 \otimes w$, $(\lambda v) \otimes w = \lambda(v \otimes w) = v \otimes (\lambda w)$, and $v \otimes (w_1 + w_2) = v \otimes w_1 + v \otimes w_2$.

Whenever $B : V \times W \to Z$ is bilinear, there is a unique linear map $T : V \otimes W \to Z$ such that $B(v, w) = T(v \otimes w)$. Moreover, if $\{v_i : i \in I\}$ and $\{w_j : j \in J\}$ are bases for V and W respectively, then $\{v_i \otimes w_j : i \in I,\, j \in J\}$ is a basis for $V \otimes W$. Therefore, $\dim(V \otimes W) = \dim(V)\dim(W)$. In fact, if $u \in V \otimes W$, then there are unique vectors $x_j \in V$ with $u = \sum_j x_j \otimes w_j$, and similarly there are unique vectors $y_i \in W$ with $u = \sum_i v_i \otimes y_i$.

Example 2.2. We have $\mathbb{C}^n \otimes \mathbb{C}^p \simeq \mathbb{C}^{np}$ with basis $\{e_i \otimes e_j : 1 \leq i \leq n,\, 1 \leq j \leq p\}$. There are two natural orderings to this basis:

(1) $e_1 \otimes e_1, ..., e_1 \otimes e_p, e_2 \otimes e_1, ..., e_2 \otimes e_p, ..., e_n \otimes e_p$. Essentially, this ordering is like the mapping $e_i \otimes e_j \mapsto e_{p(i-1)+j}$ (so a doubly indexed set becomes singly indexed). Looking at tuples, this is the mapping

$$(\lambda_1, ..., \lambda_n) \otimes (\mu_1, ..., \mu_p) \mapsto (\lambda_1\mu_1, ..., \lambda_1\mu_p, ..., \lambda_n\mu_1, ..., \lambda_n\mu_p).$$

In this way we have $\mathbb{C}^n \otimes \mathbb{C}^p \simeq \mathbb{C}^{np}$.

(2) $e_1 \otimes e_1, e_2 \otimes e_1, ..., e_n \otimes e_1, ..., e_1 \otimes e_p, ..., e_n \otimes e_p$. This is the mapping $e_i \otimes e_j \mapsto e_{i+n(j-1)}$, and looking at tuples we have the mapping

$$(\lambda_1, ..., \lambda_n) \otimes (\mu_1, ..., \mu_p) \mapsto (\lambda_1\mu_1, ..., \lambda_n\mu_1, ..., \lambda_1\mu_p, ..., \lambda_n\mu_p),$$

which gives the identification $\mathbb{C}^n \otimes \mathbb{C}^p \simeq \mathbb{C}^{np}$.

The maps above give an interesting permutation via

$$\mathbb{C}^{np} \underbrace{\simeq}_{(1)} \mathbb{C}^n \otimes \mathbb{C}^p \underbrace{\simeq}_{(2)} \mathbb{C}^{np},$$

and we call this the *canonical shuffle*. Another way to see this is the following. Suppose that $\dim(V) = n$ and $\dim(W) = p$, while $u \in V \otimes W$ with $u = \sum x_j \otimes w_j$, where the sets $\{v_i : i \in I\}$ and $\{w_i : i \in I\}$ are bases for V and W respectively. Then we can write $u = (x_1, ..., x_p) \in \underbrace{V \oplus \cdots \oplus V}_{p \text{ times}}$, so that $V \otimes W \simeq \underbrace{V \oplus \cdots \oplus V}_{p \text{ times}}$. We can also write $u = \sum_{i=1}^n v_i \otimes y_i$ so that $V \otimes W \simeq \underbrace{W \oplus \cdots \oplus W}_{n \text{ times}}$.

Definition 2.3. Let V be a vector space. An *inner product* is a map $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{C}$ that satisfies, for all $v, v_1, v_2 \in V$, $w, w_1, w_2 \in W$ and $\lambda \in \mathbb{C}$,

(1) $\langle v_1 + v_2, w \rangle = \langle v_1, w \rangle + \langle v_2, w \rangle$ and $\langle v, w_1 + w_2 \rangle = \langle v, w_1 \rangle + \langle v, w_2 \rangle$.
(2) $\langle \lambda v, w \rangle = \overline{\lambda} \langle v, w \rangle$, and $\langle v, \lambda w \rangle = \lambda \langle v, w \rangle$. ((1) and (2) is sesquilinearity.) (Note: conjugate linearity is in the *left* variable for this course)
(3) (Positive definite): $\langle v, v \rangle \geq 0$ with equality if and only if $v = 0$.

By Cauchy-Schwarz, we have $|\langle v, w \rangle| \leq \langle v, v \rangle^{\frac{1}{2}} \langle w, w \rangle^{\frac{1}{2}}$ so defining $\|v\| = \langle v, v \rangle^{\frac{1}{2}}$ gives a norm on V. We say that V is a *Hilbert space* if it is eqiupped with an inner product whose corresponding norm is complete. That is, whenever $(v_n)_{n=1}^{\infty}$ is a Cauchy sequence in V, then $(v_n)_{n=1}^{\infty}$ converges in V.

Example 2.4. $\mathbb{C}^n$ is a Hilbert space with inner product given by

$$\langle (\lambda_1, ..., \lambda_n), (\mu_1, ..., \mu_n) \rangle = \sum_{i=1}^n \overline{\lambda}_i \mu_i = v^* w,$$

where $v = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}$ and $w = \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_n \end{pmatrix}$. (Note we also write $|w\rangle$ for w as a column vector, and we can write $\langle v| = v^*$, so that $\langle v|w \rangle = v^* w$.) We also have $|w\rangle\langle v|$ as the rank one matrix wv^*.

Example 2.5. If V, W are Hilbert spaces, then so is $V \oplus W$ with inner product

$$\langle (v_1, w_1), (v_2, w_2) \rangle = \langle v_1, v_2 \rangle_V + \langle w_1, w_2 \rangle_W.$$

Similarly, $V \otimes W$ is a Hilbert space with inner product

$$\langle v_1 \otimes w_1, v_2 \otimes w_2 \rangle = \langle v_1, v_2 \rangle_V \cdot \langle w_1, w_2 \rangle_W.$$

(Note in infinite dimensions you have to take the completion to make $V \otimes W$ into a Hilbert space, but in finite dimensions it doesn't matter.)

3. Basics of Quantum Viewpoint

3.1. **Postulates of Quantum Mechanics.**

Postulate 1. To each isolated physical system, there corresponds a Hilbert space $\mathcal{H}$, called the *state space*, and each unit vector in $\mathcal{H}$ represents a possible state, called the *state vector* or *pure state*.

Quantum Measurements. When we want to observe a system, i.e., connect to the "outside world", the system is no longer closed because we interact with it. By *closed*, we mean "not interacting with anything outside the system". By *open*, we mean it is a piece of a larger system.

Postulate 2. Quantum measurements are always described by a class of operators $\{M_i\}_{i=\text{one of the outcomes}}$.
The probability that we observe the outcome i, given that the system is in state $|\psi\rangle$ before we measure, is given by $p_i = \|M_i\psi\|^2$ and if we observe the outcome i, then the system changes to the state $\frac{M_i\psi}{\|M_i\psi\|}$. Moreover, as the sum of the probabilities of all possible outcomes must equal 1, we have $\sum_i p_i = 1$.

Observation I Keeping in mind that quantum mechanics is inherently probabilistic, we consider a quantum experiment with at most k possible outcomes. Let $\mathcal{H}_s$ and $\mathcal{H}_o$ be Hilbert spaces representing the state space and the outcome space, respectively, and let $\{M_i \in \mathcal{B}(\mathcal{H}_s, \mathcal{H}_o) : 1 \leq i \leq k\}$ be a collection of bounded operators. If the system is in state $\psi \in \mathcal{H}_s, \|\psi\| = 1$ before we measure, then we have

$$1 = \sum_{i=1}^{k} p_i = \sum_{i=1}^{k} \|M_i\psi\|^2 = \sum_{i=1}^{k} \langle M_i\psi, M_i\psi \rangle = \sum_{i=1}^{k} \langle \psi, M_i^* M_i \psi \rangle .$$

Since the above equality holds for every $\psi \in \mathcal{H}$ with $\|\psi\| = 1$, the following lemma forces $\sum_{i=1}^{k} M_i^* M_i = I$.

Lemma 3.1. *If $T \in \mathcal{B}(\mathcal{H})$, then $T = I \iff \langle \psi, T\psi \rangle = 1$ for every $\|\psi\| = 1$.*

Proof. **Homework problem 1**; due 14th January, Thursday. □

Observation II Given any class of operators $\{M_i \in \mathcal{B}(\mathcal{H}_s, \mathcal{H}_o) : 1 \leq i \leq k\}$ such that $\sum_{i=1}^{k} M_i^* M_i = I$, there exists a k-outcome quantum experiment with these measurement operators.

3.2. **Measurement Systems and Distinguishable States.**

Definition 3.2. (Measurement System) Suppose that $\mathcal{H}$ and $\mathcal{K}$ are finite-dimensional Hilbert spaces. A finite family $\{M_i : 1 \leq i \leq k\}$ of operators $M_i : \mathcal{H} \to \mathcal{K}$ is called a *measurement system* if $\sum_i M_i^* M_i = I$. If $\mathcal{H} = \mathcal{K}$, we say that $\{M_i\}$ is a measurement system *on* $\mathcal{H}$.

Definition 3.3. (Perfectly Distinguishable States) A collection of states $\{\psi_1, \ldots, \psi_N\} \subseteq \mathcal{H}$ is called *perfectly distinguishable* if there exists a measurement system $\{M_i : 1 \leq i \leq k\}, k \geq N$ on $\mathcal{H}$ such that $\|M_i(\psi_j)\|^2 = \delta_{i,j}$ for $i, j \in \{1, ..., N\}$.

Theorem 3.4. *A collection of states $\{\psi_1, \ldots, \psi_N\} \subseteq \mathcal{H}$ is perfectly distinguishable if and only if $\psi_i \perp \psi_j$ for all $i \neq j$.*

Proof. ($\implies$) For the forward direction, let us assume that there is a measurement system $\{M_i : 1 \leq i \leq N\}$ such that $\|M_i(\psi_j)\| = \delta_{i,j}$ for $i, j \in \{1, ..., N\}$. Consider ψ_1 and ψ_2. ψ_2 can then be expressed as $\psi_2 = \alpha\psi_1 + \beta\eta$ where $\eta \perp \psi_1, \|\eta\| = 1$. Since $1 = \|\psi\|^2 = |\alpha|^2 + |\beta|^2$, we have $1 = \|M_2(\psi_2)\|^2 = \|M_2(\alpha\psi_1 + \beta\eta)\|^2 = |\beta|^2\|M_2(\eta)\|^2 \leq |\beta|^2\|\eta\|^2 = \|\beta\|^2 \leq 1$. This forces the above inequalities to be equalities so that $|\beta|^2 = 1$ which in turn implies that $\alpha = 0$ which means that ψ_2 and η are collinear and hence $\psi_2 \perp \psi_1$.

($\impliedby$) Let M_i be the (orthogonal) projection onto the one-dimensional subspace spanned by ψ_i. Then $M_i = M_i^* = M_i^* M_i$ for $i = 1, ..., N$ and $\sum_{i=1}^{N} M_i^* M_i$ is the orthogonal projection onto $\text{span}\{\psi_1, ..., \psi_N\}$. Let M_0 be the orthogonal projection onto $\{\psi_1, \ldots, \psi_N\}^\perp$. Then $\sum_{j=0}^{N} M_j^* M_j = \sum_{j=0}^{N} M_j = I$. Furthermore, $M_i(\psi_j) = \delta_{i,j}\psi_j$ for all $i, j \in \{1, ..., N\}$, so that $\|M_i(\phi_j)\|^2 = \delta_{i,j}$ for all $i, j \in \{1, ..., N\}$. This proves that $\{M_i\}_{i=0}^{N}$ is a measurement system. □

Corollary 3.5. *If $\dim(\mathcal{H}_s) = N$, then the system can have at most N perfectly distinguishable states.*

Theorem 3.6. *Suppose that $\{\psi_1, \ldots, \psi_N\}$ is a collection of linearly independent states. Then there exists a measurement system $\{M_i : 0 \leq i \leq N\}$ such that for $i \neq 0$, $\|M_i(\psi_j)\| \neq 0$ if and only if $i = j$.*

Proof. For $i = 1, ..., N$, let $V_i = \text{span}\{\psi_j : j \neq i\}$, and let E_i be the projection onto $V_i^\perp$. Then for $j \neq i$, $\psi_j \in V_i \implies E_i(\psi_j) = 0 \implies \|E_i(\psi_j)\|^2 = 0$. Now $0 \leq E_i \leq I \implies 0 \leq E_1 + \cdots + E_N \leq N \cdot I$. Let

$M_i = \frac{1}{\sqrt{N}} E_i$ for $i = 1, ..., N$. Then $M_i^* M_i = \frac{1}{N} E_i$, so $\sum_{i=1}^N M_i^* M_i = \frac{1}{N} \sum_{i=1}^N E_i \leq I$, and hence $I - \sum_{i=1}^N M_i^* M_i \geq 0$. Now let $M_0 = (I - \sum_{i=1}^N M_i^* M_i)^{\frac{1}{2}}$. Then $\sum_{i=0}^N M_i^* M_i = \left((I - \sum_{i=1}^N M_i^* M_i)^{\frac{1}{2}}\right)^2 + \sum_{i=1}^N M_i^* M_i = I$, so $\{M_i\}_{i=0}^N$ is a measurement system. For $i \neq 0$, if $j \neq i$, then $\|M_i(\psi_j)\| = \frac{1}{\sqrt{N}} \|E_i(\psi_j)\| = 0$. Therefore by contrapositive, $\|M_i(\psi_j)\| \neq 0$ implies that $i = j$. Conversely, $\|M_i(\psi_i)\| = \frac{1}{\sqrt{N}} \|E_i(\psi_i)\| \neq 0$ since $\psi_i \notin V_i$ and so it has non-zero projection onto $V_i^\perp$. □

So far we have talked about pure states, now we will talk about ensembles (or mixed states).

3.3. **Ensembles or Mixed States.** As motivation for this topic, let $\{M_i : 1 \leq i \leq k\}$ be a measurement system with $M_i : \mathcal{H}_s \longrightarrow \mathcal{H}_o$. Suppose we have the state $\psi \in \mathcal{H}_s$ as input. Recall that $p_i = \|M_i(\psi)\|^2$ should be interpreted as the probability of observing the outcome i, and that if we do observe i, the system is now in the state, $\frac{M_i(\psi)}{\|M_i(\psi)\|}$. That is,
input:$\psi \in \mathcal{H}_s$; output: $\frac{M_i(\psi)}{\|M_i(\psi)\|}$ with probability $p_i = \|M_i(\psi)\|^2$.
So after observation, we will have what now looks like a mixed bag of states $\left\{\frac{M_i(\psi)}{\|M_i(\psi)\|}\right\}_i$, with $\frac{M_i(\psi)}{\|M_i(\psi)\|}$ occurring with probability p_i.

Definition 3.7. An *ensemble of states*, or a *mixed state*, is a finite collection $\{\psi_i, p_i : 1 \leq i \leq N\}$ of states ψ_i with probabilities p_i where $\|\psi_i\| = 1$, $p_i \geq 0$ and $\sum_{i=1}^N p_i = 1$.

Suppose we have a measurement system $\{M_i : 1 \leq i \leq N\}$ and an ensemble of states $\{\psi_j, p_j : 1 \leq j \leq k\}$ with $\sum_{j=1}^k p_j = 1$, then what is the probability of observing the outcome i?
If ψ_j is our input, then the probability getting outcome i is $\|M_i(\psi_j)\|^2$. So, the probability that we have input ψ_j and outcome i is,

$$\sum_{j=1}^k p_j \|M_i(\psi_j)\|^2.$$

In the next subsection we discuss a better way to compute the probabilities of outcomes.

3.4. **Von Neumann's Notation: Density Matrices.** For a given state $\psi \in \mathcal{H}_s, \|\psi\| = 1$, a typical unit vector in the one-dimensional subspace spanned by ψ is given by $e^{i\theta}\psi$. In general $e^{i\theta}\psi \neq \psi$ but for any measurement M_j, we can see that $\|M_j(\psi)\|^2 = \|M_j(e^{i\theta}\psi)\|^2$. This shows that measurements don't distinguish between different unit vectors from the one-dimensional subspace spanned by the given state vector ψ and hence states should really refer to one-dimensional subspace and not just a unit vector. This means that *the probabilities of outcomes really depend on the one-dimensional subspace generated by a vector.*

Replacing states by rank one projections and lengths by trace: Recall that given a matrix $A = (a_{ij}) \in M_n$, the *trace* of that matrix is the sum of the diagonal entries: $Tr(Y) = \sum_i a_{ii}$. It is a popular fact that given any two square matrices A and B of the same size, $Tr(AB) = Tr(BA)$. The next proposition establishes this fact for compatible non-square matrices as well.

Proposition 3.8. *If $A \in M_{n,p}$ and $B \in M_{p,n}$, so that $AB \in M_n$ and $BA \in M_p$, then $Tr(AB) = Tr(BA)$.*

Proof. **Homework problem 2**; due 14th January, Thursday. □

Next, if $\psi \in \mathbb{C}^n, \|\psi\| = 1$, and P_ψ denotes the orthogonal projection onto the subspace spanned by ψ, then $P_\psi = \psi\psi^* = |\psi\rangle\langle\psi|$. ($P_\psi h = \psi\psi^* h = \langle\psi|h\rangle\psi$ where $\langle\psi|h\rangle$ is the component of h in the direction of ψ.) Furthermore,

$$Tr(P_\psi) = Tr(\psi\psi^*) = Tr(\psi^*\psi) = (\psi^*\psi) = \langle\psi|\psi\rangle = 1.$$

Back to Ensemble: Let's get back to the situation where we had a measurement system $\{M_i : 1 \leq i \leq N\}$ and an ensemble of states $\{\psi_j, p_j : 1 \leq j \leq k\}$ with $\sum_{j=1}^k p_j = 1$. We know that the probability of observing the outcome i is,

$$\sum_{j=1}^{k} p_j \|M_i(\psi_j)\|^2.$$

Simplifying this expression, we get

$$\begin{aligned}
\sum_{j=1}^{k} p_j \|M_i(\psi_j)\|^2 &= \sum_{j=1}^{k} p_j (M_i\psi_j)^*(M_i\psi_j) \\
&= \sum_{j=1}^{k} p_j Tr((M_i\psi_j)^*(M_i\psi_j)) \\
&= \sum_{j=1}^{k} p_j Tr((M_i\psi_j)(M_i\psi_j)^*) \\
&= \sum_{j=1}^{k} p_j Tr(M_i\psi_j\psi_j^* M_i^*) \\
&= \sum_{j=1}^{k} p_j Tr(M_i^* M_i\psi_j\psi_j^*) \\
&= \sum_{j=1}^{k} Tr\left(M_i^* M_i(p_j\psi_j\psi_j^*)\right) \\
&= Tr\left(M_i^* M_i\left(\sum_{j=1}^{k} p_j\psi_j\psi_j^*\right)\right).
\end{aligned}$$

Note that $\psi_j\psi_j^* = P_{\psi_j}$. If we set $P = \sum_{j=1}^{k} p_j\psi_j\psi_j^*$, then we have shown that:

Theorem 3.9. *Given an ensemble of states $\{\psi_j, p_j : 1 \leq j \leq k\}$ and a measurement system $\{M_i : 1 \leq i \leq N\}$, the probability of observing the i-th outcome is $Tr\,(M_i^* M_i P)$ where $P = \sum_{j=1}^{k} p_j\psi_j\psi_j^*$.*

The square matrix $P = \sum_{j=1}^{k} p_j\psi_j\psi_j^*$ associated to an ensemble of states is of great importance and it deserves a name of its own! Von Neumann called it the density matrix of the given ensemble.

Definition 3.10. (Density Matrix Of an Ensemble) Given an ensemble of states $\{\psi_j, p_j : 1 \leq j \leq k\}$, the matrix

$$P = \sum_{j=1}^{k} p_j P_{\psi_j}$$

is called the *density matrix* of the ensemble.

We observe that:

(1) If two ensembles have the same density matrix, then we get the same probability for outcomes for any measurement system. In other words, a measurement system cannot differentiate between two ensembles with same density matrix.

(2) If $\{M_i : 1 \leq i \leq k\}$ and $\{\tilde{M}_i : 1 \leq i \leq k\}$ are two measurement systems such that for every i, $M_i^* M_i = \tilde{M}_i^* \tilde{M}_i$, then also we get the same probability for outcomes for any ensemble. This means that an ensemble cannot practically distinguish one measurement system from the other in such situation.

The following example illustrates the first observation.

Example 3.11. If $\{u_1, ..., u_N\}$ is an orthonormal basis for $\mathbb{C}^N$, then the density matrix P for the ensemble $\{u_j, \frac{1}{N} : 1 \leq j \leq N\}$ is given by $P = \sum_{j=1}^N \frac{1}{N} u_j u_j^* = \frac{1}{N} I_N$. If $\{\tilde{u}_1, ..., \tilde{u}_N\}$ is another orthonormal basis for $\mathbb{C}^N$, then the density matrix $\tilde{P}$ for the ensemble $\{\tilde{u}_j, \frac{1}{N} : 1 \leq j \leq N\}$ also turns out to be $\tilde{P} = \sum_{j=1}^N \frac{1}{N} \tilde{u}_j \tilde{u}_j^* = \frac{1}{N} I_N$. This example guarantees the existence of two different ensembles with same density matrix.

Problem 3.12. *Fix $N \geq 3$ and let $u_j = \begin{pmatrix} \cos(\frac{2\pi j}{N}) \\ \sin(\frac{2\pi j}{N}) \end{pmatrix} \in \mathbb{C}^2$. Prove that the density matrix for the ensemble $\{u_j, \frac{1}{N} : 1 \leq j \leq N\}$ is given by $\frac{1}{2} I_2$.*

Proof. **Homework problem 3**; due 19th January, Tuesday. □

The above problem the density matrix of does not distinguish between standard orthonormal basis or any other orthonormal basis as input. So, for all intents and purposes, it is the density matrix which is important and not the ensembles.

At this point, let us pause for a while and try to visualise quantum experiments in terms of density matrices. Recall that, if a system is initially in the state ψ, that is, $\psi \in \mathcal{H}_s, \|\psi\| = 1$, and if there is given a measurement system $\{M_i : 1 \leq i \leq k\}$, then after measurement, the system becomes the ensemble $\left\{ \frac{M_i(\psi)}{\|M_i(\psi)\|}, \|M_i\psi\|^2 : 1 \leq i \leq k \right\}$. By associating density matrices with the states of the system before and after the measurement we note that the input is the state ψ and the density matrix corresponding to it is given by $P = \psi\psi^*$. After the measurement, the system becomes the ensemble $\left\{ \frac{M_i(\psi)}{\|M_i(\psi)\|}, \|M_i\psi\|^2 : 1 \leq i \leq k \right\}$, and hence the output is this ensemble which is identified by the density

matrix

$$\sum_{i=1}^{k} \|M_i\psi\|^2 \left(\frac{M_i(\psi)}{\|M_i(\psi)\|}\right)\left(\frac{M_i(\psi)}{\|M_i(\psi)\|}\right)^*$$
$$= \sum_{i=1}^{k} (M_i\psi)(M_i\psi)^* = \sum_{i=1}^{k} (M_i\psi)(\psi^* M_i^*)$$
$$= \sum_{i=1}^{k} M_i(\psi\psi^*)M_i^* = \sum_{i=1}^{k} M_i P M_i^*.$$

Thus, in terms of density matrices, we observed that if input is identified by the density matrix P, then after measurement, the output is identified by the density matrix $\sum_{i=1}^{k} M_i P M_i^*$. This observation is the key to our next theorem.

Theorem 3.13. *Given an ensemble of states $\{\psi_j, p_j : 1 \le j \le J\}$ and a measurement system $\{M_i : 1 \le i \le k\}$ on $\mathcal{H}_s$ with density matrix $P = \sum_{j=1}^{J} p_j \psi_j \psi_j^*$, then after measurement, the system becomes the ensemble $\left\{\frac{M_i(\psi_j)}{\|M_i(\psi_j)\|}, p_j\|M_i\psi_j\|^2 : 1 \le i \le k, 1 \le j \le J\right\}$ with density matrix $\sum_{i=1}^{k} M_i P M_i^*$.*

Proof. The density matrix for the output ensemble is given by

$$\sum_{j=1}^{J}\sum_{i=1}^{k} p_j ||M_i\psi_j||^2 \left(\frac{M_i\psi_j}{||M_i\psi_j||}\right)\left(\frac{M_i\psi_j}{||M_i\psi_j||}\right)^* = \sum_{j=1}^{J}\sum_{i=1}^{k} p_j (M_i\psi_j)(M_i\psi_j)^*$$
$$= \sum_{i=1}^{k}\sum_{j=1}^{J} M_i(\psi_j p_j \psi_j^*)M_i^*$$
$$= \sum_{n=1}^{N} M_i P M_i^*. \qquad \square$$

So, a measurement system takes density matrix and yields another density matrix. A natural question to ask at this stage is "what kind of matrices are density matrices?" To answer this question, let us recall the definition of positive semidefinite matrix.

Definition 3.14. (Positive Semidefinite Matrix) $P \in M_n$ is said to be a positive semidefinite (≥ 0) matrix if $\langle h, Ph\rangle \ge 0$ for every $h \in \mathbb{C}$.

Fact: $P \in M_n \ge 0 \iff P$ has an orthonormal basis consisting entirely of non-negative eigenvalues.

Lemma 3.15. *If $\{u_1, ..., u_n\}$ is an orthonormal basis consisting entirely of non-negative eigenvalues of a positive semidefinite matrix P such that $Pu_j = \lambda_j u_j$ for every j, then $P = \sum_{j=1}^{n} \lambda_j u_j u_j^*$.*

Proof. For each $h \in \mathcal{H}$, write $h = \sum_{j=1}^{n} \langle u_j, h\rangle u_j = \sum_{j=1}^{n} u_j^* h u_j = \sum_{j=1}^{n} u_j u_j^* h$. Then we have,

$$Ph = \sum_{j=1}^{n} P(\langle u_j, h\rangle u_j) = \sum_{j=1}^{n} \lambda_j \langle u_j, h\rangle u_j = \left(\sum_{j=1}^{n} \lambda_j u_j u_j^*\right) h,$$

which implies that $P = \sum_{j=1}^{n} \lambda_j u_j u_j^*$. □

The following proposition answers the question addressed earlier: "what kind of matrices are density matrices?"

Proposition 3.16. *Let $P \in M_n$. Then P is a density matrix of some ensemble if and only if $P \geq 0$ and $Tr(P) = 1$.*

Proof. Suppose P is a density matrix of some ensemble $\{\psi_j, p_j\}_{j=1}^{J}$. Then by definition, $P = \sum_{j=1}^{J} p_j \psi_j \psi_j^*$. Then we have,

$$\begin{aligned}\langle h, Ph\rangle &= \sum_{j=1}^{J} p_j \langle h, \psi\psi^* h\rangle = \sum_{j=1}^{J} p_j \langle h, \psi \langle \psi, h\rangle\rangle = \sum_{j=1}^{J} p_j \langle \psi, h\rangle \langle h, \psi\rangle \\ &= \sum_{j=1}^{J} p_j \langle \psi, h\rangle \overline{\langle \psi, h\rangle} = \sum_{j=1}^{J} p_j |\langle \psi, h\rangle|^2 \geq 0 \text{ for every } h \in \mathbb{C}^n.\end{aligned}$$

This proves that $P \geq 0$. (One can easily prove the positivity of P by observing that P is a sum of positive rank-one operators.)
Moreover,

$$\begin{aligned}\mathrm{Tr}(P) &= \sum_{j=1}^{J} p_j \mathrm{Tr}(\psi_j \psi_j^*) = \sum_{j=1}^{J} p_j \mathrm{Tr}(\psi_j^* \psi_j) \\ &= \sum_{j=1}^{J} p_j \langle \psi_j, \psi_j\rangle = \sum_{j=1}^{J} p_j = 1.\end{aligned}$$

Note that the second line follows from the fact that $\psi_j^* \psi_j$ is just the inner product $\langle \psi_j, \psi_j\rangle$. Conversely, let $P \geq 0$ with $\mathrm{Tr}(P) = 1$. Then there exists eigenvectors $\{u_j\}_{j=1}^{J}$ that is an orthonormal basis for $\mathbb{C}^J$ with corresponding eigenvalues $\lambda_j \geq 0$ so that $P = \sum_{j=1}^{J} \lambda_j u_j u i_j^*$. Then $\mathrm{Tr}(P) = \sum_{j=1}^{J} \lambda_j = 1$ suggests that λ_i's are probabilities. Consequently, $\{u_j, \lambda_j\}_{j=1}^{J}$ forms an emsemble with density matrix P. □

So, in the terminology of operator theory, density matrices are precisely positive operators of trace one. This observation allows us to analyze ensembles and quantum events from a completely mathematical approach. It is, hence, easy to see that a measurement system $\{M_i\}_i$, $M_i : \mathcal{H}_s \longrightarrow \mathcal{H}_o$ can be identified by a linear map $\Phi : \mathcal{L}(\mathcal{H}_s) \longrightarrow \mathcal{L}(\mathcal{H}_o)$, given by $\Phi(P) = \sum_i M_i P M_i^*$ which sends density matrices to density matrices. This leads to the next postulate of Quantum Mechanics.

Postulate 3. If a quantum event occurs transforming pure states in lab $\mathcal{H}_1$ to lab $\mathcal{H}_2$, then there exists a linear map $\Phi : \mathcal{L}(\mathcal{H}_1) \longrightarrow \mathcal{L}(\mathcal{H}_2)$ such that Φ maps density matrices to density matrices.

The following proposition is an immediate consequence of the above postulate.

Proposition 3.17. *If $\Phi\colon \mathcal{L}(\mathcal{H}_1) \longrightarrow \mathcal{L}(\mathcal{H}_2)$ is a linear map that sends densities to densities (that is, describes a quantum event), then*

1. *For every $P \in \mathcal{L}(\mathcal{H}_1), P \geq 0$, we have $\Phi(P) \geq 0$. (So, Φ is a positive linear map.)*
2. *For every $X \in \mathcal{L}(\mathcal{H}_1)$, we have $Tr(\Phi(X)) = Tr(X)$. (So, Φ is a trace-preserving map.)*

Such a map is called positive, trace-preserving map.

Proof. (1) If Φ describes a quantum event, then it maps density matrices to density matrices. We first prove positivity. If $P \geq 0$ and $\mathrm{Tr}(P) = 0$, then $P = 0$ and $\Phi(P) = 0 \geq 0$. Thus without loss of generality, assume $P \geq 0$ and $\mathrm{Tr}(P) > 0$. Consider the matrix $\frac{1}{Tr(P)}P$. It is positive and of trace one, hence a density matrix. By hypothesis, $\Phi(\frac{1}{Tr(P)}P) := Q$ is a density matrix so $\Phi(P) = Tr(P) \cdot Q \geq 0$.

(2) For the trace-preserving property, we first suppose that $P \geq 0$. Note that $\Phi(\frac{1}{\mathrm{Tr}(P)}P) := Q$ is a density matrix, so we have,

$$1 = \mathrm{Tr}(Q) = \mathrm{Tr}\left(\Phi(\frac{1}{Tr(P)}P)\right) = \frac{\mathrm{Tr}(\Phi(P))}{\mathrm{Tr}(P)},$$

which implies that $\mathrm{Tr}(\Phi(P)) = \mathrm{Tr}(P)$ and hence trace is preserved for positive semidefinite matrices. (Note that if $\mathrm{Tr}(P) = 0$ then $P = 0$ so $\mathrm{Tr}(\Phi(P)) = \mathrm{Tr}(0) = 0$.)
It is well known fact that if $H = H^*$, then $H = P_1 - P_2$ where $P_1, P_2 \geq 0$. Then

$$\mathrm{Tr}(\Phi(H)) = \mathrm{Tr}(\Phi(P_1) - \Phi(P_2)) = \mathrm{Tr}(P_1) - \mathrm{Tr}(P_2) = \mathrm{Tr}(H),$$

so trace is preserved on self-adjoints.

Finally, if $X \in \mathcal{L}(\mathcal{H}_1)$, consider the Cartesian decomposition $X = H + iK$, where $H = \frac{X+X^*}{2}$, $K = \frac{X-X^*}{2}$ so that $H = H^*$ and $K = K^*$. Since each of H and K can be written a difference of two positive operators, by linearity of trace we deduce that for any $X \in \mathcal{L}(\mathcal{H}_1)$, we have $\mathrm{Tr}(\Phi(X)) = \mathrm{Tr}(\Phi(H) + i\Phi(K)) = \mathrm{Tr}(H) + i\mathrm{Tr}(K) = \mathrm{Tr}(H + iK) = \mathrm{Tr}(X)$.

□

3.5. **Tensor Products of Matrices (Kronecker Product).** Suppose that $T : V_1 \longrightarrow W_1$ and $R : V_2 \longrightarrow W_2$ are linear maps between vector spaces, then there is a linear map $T \otimes R : V_1 \otimes V_2 \longrightarrow W_1 \otimes W_2$ defined by $(T \otimes R)(v_1 \otimes v_2) = T(v_1) \otimes R(v_2)$ for all $v_1 \in V_1$ and $v_2 \in V_2$. If $\mathcal{H}$ and $\mathcal{K}$ are finite-dimensional Hilbert spaces with $X : \mathcal{H} \to \mathcal{H}$ and $Y : \mathcal{K} \to \mathcal{K}$, linear. Then there is a well-defined linear map denoted $X \otimes Y : \mathcal{H} \otimes \mathcal{K} \to \mathcal{H} \otimes \mathcal{K}$ satisfying $(X \otimes Y)(h \otimes k) = X(h) \otimes Y(k)$.

Comfortingly, the image of the tensor product is the tensor product of the images; the proof is a straight-forward application of the definitions involved and its tedium is left to the unbeliever.

In particular, if $T : \mathcal{H} \longrightarrow \mathcal{H}$ and $R : \mathbb{C}^n \longrightarrow \mathbb{C}^n$ are linear, then we can define $T \otimes R : \mathcal{H} \otimes \mathbb{C}^n \longrightarrow \mathcal{H} \otimes \mathbb{C}^n$, in a similar way. Our goal in this subsection is to find a matrix representation for the map $T \otimes R$ in this setting. To do this, let us first address the following question:

I. What is a natural identification of a typical element of $\mathcal{H} \otimes \mathbb{C}^n$?

Recall that if we take the canonical orthonormal basis $\{e_1, ..., e_n\}$ for $\mathbb{C}^n$, then every vector $u \in \mathcal{H} \otimes \mathbb{C}^n$ has a unique representation given by $u = \sum_{i=1}^n h_i \otimes e_i$ where $h_i \in \mathcal{H}$, and

$$\|u\|^2 = \left\langle \sum_{i=1}^n h_i \otimes e_i, \sum_{j=1}^n h_j \otimes e_j \right\rangle = \sum_{i,j=1}^n \langle h_i, h_j \rangle \langle e_i, e_j \rangle = \sum_{i=1}^n \|h_i\|^2 = \|(h_1, ..., h_n)\|^2.$$

In other words, we have the Hilbert space isomorphism

$$\mathcal{H} \otimes \mathbb{C}^n \simeq \underbrace{\mathcal{H} \oplus \cdots \oplus \mathcal{H}}_{n \text{ times}} (= \bigoplus_1^n \mathcal{H}),$$

via the natural identification $\sum_{i=1}^n (h_i \otimes e_i) \simeq \begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix}$.

The next question which we want to address is:

II. What is a natural identification of a linear map in $\mathcal{L}(\mathcal{H} \otimes \mathbb{C}^n)$?
Given $A_{ij} \in \mathcal{L}(\mathcal{H})$ for $1 \leq i, j \leq n$, we can consider $A = (A_{ij}) \in M_n(\mathcal{L}(\mathcal{H}))$ as an operator defined by

$$A \begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix} = \begin{pmatrix} \sum_{j=1}^n A_{1j} h_j \\ \vdots \\ \sum_{j=1}^n A_{nj} h_j \end{pmatrix} \in \underbrace{\mathcal{H} \oplus \cdots \oplus \mathcal{H}}_{n \text{ times}}.$$

Therefore, we have $M_n(\mathcal{L}(\mathcal{H})) \hookrightarrow \mathcal{L}(\mathcal{H} \otimes \mathbb{C}^n)$ in a natural way. In fact, every linear map on $\mathcal{H} \otimes \mathbb{C}^n$ has such a matrix representation. The proof is "grubby" but here is the idea: If $A : \mathcal{H} \oplus \cdots \oplus \mathcal{H} \to \mathcal{H} \oplus \cdots \oplus \mathcal{H}$ is linear, then $A \begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix} = \begin{pmatrix} k_1 \\ \vdots \\ k_n \end{pmatrix}$. The map $\begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix} \mapsto k_1$ is linear. Similarly, mapping the column vector to k_2 is linear, and so on and so forth. The map $\begin{pmatrix} h_1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \mapsto k_1$ is linear, so there is $A_{11} : \mathcal{H} \to \mathcal{H}$ enacting this transformation. If we continue to do this for every h_i and k_j, then we get linear maps $A_{ij} : \mathcal{H} \to \mathcal{H}$ and one can check that $A = (A_{ij})$. Hence, we have a natural identification $\mathcal{L}(\mathcal{H} \otimes \mathbb{C}^n) \simeq M_n(\mathcal{L}(\mathcal{H}))$ via $A \simeq (A_{ij})$, thereby allowing us to identify any linear operator $A \in \mathcal{L}(\mathcal{H} \otimes \mathbb{C}^n)$ by an $n \times n$ block matrix $(A_{ij}) \in M_n(\mathcal{L}(\mathcal{H}))$ whose entries are given by linear maps.

III. Matrix Representation of $T \otimes R$(Back to the beginning): As in the beginning of this subsection, suppose that $T : \mathcal{H} \longrightarrow \mathcal{H}$ and $R : \mathbb{C}^n \longrightarrow \mathbb{C}^n$ are linear, $R \in M_n(\mathbb{C}), R = (r_{ij})$, then $T \otimes R : \mathcal{H} \otimes \mathbb{C}^n \longrightarrow \mathcal{H} \otimes \mathbb{C}^n$ has a natural representation as an $n \times n$ block matrix $T \otimes R \in M_n(\mathcal{L}(\mathcal{H}))$ whose entries are given by linear maps.
We know that $(T \otimes R)(h \otimes y) = T(h) \otimes R(y)$, therefore,

$$(T \otimes R)(h \otimes e_j) = T(h) \otimes R(e_j) = T(h) \otimes \left(\sum_{i=1}^n r_{ij} e_i \right)$$

$$= \sum_{i=1}^n r_{ij} T(h) \otimes e_i \simeq \begin{pmatrix} r_{1j} Th \\ \vdots \\ r_{nj} Th \end{pmatrix} = (r_{ij} T) \begin{pmatrix} 0 \\ \vdots \\ h \\ \vdots \\ 0 \end{pmatrix},$$

where h is in the j-th position and there are 0's everywhere else in the column vector. The ***Kronecker product*** of T and R, then, is the block matrix in $M_n(\mathcal{L}(\mathcal{H}))$ given by $(r_{ij}T)$ (so, there are n blocks, each block is of size equal to the dimension of $\mathcal{H}$, and the (i,j)-block is $r_{ij}T$). In other words, the Kronecker product is equal to the tensor product of the linear maps (with respect to the canonical basis for $\mathbb{C}^n$).

Alternatively, if $T \in M_k$ and $R \in M_n$. Then $T \otimes R$ has matrix representation given by $T \otimes R = \begin{pmatrix} r_{11}T & \cdots & r_{1n}T \\ \vdots & \ddots & \vdots \\ r_{n1}T & \cdots & r_{nn}T \end{pmatrix}$, a block matrix with n blocks, each of size k.

Remark 3.18. We observed that if we let $X_{i,j} \in \mathcal{L}(\mathcal{H})$ and $X := (X_{i,j}) \in M_n(\mathcal{L}(\mathcal{H}))$, we can view X as a linear map, $X : \underbrace{\mathcal{H} \oplus \cdots \oplus \mathcal{H}}_{\text{n times}} \to \underbrace{\mathcal{H} \oplus \cdots \oplus \mathcal{H}}_{\text{n times}}$ and hence we can make the following identification,

$$M_n(\mathcal{L}(\mathcal{H})) \cong \mathcal{L}(\mathcal{H} \oplus \cdots \oplus \mathcal{H}) \cong \mathcal{L}(\mathcal{H} \otimes \mathbb{C}^n).$$

Conversely, notice that any $X \in M_n(\mathcal{L}(\mathcal{H}))$ has a unique representation as $X = \sum_{i,j=1}^n E_{i,j} \otimes X_{i,j}$ with $X_{i,j} \in \mathcal{L}(\mathcal{H})$, which means we can identify $\sum_{i,j=1}^n E_{i,j} \otimes X_{i,j} \cong (X_{i,j})$ and therefore, $M_n(\mathcal{L}(\mathcal{H})) \simeq \mathcal{L}(\mathcal{H}) \otimes M_n \simeq \mathcal{L}(\mathcal{H}) \otimes \mathcal{L}(\mathbb{C}^n)$. Therefore, we have the identification

$$\mathcal{L}(\mathcal{H} \otimes \mathbb{C}^n) \simeq M_n(\mathcal{L}(\mathcal{H})) \simeq \mathcal{L}(\mathcal{H}) \otimes \mathcal{L}(\mathbb{C}^n).$$

3.6. **Axiomatization of Quantum Channels.** We are now in a position to axiomatize what are known as *quantum channels*.

- **Axiom I**
  Each measurement system $\{M_i : \mathcal{H}_s \longrightarrow \mathcal{H}_o, 1 \leq i \leq K\}$ (where $\sum_{i=1}^K M_i^* M_i = I$) defines a *quantum channel* given by $\Phi : \mathcal{L}(\mathcal{H}_s) \to \mathcal{L}(\mathcal{H}_o)$ where $\Phi(W) = \sum_{i=1}^K M_i W M_i^*$.
- **Axiom II**
  If $\Phi : \mathcal{L}(\mathcal{H}_s) \to \mathcal{L}(\mathcal{H}_o)$ defines a quantum channel, then we assume that Φ is a positive and trace-preserving map.

Suppose we have two laboratories A and B (for Alice and Bob respectively). We will denote by $\mathcal{H}_{s,A}, \mathcal{H}_{s,B}, \mathcal{H}_{o,A}, \mathcal{H}_{o,B}$, respectively, the state space of lab A, the state space of lab B, the outcome space of lab A, and the outcome space of lab B. Let $\{M_i : \mathcal{H}_{s,A} \longrightarrow H_{o,A}\}_{i=1}^K$ be the measurement system of A and $\{N_j : \mathcal{H}_{s,B} \longrightarrow H_{o,B}\}_{j=1}^J$ be the measurement system of B.
Now, we wish to combinedly view these two labs as one single lab, say lab AB. The state space of this lab would be then $\mathcal{H}_{s,AB} = \mathcal{H}_{s,A} \otimes \mathcal{H}_{s,B}$,

the output space would be $\mathcal{H}_{o,AB} = \mathcal{H}_{o,A} \otimes \mathcal{H}_{o,B}$ with measurement operators $\{M_i \otimes N_j : \mathcal{H}_{s,A} \longrightarrow H_{o,A}\}$, so that there are KJ outcomes. Needless to say that $\sum_{i,j}(M_i \otimes N_j)(M_i \otimes N_j) = I$. This measurement system of lab AB, then, defines a quantum channel $\Phi_{AB} : \mathcal{L}(\mathcal{H}_{s,AB}) \longrightarrow \mathcal{L}(\mathcal{H}_{o,AB})$ given by

$$\Phi_{AB}(W) = \sum_{i,j}(M_i \otimes N_j)W(M_i \otimes N_j)^*.$$

Since, $\mathcal{L}(\mathcal{H}_{s,AB}) \simeq \mathcal{L}(\mathcal{H}_{s,A}) \otimes \mathcal{L}(\mathcal{H}_{s,B})$ and $\mathcal{L}(\mathcal{H}_{o,AB}) \simeq \mathcal{L}(\mathcal{H}_{o,A}) \otimes \mathcal{L}(\mathcal{H}_{o,B})$, it is easy to see that if $W \in \mathcal{H}_{s,AB}$, then W can be expressed as $W = W_1 \otimes W_2$, for some $W_1 \in \mathcal{H}_{s,A}, W_2 \in \mathcal{H}_{s,B}$, so that

$$\Phi_{AB}(W_1 \otimes W_2) = \sum_{i,j}(M_i W_i M_i^*) \otimes (N_j W_2 N_j^*) = \Phi_A(W_1) \otimes \Phi_B(W_2).$$

Therefore,

$$\Phi_{AB} = \Phi_A \otimes \Phi_B.$$

The above discussion serves the purpose of axiomatizing the measurement systems of two labs when they are combinedly viewed. The observation is, hence, summarized as the third axiom.

- **Axiom III**
  If $\Phi_1 : \mathcal{L}(\mathcal{H}_1) \to \mathcal{L}(\mathcal{H}_2)$ and $\Phi_2 : \mathcal{L}(\mathcal{K}_1) \to \mathcal{L}(\mathcal{K}_2)$ are both quantum channels (that is, positive and trace-preserving map), then $\Phi_1 \otimes \Phi_2 : \mathcal{L}(\mathcal{H}_1 \otimes \mathcal{K}_1) \to \mathcal{L}(\mathcal{H}_2 \otimes \mathcal{K}_2)$ should also be a quantum channel.

Remark 3.19. The third axiom of quantum channels anticipates $\Phi_1 \otimes \Phi_2$ to be a quantum channel, (that is, a positive and trace-preserving map) whenever Φ_1, Φ_2 are quantum channels. As we will see later, this condition is same as requiring Φ_1 and Φ_2 to be something called *completely positive* maps. In the next section we study extensively the theory of completely positive maps. We then would be able to see that

- Quantum channels arising from measurement systems are completely positive and trace preserving;
- If Φ_1, Φ_2 are completely positive, then so is $\Phi_1 \otimes \Phi_2$;
- If Φ is completely positive and trace-preserving, then there is a measurement system whose quantum channel is Φ.

4. Theory of CP maps

We begin this section by an example.

Example 4.1 (The trivial measurement system of dimension n)**.** Let $\mathcal{H}_{s,B} = \mathcal{H}_{o,B} = \mathbb{C}^n$ be the state space and outcome space for lab

B and $M_1 = I_n$ be its measurement system. Then the corresponding quantum channel is given by the map $\Phi_B : \mathcal{L}(\mathbb{C}^n) \to \mathcal{L}(\mathbb{C}^n)$, defined by $\Phi_B(W) = I_n^* W I_n = W$. Now if there is another lab (say, lab A) with quantum channel $\Phi_A : \mathcal{L}(\mathcal{H}_{s,A}) \to \mathcal{L}(\mathcal{H}_{o,A})$ that is positive and trace-preserving, then $\Phi_A \otimes \Phi_B : \mathcal{L}(\mathcal{H}_{s,A} \otimes \mathbb{C}^n) \to \mathcal{L}(\mathcal{H}_{o,A} \otimes \mathbb{C}^n)$ must also be positive and trace-preserving by axiom II. Note that for every $T \in \mathcal{L}(\mathcal{H}_{s,A})$ and every $R \in \mathcal{L}(\mathbb{C}^n)$, we have

$$(\Phi_A \otimes \Phi_B)(T \otimes R) = \Phi_A(T) \otimes \Phi_B(R) = \Phi_A(T) \otimes R.$$

So, if $W \in \mathcal{L}(\mathcal{H}_{s,A} \otimes \mathbb{C}^n) \simeq \mathcal{L}(\mathcal{H}_{s,A}) \otimes M_n$ and write $W = \sum_{i,j} W_{ij} E_{ij}$, then

$$(\Phi_A \otimes \Phi_B)\left(\sum_{i,j} W_{ij} \otimes E_{ij}\right) = \sum_{i,j} \Phi_A(W_{ij}) \otimes E_{ij}.$$

In terms of matrices, we can think of the quantum channel $\Phi_A \otimes \Phi_B$ as the map $W = (W_{ij}) \mapsto (\Phi_A(W_{ij}))$. We need this map to be positive and trace-preserving; as the third axiom suggests.

Definition 4.2 ($\Phi^{(n)}$)**.**
For any linear map $\Phi : \mathcal{L}(\mathcal{H}_1) \to \mathcal{L}(\mathcal{H}_2)$, we define map $\Phi^{(n)} : M_n(\mathcal{L}(\mathcal{H}_1)) \to M_n(\mathcal{L}(\mathcal{H}_2))$ by $\Phi^{(n)}((W_{i,j})) := (\Phi(W_{i,j}))$.

Definition 4.3 (n-positive; Completely Positive)**.** A linear map $\Phi : \mathcal{L}(\mathcal{H}_1) \to \mathcal{L}(\mathcal{H}_2)$ is called *n-positive* provided that $\Phi^{(n)} : M_n(\mathcal{L}(\mathcal{H}_1)) \to M_n(\mathcal{L}(\mathcal{H}_2))$ is positive., i.e. for any $(W_{i,j}) \in M_n(\mathcal{L}(\mathcal{H}_1))$ with $(W_{i,j}) \geq 0$ then $(\Phi(W_{i,j})) \geq 0$. We say that Φ is *completely positive* (abbreviated as CP) if Φ is n-positive $\forall n \in \mathbb{N}$.

Remark 4.4. Recall that in Example **??** where we discussed about the trivial measurement system of dimension n, we concluded that the map Φ_A must be positive. It should, actually be, n-positive. But since $n \in \mathbb{N}$ is arbitrary, we really require the map to be n-positive for all $n \in \mathbb{N}$, that is, completely positive.

Recall that in $\mathcal{H} \otimes \mathbb{C}^n \simeq \underbrace{\mathcal{H} \oplus \cdots \oplus \mathcal{H}}_{n \text{ times}}$. Then for $\begin{pmatrix} k_1 \\ \vdots \\ k_n \end{pmatrix}, \begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix} \in \mathcal{H} \otimes \mathbb{C}^n$, we have

$$\left\langle \begin{pmatrix} k_1 \\ \vdots \\ k_n \end{pmatrix}, \begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix} \right\rangle = \sum_{i=1}^{n} \langle k_i, h_i \rangle.$$

So, if $(X_{ij}) \in M_n(\mathcal{L}(\mathcal{H}))$, then

$$\left\langle \begin{pmatrix} k_1 \\ \vdots \\ k_n \end{pmatrix}, (X_{ij}) \begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix} \right\rangle = \left\langle \begin{pmatrix} k_1 \\ \vdots \\ k_n \end{pmatrix}, \begin{pmatrix} \sum_{j=1}^n X_{1j}h_j \\ \vdots \\ \sum_{j=1}^n X_{nj}h_j \end{pmatrix} \right\rangle = \sum_{i,j=1}^n \langle k_i, X_{ij}h_j \rangle.$$

Therefore, $(X_{ij}) \geq 0$ if and only if $\sum_{i,j=1}^n \langle h_i, X_{ij}h_j \rangle \geq 0$ for all $h_1, ..., h_n \in \mathcal{H}$.

Lemma 4.5. (1) *Let $Y, B \in \mathcal{L}(\mathcal{H})$. Then $Y \geq 0 \implies B^*YB \geq 0$.*
(2) *Let $Y = (Y_{ij}) \in M_n(\mathcal{L}(\mathcal{H}))$. Then $(Y_{ij}) \geq 0 \implies (B^*Y_{ij}B) \geq 0$ for all $B \in \mathcal{L}(\mathcal{H})$.*

Proof. To see (1), note that for every $h \in \mathcal{H}$, we have $\langle h, B^*YBh \rangle = \langle Bh, Y(Bh) \rangle \geq 0$. For (2), we have

$$\left\langle \begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix}, (B^*Y_{ij}B) \begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix} \right\rangle = \left\langle \begin{pmatrix} Bh_1 \\ \vdots \\ Bh_n \end{pmatrix}, (Y_{ij}) \begin{pmatrix} Bh_1 \\ \vdots \\ Bh_n \end{pmatrix} \right\rangle \geq 0,$$

so that $(B^*Y_{ij}B) \geq 0$. □

Proposition 4.6. *A map $\Phi : \mathcal{L}(\mathcal{H}) \longrightarrow \mathcal{L}(\mathcal{H})$ of the form $\Phi(X) = \sum_{\ell=1}^L B_\ell^* X B_\ell$ is completely positive. In particular, quantum channels arising from measurement systems are completely positive.*

Proof. We need Φ to be n-positive for all n. Let $(Y_{ij}) \in M_n(\mathcal{L}(\mathcal{H}))$ be positive. Then

$$(\Phi(Y_{ij})) = \left(\sum_{\ell=1}^L B_\ell^* Y_{ij} B_\ell \right) = \sum_{\ell=1}^L (B_\ell^* Y_{ij} B_\ell) \geq 0,$$

where the last inequality follows from the fact that sum of positive semidefinite matrices is positive semidefinite. □

Proposition 4.7. *Let $\Phi : M_p \longrightarrow M_p$ defined by $\Phi(X) = X^t$. Then Φ is positive but not 2-positive.*

Proof. To see Φ is positive must show that for $X = (x_{ij}) \in M_p, X \geq 0$ then $X^t = (x_{ji}) \geq 0$. Let $\begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_p \end{pmatrix} \in \mathbb{C}^p$, so that,

$$\begin{aligned}\left\langle \begin{pmatrix}\lambda_1\\ \vdots \\ \lambda_p\end{pmatrix}, X^t \begin{pmatrix}\lambda_1\\ \vdots \\ \lambda_p\end{pmatrix} \right\rangle &= \sum_{i,j=1}^{p} \overline{\lambda_i} x_{ij}^t \lambda_j \\ &= \sum_{i,j=1}^{p} \overline{\lambda_i} x_{ji} \lambda_j \\ &= \sum_{i,j=1}^{p} \overline{\lambda_j} x_{ij} \lambda_i \\ &= \left\langle \begin{pmatrix}\overline{\lambda_1}\\ \vdots \\ \overline{\lambda_p}\end{pmatrix}, (x_{ij}) \begin{pmatrix}\overline{\lambda_1}\\ \vdots \\ \overline{\lambda_p}\end{pmatrix} \right\rangle \geq 0,\end{aligned}$$

which implies that $X^t \geq 0$ and hence Φ is positive.

To see that it is not 2-positive, we let $\{E_{ij} : 1 \leq i, j \leq p\}$ be the usual matrix units for M_p. We claim that

(1) $E = \begin{pmatrix} E_{11} & E_{12} \\ E_{21} & E_{22}\end{pmatrix} \in M_2(\mathcal{L}(\mathbb{C}^p))$ is positive; but

(2) $\Phi^{(2)}(E) = \Phi^{(2)}\left(\begin{pmatrix} E_{11} & E_{12} \\ E_{21} & E_{22}\end{pmatrix}\right) = \begin{pmatrix} E_{11}^t & E_{12}^t \\ E_{21}^t & E_{22}^t\end{pmatrix} = \begin{pmatrix} E_{11} & E_{21} \\ E_{12} & E_{22}\end{pmatrix}$ is not positive.

First, let $h_1, h_2 \in \mathbb{C}^p$, and we obtain

$$\begin{aligned}\left\langle \begin{pmatrix}h_1\\ h_2\end{pmatrix}, \begin{pmatrix} E_{11} & E_{12} \\ E_{21} & E_{22}\end{pmatrix} \begin{pmatrix}h_1\\ h_2\end{pmatrix} \right\rangle &= (\langle h_1, e_1\rangle + \langle h_2, e_2\rangle)(\langle e_1, h_1\rangle + \langle e_2, h_2\rangle) \\ &= (\langle h_1, e_1\rangle + \langle h_2, e_2\rangle)(\overline{\langle h_1, e_1\rangle + \langle h_2, e_2\rangle}) \\ &= |\langle h_1, e_1\rangle + \langle h_2, e_2\rangle|^2 \geq 0.\end{aligned}$$

Therefore, $E \geq 0$ and this proves (1). To see (2), we observe that

$$\Phi^{(2)}(E) \begin{pmatrix} e_2 \\ -e_1\end{pmatrix} = \begin{pmatrix} E_{11} & E_{21} \\ E_{12} & E_{22}\end{pmatrix} \begin{pmatrix} e_2 \\ -e_1\end{pmatrix} = \begin{pmatrix} -e_2 \\ e_1\end{pmatrix} = -\begin{pmatrix} e_2 \\ -e_1\end{pmatrix}.$$

Therefore, -1 is an eigenvalue for $\Phi^{(2)}(E)$, so that $\Phi^{(2)}(E) \ngeq 0$ (if it were positive then all of its eigenvalues would be non-negative). □

4.1. **Choi-Krauss Representation.**

Theorem 4.8 (Choi, 1975)**.** *Let $\Phi : M_n \to M_d$ be a linear map. Then the following are equivalent:*

(1) Φ *is completely positive.*

(2) Φ *is n-positive.*
(3) $P_\Phi = (\Phi(E_{i,j}))_{i,j=1}^n \in M_n(M_d)$ *is positive.*
(4) *There exist* $B_1, ..., B_K \in M_{d,n}$ *such that* $\Phi(X) = \sum_{k=1}^K B_k X B_k^*$ *for all* $X \in M_n$.

Proof. $(1 \implies 2)$:
True by definition.

$(2 \implies 3)$:
Notice that it suffices to show that $Q := (E_{i,j})_{i,j=1}^n \in M_n(M_n)$ is positive, for then, due to the fact that Φ is n-positive, it is easy to see that

$$\Phi^{(n)}((E_{i,j})) = (\Phi(E_{i,j})) = P_\Phi \geq 0.$$

To this end, consider the matrix $Q := (E_{i,j})_{i,j=1}^n \in M_n(M_n)$ (e.g. the matrix of matrix units) and notice that,

$$Q^* = (E_{i,j})^* = (E_{j,i}^*) = (E_{i,j}) = Q$$

Also,

$$Q^2 = (\sum_{k=1}^n E_{i,k}E_{k,j}) = (\sum_{k=1}^n E_{i,j}) = (nE_{i,j}) = nQ.$$

Note that any eigenvector x of H, and its corresponding eigenvalue λ, then by spectral mapping theorem, we have

$$\lambda^2 x = Q^2 x = nQx = n\lambda x$$

so that $\lambda^2 - n\lambda = 0$ which implies that $\lambda \in \{0, n\}$. Since all the eigenvalues of Q are non-negative, we conclude that $Q \geq 0$. This proves the implication.

$(4 \implies 1)$:
Let $(X_{ij}) \in M_p(M_n)$ be positive. Then

$$\Phi^{(m)}((X_{ij})) = (\Phi(X_{ij})) = \left(\sum_{k=1}^K B_k X_{ij} B_k^*\right) = \sum_{k=1}^K (B_k X_{ij} B_k^*) \geq 0,$$

where the last inequality follows from the fact that sum of positive semidefinite matrices is positive semidefinite. This implies that $\Phi^{(m)}$ is positive. Since m is arbitrary, Φ is completely positive.

$(3 \implies 4)$:
Since $P_\Phi \geq 0$ and is an $nd \times nd$ matrix, there exist vectors $v_1, \ldots, v_K \in$

$\mathbb{C}^{nd}$ such that

$$P_\Phi = v_1 v_1^* + \cdots + v_K v_K^* = \sum_{k=1}^{K} v_k v_k^* = \sum_{k=1}^{K} |v_k\rangle\langle v_k|.$$

Since each $v_k \in \mathbb{C}^{nd}$, we can express these as an $n \times 1$ block matrix with each block being a $d \times 1$ matrix, that is, $v_k = \begin{pmatrix} h_1^k \\ h_2^k \\ \vdots \\ h_n^k \end{pmatrix}$ where $h_i^k \in \mathbb{C}^d$ for all i. We want to convert this $nd \times 1$ matrix into $d \times n$ matrix. So, we define $B_k := (h_1^k \ h_2^k \ \cdots \ h_n^k)$. Notice that B_k is a $d \times n$ matrix. Next, we claim that

$$\Phi(X) = \sum_{k=1}^{K} B_k X B_k^*.$$

In order to prove this claim, notice that

$$\begin{aligned} v_k v_k^* &= \begin{pmatrix} h_1^k \\ h_2^k \\ \vdots \\ h_n^k \end{pmatrix} \begin{pmatrix} h_1^{k*} & h_2^{k*} & \cdots h_n^{k*} \end{pmatrix} \\ &= \begin{pmatrix} h_1^k h_1^{k*} & h_1^k h_2^{k*} & \cdots & h_1^k h_n^{k*} \\ h_2^k h_1^{k*} & h_2^k h_2^{k*} & \cdots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ h_n^k h_1^{k*} & h_n^k h_2^{k*} & \cdots & h_n^k h_n^{k*} \end{pmatrix} = (h_i^k h_j^{k*}) \in M_n(M_d). \end{aligned}$$

Since $(\Phi(E_{ij})) = P_\Phi = \sum_{k=1}^{K} v_k v_k^* = \sum_{k=1}^{K} (h_i^k h_j^{k*})$, it follows that

$$\Phi(E_{ij}) = \sum_{k=1}^{K} h_i^k h_j^{k*} \text{ for all } 1 \le i, j \le n.$$

Now let us observe what these B_k do to a matrix unit;

$$\begin{aligned} B_k E_{ij} B_k^* &= (B_k E_{ii}) E_{ij} (E_{jj} B_k^*) \\ &= \begin{pmatrix} 0 & \cdots & h_i^k & \cdots & 0 \end{pmatrix} E_{ij} \begin{pmatrix} 0 \\ \vdots \\ h_j^{k*} \\ \vdots \\ 0 \end{pmatrix} = (h_i^k h_j^{k*})_{d \times d}, \end{aligned}$$

because $B_k E_{ii}$ is the matrix with 0's in every column, except that column i is preserved from B_k, and $E_{jj} B_k^*$ is the matrix with 0's in

every row, except that the i-th row becomes $\begin{pmatrix} h_1^{k*} & \cdots & h_n^{k*} \end{pmatrix}$. Thus we have

$$\sum_{k=1}^{K} B_k E_{ij} B_k^* = \sum_{k=1}^{K} (h_i^k h_j^{k*}) = \Phi(E_{ij}) \text{ for all } i, j$$

which in turn implies that

$$\sum_{k=1}^{K} B_k X B_k^* = \Phi(X).$$

□

Remark 4.9 (The Matrix of Matrix Units)**.** Observe that the $n \times n$ matrix of matrix units $(E_{ij})_{i,j=1}^n$ can be expressed as $(E_{ij}) = \sum_{i,j=1}^n E_{ij} \otimes E_{ij}$. Since $E_{ij} = e_i e_j^* = |i\rangle\langle j|$, it follows then that

$$(E_{ij}) = \sum_{i,j=1}^{n} E_{ij} \otimes E_{ij} = \sum_{i,j=1}^{n} |i\rangle\langle j| \otimes |i\rangle\langle j| = \sum_{i,j=1}^{n} (|i\rangle \otimes |i\rangle)(\langle j| \otimes \langle j|).$$

This matrix corresponds to "bell states" or "maximal entangled states" in quantum physics.

Definition 4.10 (Choi Rank)**.** Let $\Phi : M_n \to M_d$ be a completely positive map. The *Choi rank* of Φ is defined by

$$cr(\Phi) = \min\{K \in \mathbb{N} : \Phi(X) = \sum_{k=1}^{K} B_k X B_k^*\}.$$

Proposition 4.11. *If $\Phi : M_n \to M_d$ be a completely positive map, then $cr(\Phi) = rank\ (P_\Phi)$*

Proof. Let $Q = (E_{ij})$. Then it is easy to see that rank $(Q) = 1$. Notice that if $r = \text{rank}\ (P_\Phi)$, then as in the proof (3) $\implies$ (4) of Choi-Kraus representation, we may write $P_\Phi = \sum_{k=1}^r v_k v_k^*$. Then $cr(\Phi) \leq \text{rank}\ (P_\Phi)$. Now, let $r = cr(\Phi)$, and suppose that $\Phi(X) =$

$\sum_{k=1}^{r} A_k X A_k^*$ for some matrices A_k. Then

$$\begin{aligned} P_\Phi = (\Phi(E_{ij})) &= \left(\sum_{k=1}^{r} A_k E_{ij} A_k^*\right) \\ &= \sum_{k=1}^{r} \begin{pmatrix} A_k & & \\ & \ddots & \\ & & A_k \end{pmatrix} \begin{pmatrix} E_{11} & \cdots & E_{1n} \\ \vdots & \ddots & \vdots \\ E_{n1} & \cdots & E_{nn} \end{pmatrix} \begin{pmatrix} A_k^* & & \\ & \ddots & \\ & & A_k^* \end{pmatrix} \\ &= \sum_{k=1}^{r} \begin{pmatrix} A_k & & \\ & \ddots & \\ & & A_k \end{pmatrix} Q \begin{pmatrix} A_k^* & & \\ & \ddots & \\ & & A_k^* \end{pmatrix}. \end{aligned}$$

It then follows that

$$\text{rank } (P_\Phi) \leq \sum_{k=1}^{r} \text{rank } \left(\begin{pmatrix} A_k & & \\ & \ddots & \\ & & A_k \end{pmatrix} Q \begin{pmatrix} A_k^* & & \\ & \ddots & \\ & & A_k^* \end{pmatrix} \right) \leq r.$$

Thus, rank $(P_\Phi) \leq cr(\Phi)$. □

Lemma 4.12. *If $Y \in M_n$ and $tr(XY) = Tr(X)$ for all $X \in M_n$, then $Y = I_n$.*

Proof. Since the equality in the hypothesis is true for every X, it holds when $X = E_{ij}$, so that if $i = j$, we have

$$1 = \text{Tr}(X) = \text{Tr}(E_{ii}) = \text{Tr}(E_{ii}Y) = y_{ii}.$$

If $i \neq j$, then

$$0 = \text{Tr}(E_{ij}) = \text{Tr}(E_{ij}Y) = y_{ji}.$$

Thus, $Y = I_n$.

□

Corollary 4.13. *$\Phi : M_n \to M_d$ is a completely positive trace preserving (CPTP) map iff $\Phi(X) = \sum_{k=1}^{K} B_k X B_k^*$ with $\sum_{k=1}^{K} B_k^* B_k = I_n$.*

Proof. ($\Longleftarrow$): We know Φ is completely positive by Choi's theorem. To see that the map is trace preserving, we have,

$$\begin{aligned} Tr(\Phi(X)) = Tr(\sum_{k=1}^{K} B_k X B_k^*) &= \sum_{k=1}^{K} Tr(B_k X B_k^*) \\ &= \sum_{k=1}^{K} Tr(X B_k^* B_k) = Tr(X \sum_{k=1}^{K} B_k^* B_k) = Tr(X). \end{aligned}$$

($\Longrightarrow$):
By Choi, we know that there exist matrices B_k such that $\Phi(X) = \sum_{k=1}^{K} B_k X B_k^*$. Next, observe that for every X,

$$\begin{aligned}\mathrm{Tr}(X) = \mathrm{Tr}(\Phi(X)) &= \mathrm{Tr}\left(\sum_{k=1}^{K} B_k X B_k^*\right) \\ &= \mathrm{Tr}\left(\sum_{k=1}^{K} X B_k^* B_k\right) \\ &= \mathrm{Tr}\left(X\left(\sum_{k=1}^{K} B_k^* B_k\right)\right)\end{aligned}$$

which implies, from the preceding lemma, $\sum_{k=1}^{K} B_k^* B_k = I_n$. □

Remark 4.14. We see that $\Phi : M_n \to M_d$ is completely positive and trace preserving if and only if there exists $K \in \mathbb{N}$ and a K-outcome measurement system such that Φ is the map (or quantum channel) that comes from the measurement system.

The following theorem is due to R.G. Douglas. The proof would developing a bit more operator theory than we want to at this point. As we shall see, it has many applications. For a proof see [**?**].

Theorem 4.15 (Douglas' Majorization/Factorization Theorem)**.** *Let $\mathcal{H}, \mathcal{H}_A$ and $\mathcal{H}_B$ be Hilbert spaces with bounded, linear operators $A : \mathcal{H} \to \mathcal{H}_A$, and $B : \mathcal{H} \to \mathcal{H}_B$. Then the following are equivalent:*

1. *range $(A^*) \subseteq$ range (B^*),*
2. *there exists $\lambda > 0$ such that $A^*A \leq \lambda^2 B^*B$,*
3. *there exists $C : \mathcal{H}_B \to \mathcal{H}_A$ such that $A = CB$.*

Moreover, $\inf\{\lambda : A^*A \leq \lambda^2 B^*B\} = \inf\{\|C\| : A = CB\}$ *and both are attained.*

Corollary 4.16. *Let $A \in M_{k,d}$ and $B \in M_{r,d}$ with $r \leq K$. If $A^*A = B^*B$, then there exists $U = \begin{pmatrix} u_1 & \cdots & u_r \end{pmatrix} \in M_{k,r}$ with u_i's orthonormal such that $A = UB$.*

4.2. **Non-Uniqueness of Choi-Kraus Representation.** Let us motivate the next proposition. Recall that many ensembles $\{\phi_j, p_j\}$, $\{\psi_k, q_k\}$ can have the same density matrix and when they do, measurement systems will not be able to distinguish between these ensembles. Example **??** and Problem **??** explain this situation
We can also have two measurement systems that transform every state or ensemble of states in the same way. Thus, we can not distinguish

between these two systems by seing how states are transformed after they pass through the systems. In particular, if we have two measurement systems $\{M_1, ..., M_r\}$ and $\{N_1, ..., N_K\}$, which generate the same quantum channel Φ, that is, Φ has two representations,

$$\Phi(X) = \sum_{\ell=1}^{r} M_\ell X M_\ell^* = \sum_{j=1}^{K} N_j X N_j^*,$$

then, if a state v passes through either of these measurement systems, we get the output $\Phi(vv^*)$. In this situation, not only are we not able to differentiate between two measurement systems, we are not even able to keep track of r and K, the number of measurement operators in each measurement systems. The next theorem describes when two measurement systems yield the same quantum channel.

Theorem 4.17. *Let $\Phi : M_n \to M_d$ be a CP map (not necessarily TP), $r = cr(\Phi)$, and suppose*

$$\Phi(X) = \sum_{j=1}^{r} V_j X V_j^* = \sum_{k=1}^{K} W_k X W_k^*$$

are two Choi-Kraus representations of Φ. Then

1. *there exists $u = (u_{ij}) \in M_{K,r}$ such that $W_i = \sum_{j=1}^{r} u_{ij} V_j$ for all i.*
2. *span$\{V_1, ..., V_r\}$ = span$\{W_1, ..., W_K\} \subseteq M_{d,n}$*
3. *$U^*U = I_r$.*

Proof. Given a matrix $V = \begin{pmatrix} h_1 & .... & h_n \end{pmatrix} \in M_{d\times n}$, let us identify it by $v = \begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix} \in M_{dn\times 1}$. In this way, we turn $V_1, ..., V_r, W_1, ..., W_k$ into vectors $v_1, ..., v_r, w_1, ..., w_k \in \mathbb{C}^{nd}$. By reversiing the proof (3) $\implies$ (4) of Choi's theorem, we can show that

$$P_\Phi = (\Phi(E_{ij})) = \sum_{j=1}^{r} v_j v_j^* = \sum_{k=1}^{K} w_k w_k^*.$$

Hence by Douglas' Factorization theorem, with $B = \begin{pmatrix} v_1^* \\ \vdots \\ v_r^* \end{pmatrix}$ and $A = \begin{pmatrix} w_1^* \\ \vdots \\ w_k^* \end{pmatrix}$ there exists $U = (u_{ij})$ such that $\begin{pmatrix} w_1^* \\ \vdots \\ w_k^* \end{pmatrix} = U \begin{pmatrix} v_1^* \\ \vdots \\ v_r^* \end{pmatrix}$. which implies $W_i = \sum_{j=1}^r u_{ij} V_j$, and this proves (1).

Since V_j's and W_i's are interchangeable in (1), the above tells us that $\text{span}\{V_1, ..., V_r\} = \text{span}\{W_1, ..., W_m\}$.

To prove (3), if $r = cr(\Phi)$ then this implies $r = rank(P_\Phi)$. Thus $P_\Phi = \sum_{j=1}^r v_j v_j^*$ is written in the minimal way in the Douglas' Factorization and this implies $U^*U = I_r$. □

Conversely, we have the following theorem:

Proposition 4.18. *If $\Phi(X) = \sum_{j=1}^r V_j X V_j^*$ and $U = (u_{ij})_{K\times r}$ with $U^*U = I_r$ and if we set $W_i = \sum_{j=1}^r u_{ij} V_j$. Then*

$$\Phi(X) = \sum_{i=1}^K W_i X W_i^*.$$

Proof. We have

$$\begin{aligned} \sum_{i=1}^K W_i X W_i^* &= \sum_{i=}^K (\sum_{j=1}^r u_{ij} V_j) X (\sum_{j=1}^r u_{i\ell} V_\ell)^* \\ &= \sum_{i=1}^K \sum_{j,\ell=1}^r u_{ij} \overline{u}_{i\ell} V_j X V_\ell^* \\ &= \sum_{j,\ell=1}^r \left(\sum_{i=1}^K u_{ij} \overline{u}_{i\ell} \right) V_j X V_\ell^* \\ &= \sum_{j=1}^r V_j X V_j^* = \Phi(X) \end{aligned}$$

where we used the fact that $\sum_{i=1}^K u_{ij} \overline{u}_{ik\ell}$ equals 0 if $j \neq \ell$ and equals 1 if $j = \ell$. □

Problem 4.19. *Show directly that if $\Phi : M_n \to M_d$ is given by $\Phi(X) = \sum_{\ell=1}^q W_\ell X W_\ell^*$, then $\sum_{\ell=1}^q w_\ell w_\ell^* = P_\Phi = (\Phi(E_{ij}))$. Here given $W = \begin{pmatrix} w_1 & \cdots & w_n \end{pmatrix} \in M_{d,n}$, we associate W with $w = \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix} \in \mathbb{C}^{nd}$.*

Proof. **Homework problem 4**; due 28th January, Thursday. □

Problem 4.20. *A linear functional $\varphi : M_n \to \mathbb{C}$ is called an (abstract) state if*

(1) *whenever $P \geq 0$ in M_n, we have $\varphi(P) \geq 0$, and*
(2) *$\varphi(I_n) = 1$.*

Prove that φ is a state if and only if $P_\varphi = \big(\varphi(E_{ij})\big) \geq 0$ and $Tr(P_\Phi) = 1$. Also, show that $\varphi(X) = Tr(XP_\varphi^t)$.

Proof. **Homework problem 5**; due 28th January, Thursday. □

Problem 4.21. *Let $\Phi : M_n \to M_n$ be given by $\Phi(X) = \frac{1}{n-1}(Tr(X)I_n - X^t)$. Prove that Φ is completely positive and trace-preserving map (Φ is commonly known as the Werner-Holevo Channel).*

Proof. **Homework problem 6**; due 28th January, Thursday. □

The next subsection requires some basic results about partial isometries.
We say that $W : \mathcal{H}_1 \to \mathcal{H}_2$ is a *partial isometry* if there are subspaces $\mathcal{V}_1 \subseteq \mathcal{H}_1$ and $\mathcal{V}_2 \subseteq \mathcal{H}_2$ such that when we write $h \in \mathcal{H}_1$ as $h = v_1 + v_1^\perp$ where $v_1 \in \mathcal{V}_1$ and $v_1^\perp \in \mathcal{V}_1^\perp$, then $Wh = Wv_1 \in \mathcal{V}_2$ and $\|Wh\| = \|Wv_1\| = \|v_1\|$ and $W(\mathcal{V}_1) = \mathcal{V}_2$. We call $\mathcal{V}_1$ the *initial space* of W and $\mathcal{V}_2$ the *final space* of W.

Proposition 4.22. *Let $W : \mathcal{H}_1 \to \mathcal{H}_2$ be a bounded linear map. The following are equivalent.*

(1) *W is a partial isometry.*
(2) *W^* is a partial isometry.*
(3) *W^*W is a projection.*
(4) *WW^* is a projection.*

Moreover, when this holds, W^*W is the projection onto the initial space of W (which is $(\ker(W))^\perp$) and WW^* is the projection onto the final space of W (which is ran (W)).

5. Applications to Quantum Error Correction

We apply these results about Choi-Kraus representations to *quantum error correction*. When we pass quantum states through a channel, errors happen, and the output states may not be the same as the input

states. The errors, however, must be quantum events and so represented by a quantum channel. So we have some error operator $\mathcal{E}$ with

$$\mathcal{E}(X) = \sum_{i=1}^{r} E_i X E_i^*$$

where $\sum_{i=1}^{r} E_i^* E_i = 1$. (In particular, if the input is of the form $vv^*, v \in \mathbb{C}^n$ and the output is $\mathcal{E}(vv^*)$ for some map $\mathcal{E}$, then since this is a quantum event, we want $\mathcal{E}$ to be CPTP, of the form mentioned above). We don't expect that we will be able to fix all the errors but in the **Knill-LaFlamme model**, they would like there to be a subspace of the state space $\mathcal{L} \subset H$ and if we have $v \in \mathcal{L}$ and pass vv^* through the channel we can still recover the original state. Such an $\mathcal{L}$ is called a *protected subspace*.

To recover, we will need another "recovery" CPTP map $\mathcal{R}$ such that

$$\mathcal{R}(\mathcal{E}(vv^*)) = vv^*$$

for all $v \in \mathcal{L}$.

5.1. **Knill-LaFlamme Protected Subspace.**

We have the following key theorem.

Theorem 5.1 (Knill-LaFlamme)**.** *Let $\mathcal{L} \subset \mathbb{C}^n$ be nonzero subspace and let $P : \mathbb{C}^n \to \mathcal{L}$ be the orthogonal projection onto $\mathcal{L}$. Let $\mathcal{E} : M_n \to M_d$ be a CPTP map given by*

$$\mathcal{E}(X) = \sum_{i=1}^{m} E_i X E_i^*.$$

Then there exists a CPTP map $\mathcal{R} : M_d \to M_n$ such that for all $v \in \mathcal{L}$, $\mathcal{R}(\mathcal{E}(vv^*)) = vv^*$ if and only if $PE_i^* E_j P = \alpha_{ij} P$ for some scalars $\alpha_{ij} \in \mathbb{C}$.

Remark 5.2. We remark that $\mathcal{R}(\mathcal{E}(vv^*)) = vv^*$ for all $v \in \mathcal{L} \iff \mathcal{R}(\mathcal{E}(T)) = T$ for all $T \in \text{span}\,\{vv^* : v \in \mathcal{L}\} \iff (\mathcal{E}(PXP)) = PXP$ for all $X \in M_n$. The first equivalence follows from linearity. For the second equivalence, we first suppose that $X \geq 0$. Then since $P = P^*$, we have $PXP \geq 0$. We write PXP as a sum of rank-one operators (using an orthonormal basis of eigenvectors). Since ran $(PXP) \subseteq \mathcal{L}$, we can write PXP as a sum of rank-one operators of the form vv^* where $v \in \mathcal{L}$. Hence, $PXP \in \text{span}\,\{vv^* : v \in \mathcal{L}\}$. If $X \in M_n$ is arbitrary, then it may be written as a linear combination of at most four positive-semidefinite matrices, so that by linearity, $PXP \in \text{span}\,\{vv^* : v \in \mathcal{L}\}$ as well, as claimed.

Proof. ($\Longrightarrow$) Assume that a recovery map $\mathcal{R}$ exists. Then by Choi-Kraus write $\mathcal{R}(Y) = \sum_l A_l Y A_l^*$ with $A_l^* A_l = I$. Then

$$\psi(X) := PXP = \mathcal{R}(\mathcal{E}(PXP)) = \sum_{\ell,i} A_\ell E_i PXP E_i^* A_\ell^*,$$

while

$$\psi(X) := PXP = \sum_{\ell,i} (A_\ell E_i P) X (P E_i^* A_\ell^*).$$

These are two Choi-Kraus representations of $\psi : M_n \to M_n$. The representation $X \mapsto PXP$ must be minimal since $(PE_{ij}P)_{i,j=1}^n \in M_n(M_n)$ cannot have zero rank. Therefore, $cr(\psi) = 1$. By the theorem on minimal Choi-Kraus representations, each $A_\ell E_i P \in \text{span}\{P\}$. Hence, there are $\beta_{\ell,i} \in \mathbb{C}$ such that $A_\ell E_i P = \beta_{\ell i} P$. Moreover, by the same theorem, letting

$$U = (\beta_{11}, \beta_{12}, ..., \beta_{1m}, \beta_{21,}..., \beta_{2m}, ..., \beta_{q1}, ..., \beta_{qm})^t,$$

we have $U^*U = I_r = 1 = \sum_{\ell,i} |\beta_{\ell,i}|^2$. Now

$$(PE_i^* A_l^*)(A_l E_j P) = (\beta_{li} P)^* (\beta_{lj} P) = \overline{\beta}_{li} \beta_{lj} P,$$

so that,

$$\begin{aligned} \sum_l (PE_i^* A_l^*)(A_l E_j P) &= PE_i^* (\sum_l A_l^* A_l) E_j P \\ &= PE_i^* E_j P \\ &= (\sum_l \overline{\beta}_{li} \beta_{lj}) P \end{aligned}$$

Setting $\alpha_{ij} = \sum_l \overline{\beta}_{li} \beta_{lj}$, we get $PE_i^* E_j P = \alpha_{ij} P$.

($\Longleftarrow$) Assuming $PE_i^* E_j P = \alpha_{ij} P$, we want to build the recovery operator R. We have

$$\begin{aligned} (\alpha_{ij} P) &= (\alpha_{ij}) \otimes P \\ &= (PE_i^* E_j P) \\ &= \begin{pmatrix} PE_1^* \\ . \\ . \\ . \\ PE_m^* \end{pmatrix} \begin{pmatrix} E_1 P & ... & E_m P \end{pmatrix} \geq 0 (\text{ since it is of the form } X^* X) \end{aligned}$$

This implies $(\alpha_{ij}) \geq 0$. Hence (α_{ij}) is unitarily equivalent to a diagonal matrix and therefore there exists $U \in M_{m \times m}$ unitary such that

$$U^*(\alpha_{ij}) U = D = diag(d_{11}, ..., d_{mm}),$$

where d_{ii} are the eigenvalues of (α_{ij}) and $d_{ii} \geq 0$. Since, $\mathcal{E}$ is CPTP, we have $\sum_{i=1}^{m} E_i E_i^* = I$, we have

$$\sum_{i=1}^{m} \alpha_{ii} P = \sum_{i=1}^{m} P E_i^* E_i P = P^2 = P.$$

Hence $\text{trace}((\alpha_{ij})) = \sum_{i=1}^{m} \alpha_{ii} = 1 = \sum_{i=1}^{m} d_{ii}$.

Set $F_i = \sum_j u_{ij} E_j$, where u_{ij} is the (i,j)-th entry of U. Then we have,

$$\begin{aligned}
\sum_{i=1}^{m} F_i X F_i^* &= \sum_{i=1}^{m} \left(\sum_{j=1}^{m} u_{ij} E_j \right) X \left(\sum_{k=1}^{m} \overline{u}_{ik} E_k^* \right) \\
&= \sum_{j,k=1}^{m} \left(\sum_{i=1}^{m} u_{ij} \overline{u}_{ik} \right) E_j X E_k^* \\
&= \sum_{j=1}^{m} E_j X E_j^*.
\end{aligned}$$

Therefore, $\mathcal{E}(X) = \sum_i F_i X F_i^*$ as well. In particular, one has

$$\begin{aligned}
P F_i^* F_j P &= P(\sum_j \overline{u}_{il} E_l)(\sum_k u_{jk} E_k) P \\
&= \sum_{jk} \overline{u}_{il} u_{jk} P E_j^* E_k P \\
&= (\sum_{j,k} \overline{u}_{ij} \alpha_{jk} u_{jk}) P \\
&= [(i,j) \text{ entry of } U^* (\alpha_{ij}) U] P.
\end{aligned}$$

Therefore, $P F_i^* F_j P = 0$ if $i \neq j$ and $d_{ii} P$ if $i \neq j$. Thus $P F_i^* F_j P = d_{ij} P$. Define $V_i = 0$ if $d_{ii} = 0$ and $\frac{1}{\sqrt{d_{ii}}} F_i P$ if $d_{ii} > 0$. Then

$$V_i^* V_i = \frac{1}{d_{ii}} P F_i^* F_i P = \frac{1}{d_{ii}} d_{ii} P = P.$$

Thus V is a partial isometry.

If $i \neq j$, then $V_i^* V_j = (1/\sqrt{d_{ii} d_{jj}}) P F_i^* F_j P = d_{ij} P / \sqrt{d_{ii} d_{jj}} = 0$. This implies that ran $V_j \subseteq \ker V_i^* = \text{ran}\, {}^{\perp} V_i$, so we have that ran $V_i \perp$ ran V_j if $i \neq j$. Thus, $V_i V_i^*$ are mutually orthogonal projections. So $R = \sum_{i=1}^{m} V_i V_i^*$ is also a projection.

Now we let $Q = I - R$ and define $\mathcal{R} : M_n \to M_n$ as the following,

$$\mathcal{R}(X) = \sum_{i=1}^{m} V_i^* X V_i + QXQ.$$

Since $\sum V_i V_i^* + Q^2 = I$, we know that $\mathcal{R}$ is CPTP.

What left for us to check is that $\mathcal{R}(\mathcal{E}(PXP)) = PXP$. To this end, we can compute that

$$\begin{aligned}
\mathcal{R}(\mathcal{E}(PXP)) &= \sum_{i=1}^{m} V_i^* \mathcal{E}(PXP) V_i + Q\mathcal{E}(PXP)Q \\
&= \sum_{i.j} V_i^* F_j PXP F_j^* V_i + \sum_j QF_j PXP F_j^* Q \\
&= \sum_{i,j} \frac{1}{d_{ii}} (PF_i^* F_j P) X (PF_j^* F_i P) + \sum_j d_{jj} QV_j X V_j^* Q, \quad QV_j = 0 \\
&= \sum_{i=1}^{m} d_{ii} PXP \\
&= PXP.
\end{aligned}$$

□

Theorem 5.3. *Let $\mathcal{L} \subset \mathbb{C}^n$ be nonzero subspace and let $P : \mathbb{C}^n \to \mathcal{L}$ be the orthogonal projection onto $\mathcal{L}$. Let $E_1, ..., E_m \in M_{d,n}$ and let $\mathcal{E} : M_n \to M_d$ be a map given by*

$$\mathcal{E}(X) = \sum_{i=1}^{m} E_i X E_i^*.$$

Let $\mathcal{R} : M_d \to M_n$ be the operator obtained in the previous theorem.Choose $\widetilde{E}_i \in span\{E_1, ..., E_m\}$ with $\sum_{i=1}^{\widetilde{m}} \widetilde{E}_i \widetilde{E}_i^* = I$, and let

$$\widetilde{\mathcal{E}}(X) = \sum_{i=1}^{\widetilde{m}} \widetilde{E}_i X \widetilde{E}_i^*.$$

Then $\mathcal{R} \circ \widetilde{E}(vv^*) = vv^*$ for all $v \in \mathcal{L}$.

Proof. Since $F_1, ..., F_m$ be the operators defined in the previous proof. It was shown that the spans are equal; that is, span $\{F_1, ..., F_m\}$ = span $\{E_1, ..., E_m\}$. Hence, we may write

$$\widetilde{E}_i = \sum_{\ell=1}^{m} \beta_{i\ell} F_\ell.$$

We obtain

$$\begin{aligned} I &= \sum_{i=1}^{\widetilde{m}} \widetilde{E}_i \widetilde{E}_i^* \\ &= \sum_i \left(\sum_\ell \overline{\beta}_{i\ell} F_\ell^* \right) \left(\sum_k \beta_{ik} F_k \right) \\ &= \sum_{i,\ell,k} \overline{\beta}_{i\ell} \beta_{ik} F_\ell^* F_k \end{aligned}$$

Therefore, if P is the orthogonal projection onto $\mathcal{L}$, we obtain

$$\begin{aligned} P = PIP &= \sum_{i,\ell,k} \overline{\beta}_{i\ell} \beta_{ik} P F_\ell^* F_k P \\ &= \sum_{i,\ell,k} \overline{\beta}_{i\ell} \beta_{ik} d_{\ell k} P, \end{aligned}$$

where the $d_{\ell k}$'s are from the diagonal matrix in the previous proof. Then we have $\sum_{i,\ell,k} \overline{\beta}_{i\ell} \beta_{ik} d_{\ell k} = 1$, so that

$$\sum_{i,k} |\beta_{ik}|^2 d_{kk} = 1,$$

since $d_{\ell k} = 0$ if $\ell \neq k$. Using the notation of the previous proof (recall that $V_i^* = \frac{1}{\sqrt{d_{ii}}} P F_i^*$), we have

$$\begin{aligned} \mathcal{R}\widetilde{\mathcal{E}}(PXP) &= \sum_{i,j} V_i^* \widetilde{E}_i P X P \widetilde{E}_j^* V_i + Q \left(\sum \widetilde{E}_j^* P X P \widetilde{E}_j \right) Q \\ &= \sum_{i,j} \frac{1}{d_{ii}} (P F_i^* \widetilde{E}_j P) X (P \widetilde{E}_j^* F_i P) \\ &= \sum_{i,j,k,\ell} \overline{\beta}_{jk} \beta_{j\ell} \frac{1}{d_{ii}} (P F_i^* F_\ell P) X P (F_k^* F_i) P \\ &= PXP. \end{aligned}$$

Hence, $\mathcal{R} \circ \widetilde{E}(vv^*) = vv^*$ whenever $v \in \mathcal{L}$ with $\|v\| = 1$.

□

The beautiful thing about Knill-LaFlamme is that it looks like correcting one error but it actually corrects the whole family.

5.2. **Brief introduction to binary codes.** We can motivate some of this model by considering transmission. Let us first consider the case of Binary communication.
In the binary case, we transmit bits (basic units), where each bit is in $\{0,1\} = \mathbb{Z}_2$, the field of two elements. We define an *n-bit* to be an n-tuple $\delta = (\delta_1, ..., \delta_n) \in \mathbb{Z}_2^n$ in the vector space $\mathbb{Z}_2^n$ of dimension n over the field $\mathbb{Z}_2$.

$$\begin{pmatrix}\text{Input}\\ \delta\end{pmatrix} \xrightarrow[ErrorOccurs]{Channel} \begin{pmatrix}\text{Output}\\ \mathcal{E}(\delta)\end{pmatrix}$$

The aim is to build codes such that when we transmit an n-bit through a channel, we obtain the same n-bit. The usual model for errors in this case is bit flips (a bit flipping $1 \leftrightarrow 0$). We do not usually think of errors as being permutations of the bits.

We define *binary codes* to be certain special subsets of $\mathbb{Z}_2^n$. The *linear codes* correspond to a linear subspace, say k-dimensional where $k \leq n$. In this case, we have information in $\mathbb{Z}_2^k$ and encode it in $\mathbb{Z}_2^n$, and we transmit the encoded information (which will now come with errors) into $\mathbb{Z}_2^n$, and decode back into $\mathbb{Z}_2^k$. Pictorially,

$$\mathbb{Z}_2^k(\text{Information}) \xrightarrow{Encode} \mathbb{Z}_2^n \xrightarrow[Errors]{Transmit} \mathbb{Z}_2^n \xrightarrow{Decode} \mathbb{Z}_2^k(\text{Get information}).$$

One example of a binary code is the *majority rule code*, where each bit is repeated an odd number of times. For example, if $k = 4$ we could encode $(1, 0, 0, 1)$ in $\mathbb{Z}_2^{12}$ by repeating each bit three times, to get

$$(111, 000, 000, 111).$$

As we transmit this, bit flips may happen; for example we may get

$$(110, 010, 110, 011).$$

To decode in this case, we "vote" (choose which number has more repetition in a given block), and get

$$(1, 0, 1, 1),$$

which is different from the input. If only one bit flip occurs in each block, after decoding we still obtain what we sent. (Notice that there has been two bit flips in the third block in this example and hence we did not get the right information after decoding.) This rule (Coding and Decoding) works for one bit-flip per block. If there are more than one bit-flip, we don't get what we sent.

5.3. **Shor's code.** In the quantum world, we transmit *qubits* (quantum bits), where a qubit is a unit vector in $\mathbb{C}^2$. We write n-qubits as a unit vector in

$$\underbrace{\mathbb{C}^2 \otimes \cdots \otimes \mathbb{C}^2}_{n \text{ times}} = (\mathbb{C}^2)^{\otimes n}.$$

Let e_0, e_1 be an orthonormal basis for $\mathbb{C}^2$, so that $e_0 = |0\rangle$ and $e_1 = |1\rangle$. We obtain an orthonormal basis for $(\mathbb{C}^2)^{\otimes n}$ formed by all elements of the form

$$e_{i_1} \otimes \cdots \otimes e_{i_n} = |i_1 i_2 \cdots i_k\rangle,\ i_j \in \{0, 1\}.$$

By an early theorem, this set of vectors is perfectly distinguishable.

This is a nice application of Knill-LaFlamme. It is a quantum analogue of the majority rule code.

Definition 5.4 (Pauli Matrices)**.** The following are called *Pauli matrices*:

$$I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, Y = \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix}, Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

We note that the Pauli matrices form an orthogonal basis for M_2 and each Pauli matrix is a unitary.

Definition 5.5 (1-Pauli Matrices)**.** The following are called *1-Pauli matrices* for $(\mathbb{C}^2)^{\otimes n}$:

$$\begin{aligned}
I &= \underbrace{I_2 \otimes \cdots \otimes I_2}_{n \text{ times}}, \\
X_i &= \underbrace{I_2 \otimes \cdots \otimes I_2}_{i-1 \text{ times}} \otimes \underbrace{X}_{\text{i-th place}} \otimes \underbrace{I_2 \otimes \cdots \otimes I_2}_{n-i \text{ times}} \\
Y_i &= \underbrace{I_2 \otimes \cdots \otimes I_2}_{i-1 \text{ times}} \otimes \underbrace{Y}_{\text{i-th place}} \otimes \underbrace{I_2 \otimes \cdots \otimes I_2}_{n-i \text{ times}} \\
Z_i &= \underbrace{I_2 \otimes \cdots \otimes I_2}_{i-1 \text{ times}} \otimes \underbrace{Z}_{\text{i-th place}} \otimes \underbrace{I_2 \otimes \cdots \otimes I_2}_{n-i \text{ times}}
\end{aligned}$$

There are $3n + 1$ 1-Pauli matrices.

For $A \in M_2$, the matrix

$$I_2 \otimes \cdots \otimes I_2 \otimes A \otimes I_2 \otimes \cdots \otimes I_2$$

is in the span of the 1-Pauli matrices. If we could do Knill-Laflamme contruction in a way that "fixed" the errors caused by the 1-Pauli matrices, then we would "fix" all such $I_2 \otimes \cdots \otimes I_2 \otimes A \otimes I_2 \otimes \cdots \otimes I_2$. This is "like" a quantum bit flip; that is, errors happen on each qubit

site but don't do anything like permuting sites.

This is like having a subspace $\mathcal{L} \subseteq \mathbb{C}^{2^n}$ and P the orthogonal projection onto $\mathcal{L}$, such that $PU_i^* U_j P = \alpha_{ij} P$ whenever U_1, U_2 are 1-Pauli matrices. We can do this by considering the following subspace, from which we get Schor's code. Let $\mathcal{L} \subseteq (\mathbb{C}^2)^{\otimes 9}$ with $\mathcal{L} = \text{span}\,\{|0\rangle_L, |1\rangle_L\}$ where

$$|0\rangle_L = \frac{1}{2\sqrt{2}}(|000\rangle + |111\rangle) \otimes (|000\rangle + |111\rangle) \otimes (|000\rangle + |111\rangle),$$

and

$$|1\rangle_L = \frac{1}{2\sqrt{2}}(|000\rangle - |111\rangle) \otimes (|000\rangle - |111\rangle) \otimes (|000\rangle - |111\rangle).$$

Theorem 5.6. (Shor)

(1) *We have* $|0\rangle_L \perp |1\rangle_L$.
(2) *If* $U_1, ..., U_{28}$ *are the* 1*-Pauli matrices on* $(\mathbb{C}^2)^{\otimes 9}$*, then* $PU_i^* U_j P \in \{0, P\}$*, where* P *is the orthogonal projection onto* $\mathcal{L}$.

Hence, $\mathcal{L}$ is a protected subspace in the Knill-Laflamme sense. (So errors can be recovered whenever $\mathcal{E} : M_{2^9} \to M_{2^9}$ is an error operator with a Choi-Kraus representation consisting of matrices from the span of the 1-Pauli's.)

In this case, there is an operator $\mathcal{R} : M_{2^9} \to M_{2^9}$ such that if $v \in \mathcal{L}$ with $\|v\| = 1$, then $\mathcal{R} \circ \mathcal{E}(vv^*) = vv^*$.

(If anything, this should tell us that doing Knill-Laflamme is hard.)

5.4. **Another Coding Viewpoint.** The idea is the following: suppose we obtain a quantum channel with a family of errors $\mathcal{E}_\lambda : M_n \to M_n$ where $\lambda \in \Lambda$ (Λ is some set). We have states in $\mathbb{C}^d$ where $d \leq n$. We want to "encode" these states in $\mathbb{C}^d$ as states in $\mathbb{C}^n$, pass them (or communicate them) through the given channel, and then "decode". Since states have to be unit vectors, we want the encoding map $V : \mathbb{C}^d \to \mathbb{C}^n$ to be an isometry (so that states go to states).

If $h \in \mathbb{C}^d$ is a state encoded as $Vh \in \mathbb{C}^n$, then notice that the rank one projections corresponding to h and Vh are related as follows:

$$P_{Vh} = (Vh)(Vh)^* = Vhh^*V^* = VP_hV^*.$$

So, the encoding map corresponds to (or can be identified with) a map $\Phi_V : M_d \longrightarrow M_n$ defined by $X \mapsto VXV^*$ which is clearly a CP map and since $\text{Tr}\,(VXV^*) = \text{Tr}\,(XV^*V) = \text{Tr}\,(XI_d) = \text{Tr}\,(X)$, this map is also trace preserving. Now we run $(Vh)(Vh)^*$ through the given channel

and obtain the output $\mathcal{E}_\lambda((Vh)(Vh)^*)$. We can do "blind decoding" by sending $\mathcal{E}_\lambda((Vh)(Vh)^*)$ to $V^*\mathcal{E}_\lambda((Vh)(Vh)^*)V$, so that if $\mathcal{E}_\lambda$ didn't do anything, then we would obtain the same output as what we put in. That is,

$$h \xrightarrow{Code} Vh \leftrightarrow (Vh)(Vh)^* \xrightarrow[\text{channel}]{\mathcal{E}_\lambda} \mathcal{E}_\lambda(Vhh^*V^*) \xrightarrow{Decode} V^*\mathcal{E}_\lambda(Vhh^*V^*)V.$$

Or, by using the identification of h by hh^*,

$$\left(hh^* \xrightarrow{Code} (Vh)(Vh)^* \xrightarrow[\text{channel}]{\mathcal{E}_\lambda} \mathcal{E}_\lambda(Vhh^*V^*) \xrightarrow{Decode} V^*\mathcal{E}_\lambda(Vhh^*V^*)V.\right)$$

In general, we have

$$P \xrightarrow{Code} VPV^* \xrightarrow[\text{channel}]{\mathcal{E}_\lambda} \mathcal{E}_\lambda(VPV^*) \xrightarrow{Decode} V^*\mathcal{E}_\lambda(VPV^*)V.$$

A measure of fidelity of this transmission can be given by

$$\|hh^* - V^*\mathcal{E}_\lambda((Vh)(Vh)^*)V\|_2,$$

which is just the square root of the sum of the squares of the entries of the above $d \times d$ matrix.

Define

$$e(V,\lambda) = \sup_{\|h\|=1} \|hh^* - V^*\mathcal{E}_\lambda((Vh)(Vh)^*)\|_2,$$

and define

$$e(V) = \sup_\lambda\{e(V,\lambda) : \lambda \in \Lambda\}.$$

The goal is when given a family of errors $\{\mathcal{E}_\lambda : \lambda \in \Lambda\}$, find the isometry $V_0 : \mathbb{C}^d \to \mathbb{C}^n$ such that $e(V_0) := \min_V e(V)$, that is, the goal is to solve $\min_V e(V)$. The solution V_0 where this minimum is attained would give the "best" encoding.

Theorem 5.7 (Bodmann-Kribs-Paulsen)**.** *For each $A \subseteq \{1,...,n\}$, let P_A be the orthogonal projection onto span $\{e_i : i \in A\}$. Let $\mathcal{E}_A : M_n \to M_n$ be given by*

$$\mathcal{E}_A(X) = P_AXP_A + (I - P_A)X(I - P_A).$$

Let $d \leq n$ and $V_0 : \mathbb{C}^d \to \mathbb{C}^n$ be an isometry. Then $e(V_0) = \min_V e(V)$ if and only if $V_0 = \begin{pmatrix} r_1 \\ \vdots \\ r_n \end{pmatrix}$ where $r_i^t \in \mathbb{C}^d$ is such that

(1) *V_0 is an isometry.*

(2) $\|r_i\|_2 = \|r_j\|_2$ *for all* i, j.
(3) $\max_{i \neq j}\{|\langle r_i | r_j \rangle|\}$ *is minimized over all possible isometries satisfying* (2).

Example 5.8. Let $d = 2$ (trying to think of a single qubit) and $n = N$. Then V_0 as above must be an $N \times 2$ matrix. In fact,

$$V_0 = \sqrt{\frac{2}{N}} \begin{pmatrix} \cos\left(\frac{2\pi}{N}\right) & \sin\left(\frac{2\pi}{N}\right) \\ \vdots & \vdots \\ \cos\left(\frac{2\pi N}{N}\right) & \sin\left(\frac{2\pi N}{N}\right) \end{pmatrix}.$$

These are optimal vectors in $\mathbb{R}^2$, but the answer is not known in $\mathbb{C}^2$ for general n.

These ideas are closely related to Zauner's conjecture.

Conjecture 5.9 (Zauner)**.** For every $n \in \mathbb{N}$, there is a set of n^2 unit vectors $\{v_1, ..., v_{n^2}\} \in \mathbb{C}^n$ with $|\langle v_i, v_j \rangle|$ constant for all $i \neq j$.

If we have such a set as in Zauner's conjecture, with $P_i = \frac{1}{n} v_i v_i^*$, then $\{P_i : 1 \leq i \leq n^2\}$ defines what is called a *symmetric informationally complete positive operator-valued measure (SIC POVM).*
The constant $|\langle v_i, v_j \rangle|^2 = \frac{1}{n+1}$ for all $i \neq j$, is known as the *Welch constant* for any set of n^2 unit vectors with constant inner product (of course for distinct vectors). If we set $V_0 = \frac{1}{n} \begin{pmatrix} v_1^t \\ \vdots \\ v_{n^2}^t \end{pmatrix} : \mathbb{C}^n \to \mathbb{C}^{n^2}$, then this is an isometry and $\max_{i \neq j} |\langle v_i, v_j \rangle|$ is as small as possible among all such isometries. (This gives the optimal isometry in the B-K-P setting above.)

Problem 5.10. *In Shor's proof, for* $\mathcal{L} = span\ \{|0\rangle_L, |1\rangle_L\}$*, he proves that the subspaces*

$$\mathcal{L},\ X_1\mathcal{L}, ..., X_9\mathcal{L},\ Y_1\mathcal{L}, ..., Y_9\mathcal{L},\ Z_1\mathcal{L}$$

are pairwise orthogonal. This is $\binom{20}{2} = 190$ *cases. Show that* $\mathcal{L}$*,* $X_1\mathcal{L}$*,* $Y_1\mathcal{L}$ *and* $Z_1\mathcal{L}$ *are pairwise orthogonal. (This is* $\binom{4}{2} = 6$ *cases.)*

Proof. **Homework problem 7**; due 4th February, Thursday. □

Problem 5.11. *Find* $\{v_1, ..., v_4\} \subseteq \mathbb{C}^2$ *unit vectors with* $|\langle v_i, v_j\rangle|^2 = \frac{1}{3}$ *for all* $i \neq j$.

Proof. **Homework problem 8**; due 4th February, Thursday. □

6. One Shot Zero Error Capacity-Classical Case

6.1. **The Classical Case.** Given a finite alphabet $X = \{x_1, \ldots, x_n\}$ for Alice, suppose that Alice sends a message through a noisy channel to Bob, who has a finite alphabet $Y = \{y_1, \ldots, y_m\}$.

Let $p(y_i|x_j)$ denote the probability that Bob receives y_i when Alice sends x_j. Defining a matrix $N = (p(y_i|x_j))_{1\leq i\leq m,\ 1\leq j\leq n} = (p_{ij})_{1\leq i\leq m,\ 1\leq j\leq m} \in M_{m,n}$, we have $\sum_{i=1}^{m} p(y_i|x_j) = 1$. We call $N = (p(y_i|x_j)) = (p_{ij}) \in M_{m,n}$ the *noise matrix* for this channel. Since $p_{ij} \geq 0$ and the entries in each column of N sum to 1. N is *column stochastic*.

If α_j is the probability that Alice sends x_j, $\vec{\alpha} = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}$, and $\vec{\beta} = N\vec{\alpha} = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_m \end{pmatrix}$, then the i^{th} entry of the vector β, β_i, is the probability that Bob receives y_i.

Definition 6.1. The *Zero error capacity* is the maximum number of different $x_i's$ that Alice can send through the channel N so that when Bob receives something, he knows exactly which x_i was sent.

To figure out how to compute the zero error capacity of a classical channel, we will need some graph theory.

Definition 6.2. A graph $G = (V, E)$ consists of a finite set V, called the "vertex set", and a set $E \subseteq V \times V$, called the "edge set", such that

(1) $(v, v) \notin E$ for all $v \in V$, (no loop) and
(2) if $(v, w) \in E$ for some $v \neq w \in V$, then $(w, v) \in E$ (undirected edges).

The elements in V are called "vertices" of the graph; the elements in E are called "edges" of the graph, and if $(v, w) \in E$, we say that v and w are "adjacent" vertices. We write $v \sim w$ if $(v, w) \in E$.

A subset $S \subseteq V$ is called *independent* if for all $v, w \in S$, $(v, w) \notin E$. The *independence number* of the graph is the number

$$\alpha(G) = \max\{card(S) : S \text{ is an independent set}\}.$$

A subset $S \subseteq V$ is called a *clique* if for all $v, w \in S$, $v \neq w$ implies $(v, w) \in E$. The *clique number* of the graph is the number

$$\omega(G) = \max\{card(S) : S \text{ is a clique}\}.$$

Definition 6.3. Given a graph $G = (V, E)$ and a set $\{1, \ldots, c\} \subset \mathbb{N}$, a *c-coloring* of G is a map $f\colon V \to \{1, \ldots, c\}$ such that whenever $v \sim w$, $f(v) \neq f(w)$. The *chromatic* or *coloring number* of G, denoted by $\chi(G)$, is the least c for which there is a c-coloring for G.

Definition 6.4 (Shannon's Confusability Graph)**.** Given a noisy channel as above, the *Shannon Confusability Graph* is defined on the alphabet $\{x_1, \ldots, x_n\}$ as the set of vertices $V = \{x_1, \ldots, x_n\}$ with $x_i \sim x_j$ if and only if there exists a $y \in Y$ such that $p(y|x_i)\, p(y|x_j) \neq 0$. This graph is also known as the *confusability graph* of the channel.

If $x_j \sim x_j$ in the confusability graph G, then when Bob receives a y guaranteed by the definition of the edge set of G (i.e., a y such that $p(y|x_i)\, p(y|x_j) \neq 0$), he won't be able to know whether x_i or x_j was sent. But if Alice chooses $S \subset X$ that is an independent set and sends something from S, then when Bob receives a y, he knows exactly which $x \in S$ was sent.

The next proposition states that in terms of graphs, the (one-shot) zero error capacity of a noisy channel N is $\alpha(G)$ when G is the confusability graph of N.

Proposition 6.5. *Suppose we have a channel with noise matrix $N = (p(y_i|x_j))$, and let G be the corresponding confusability graph, then the* one shot zero error capacity of *this channel is $\alpha(G)$, the independence number of G.*

Proof. Suppose that $x_1, ..., x_k$ are letters that Alice can send so that Bob knows what letters were sent, such that k is the one-shot zero error capacity. If there were $y \in Y$ with $p(y|x_i) \cdot p(y|x_j) \neq 0$ for $1 \leq i \neq j \leq k$, then Bob is unable to know whether x_i or x_j was sent. By definition of one-shot zero error capacity, this cannot happen. Hence, $p(y|x_i) \cdot p(y|x_j) = 0$ for all $y \in Y$, so that $x_i \not\sim x_j$. Thus, $\{x_1, ..., x_k\}$ is independent in G so that $\alpha(G) \geq k$. Conversely, if $x_1, ..., x_\ell \in G$ are independent with $\ell = \alpha(G)$, then $p(y|x_i) \cdot p(y|x_j) \neq 0$ for all $1 \leq i \neq j \leq \ell$. By the same argument, $x_1, ..., x_\ell$ is a set of letters that Alice can send such that Bob can still know which letters were sent through the channel. Hence, ℓ is at most the one-shot zero error capacity. This shows that the one-shot zero error capacity for the channel is exactly $\alpha(G)$.

□

A natural question arises–what can be said about capacity and multiple uses of N?

Suppose we want to use the noisy channel with the noise matrix $N = (p\,(y_i|x_j))$ to send two letters. Then we have

$$(x_1, x_2) \xrightarrow{N\times N} (y_1, y_2),$$

with $p\big((y_1, y_2)|(x_1, x_2)\big)$ being the probability that (y_1, y_2) is received by Bob when Alice sent (x_1, x_2). If we assume that the noise acts independently, then

$$p\big((y_1, y_2)|(x_1, x_2)\big) = p(y_1|x_1) \cdot p(y_2|x_2).$$

The induced confusability graph $G^{(2)} = (V^{(2)}, E^{(2)})$ has vertex set $V^{(2)} = \{(x_1, x_2) : x_i \in X\} = V \times V$ with $(x_1, x_2) \sim (x_1', x_2')$ if and only if $(x_1, x_2) \neq (x_1', x_2')$ and $N \times N$ can send both to a common point (y_1, y_2); that is,

$$\begin{aligned}
(x_1, x_2) \sim (x_1', x_2') &\iff \exists (y_1, y_2) \ni p\big((y_1, y_2)|(x_1, x_2)\big) \cdot p\big((y_1, y_2)|(x_1', x_2')\big) \neq 0 \\
&\iff \exists (y_1, y_2) \ni p(y_1|x_1) \cdot p(y_2|x_2) \cdot p(y_1|x_1') \cdot p(y_2|x_2') \neq 0 \\
&\iff \exists (y_1, y_2) \ni \underbrace{p(y_1|x_1) \cdot p(y_1|x_1')}_{\neq 0 \,\Leftrightarrow\, x_1 \sim x_1' \text{ or } x_1 = x_1'} \cdot \underbrace{p(y_2|x_2) \cdot p(y_2|x_2')}_{\neq 0 \,\Leftrightarrow\, x_2 \sim x_2' \text{ or } x_2 = x_2'} \neq 0 \\
&\iff \left\{ \begin{array}{c} (x_1 \sim x_1' \ \textit{and}\ x_2' \sim x_2) \\ \text{OR} \\ (x_1 = x_1' \ \textit{and}\ x_2' \sim x_2) \\ \text{OR} \\ (x_1 \sim x_1' \ \textit{and}\ x_2' = x_2) \end{array} \right\} \\
&\iff \left\{ \begin{array}{c} (x_1 \sim x_1' \textit{or}\ x_1 = x_1') \\ \text{AND} \\ (x_2 \sim x_2' \ \textit{or}\ x_2 = x_2') \\ \text{Such that} \\ (\text{either } x_1 \neq x_1' \ \textit{or}\ x_2 \neq x_2') \end{array} \right\}
\end{aligned}$$

Definition 6.6 ($\simeq$)**.** If $x, x' \in G$, a graph, then $x \simeq x'$ means $x \sim x'$ or $x = x'$. It follows then that

$(x_1, x_2) \simeq (x_1', x_2') \Leftrightarrow \big(x_1 \simeq x_1'$ and $x_2 \simeq x_2'\big)$;

$(x_1, x_2) \sim (x_1', x_2') \Leftrightarrow \big(x_1 \simeq x_1'$ and $x_2 \simeq x_2'$ and $(x_1, x_2) \neq (x_1', x_2')\big)$.

This motivates a new definition, making immediate use of our new notation.

Definition 6.7. If $G = (V_1, E_1)$ and $H = (V_2, E_2)$ are graphs, we define the *strong product* of the graphs, denoted $G \boxtimes H = (V, E)$, as

the graph with vertex set $V := V_1 \times V_2$ and edge set

$$E := \{(x_1, x_2) \sim (x_1', x_2') \mid x_1 \simeq_G x_1',\ x_2 \simeq_H x_2',\ (x_1, x_2) \neq (x_1', x_2')\}.$$

Remark 6.8. Our intent is to use this notation and vocabulary to describe confusability graphs. As such, it is important to note that our definition of the strong product takes care of avoiding loops (no edge from a vertex to itself) since $(a, b) \simeq (a, b)$ by definition.

Proposition 6.9. *If we assume errors act independently and we use the noisy channel N to send messages of length 2 (we call this channel "product channel" $N \times N$), then $N \times N$ has confusability graph $G \boxtimes G$ and hence the one-shot zero error capacity of $N \times N$ is $\alpha(G \boxtimes G)$, where G is the confusability graph of N.*

In the same way, we can talk about the *Shannon capacity* of a Channel N. If we use an alphabet to make words of length n, then the one shot zero error capacity of $N \times \cdots \times N$ (n times) is

$$\alpha(\underbrace{G \boxtimes \cdots \boxtimes G}_{n \text{ times}}) = \alpha(G^{\boxtimes n}).$$

6.2. **Shannon Capacity.** We would like this to behave like Θ^n for some Θ. That is, we would like

$$\alpha(G^{\boxtimes n}) \sim \Theta^n.$$

We define (for a channel N) the *Shannon capacity* to be

$$\Theta_1(N) = \lim_{n \to \infty} \frac{\log(\alpha(G^{\boxtimes n}))}{n}.$$

Similarly, we define

$$\Theta_2(N) = \lim_{n \to \infty} \sqrt[n]{\alpha(G^{\boxtimes n})}.$$

These are related by the equation $\Theta_1(N) = \log(\Theta_2(N))$.

It is worth mentioning that only the confusability graph G matters. In particular, we may also define

$$\Theta_1(G) = \lim_{n \to \infty} \frac{\log(\alpha(G^{\boxtimes n})}{n},$$

and simiarly

$$\Theta_2(G) = \lim_{n \to \infty} \sqrt[n]{\alpha(G^{\boxtimes n})}.$$

Shannon was able to compute these numbers for the 3-cycle graph, the 4-cycle graph and the complete graph on 4 vertices. He couldn't get a bound for the Shannon capacity on the 5-cycle graph.

We will examine some estimates given by Lovasz. If $A = A^*$ is a self-adjoint matrix, then we define $\lambda_{\max}(A)$ to be the largest eigenvalue of A.
We summarize Lovász's results in [**?**] in this theorem.

Theorem 6.10 (Lovász)**.** *Let G be a graph on n vertices $\{1, \ldots, n\}$ and define*

$$\theta(G) := \inf\{\lambda_{\max}(A) \colon A = [a_{ij}] = A^* \in M_n, a_{ij} = 1, \forall i \not\sim j\}.$$

Then the following holds:

1. $\theta(G) = \sup\{||I + H|| \colon H = [h_{ij}] = H^* \in M_n, h_{ij} = 0, \forall i \sim j \text{ and } h_{ii} = 0, \text{ for all } i, I + H \geq 0\}$.
2. *If H is another graph on m vertices, then $\theta(G \boxtimes H) = \theta(G) \cdot \theta(H)$.*
3. $\alpha(G) \leq \theta(G)$.

Hence, $\alpha(G^{\boxtimes n}) \leq \theta(G^{\boxtimes n}) = \theta(G)^n$. Taking n-th root and letting $n \to \infty$ it is easy to see that

$$\Theta_2(G) = \lim_{n\to\infty} \sqrt[n]{\alpha(G^{\boxtimes n})} \leq \theta(G).$$

To illustrate an application of the chromatic number, let us address another noise problem:
Suppose that Alice sends a letter x_j through the noisy channel N, and Bob receives the letter y_i. Also suppose that Alice has another, "very expensive" channel that is perfect (no errors). (For example, sending every bit a thousand times would dramatically decrease the possibility of being unable to determine what Alice sent.) We wish to find the smallest integer k and a function $f : \{x_1, ..., x_m\} \to \{1, ..., k\}$ (where Alice's alphabet is $\{x_1, ..., x_m\}$) such that if Alice sends

$$(x_j, f(x_j)) \to (y_i, f(x_j)),$$

(first coordinate through N and second coordinate through perfect channel), then Bob is able to retrieve with certainty which x_j was sent. Obviously there is a trivial solution — choose$f(x) = x$ — where Alice can simply send x to Bob using her perfect channel. This needs $k = card(V)$ where V is the set of vertices of the cofusbility graph, because each message in V has a distinct label. But the question here is: What is the smallest k so that the cost is minimized. (The smaller k is, the less number of bits that you need to send through the perfect channel.) We call this smallest number *packing number* of N and denote it by $\chi^*(N)$.

Theorem 6.11. *If $N = (p(y_i|x_j))$ is a noisy channel and G is its confusability graph, then $\chi^*(N) = \chi(G)$, the chromatic number of the graph G.*

Proof. Let $\chi(G) = c$ and $f : \{x_1, ..., x_n\} \to \{1, ..., c\}$ be a colouring such that if $x_i \sim x_j$, then $f(x_i) \neq f(x_j)$. Suppose that Alice sends $(x_j, f(x_j))$ and Bob receives $(y_i, f(x_j))$. Let $S = \{x : p(y_i|x) \neq 0\}$; so Bob knows that $x_j \in S$. By definition of the confusability graph, if $x_p, x_\ell \in S$, then they can be confused by Bob when they are sent, so that $x_p \sim x_\ell$. Hence, each element of S has a unique colour. Hence, with $k = c$, Bob knows which element of S was sent, since he has also received the colour of the input. By definition of $\chi^*(N)$, we must have $\chi^*(N) \leq c$.
Conversely, suppose that $f : \{x_1, ..., x_n\} \to \{1, ..., k\}$ is such that whenever the input/output is

$$(x_j, f(x_j)) \to (y_i, f(x_j)),$$

then we can find which x_j was sent. We claim that f is a colouring of G. Indeed, suppose that $x_p \sim x_\ell$ in G. Then there is y_q such that $x_p \to y_q$ and $x_\ell \to y_q$. Hence, with sending $(x_p, f(x_p))$ and $(x_\ell, f(x_\ell))$ through, the only way that x_p and x_ℓ can be distinguished is by their images under f. Hence, $f(x_\ell) \neq f(x_p)$ so that f is a colouring for G. Thus, $\chi^*(N) \geq \chi(G)$, so that $\chi^*(N) = \chi(G)$.

□

Definition 6.12. The *Witsenhausen rate* of a noisy channel N is

$$R(N) = \lim_{n\to\infty} \frac{\log(\chi^*(N^{(n)}))}{n} = \lim_{n\to\infty} \frac{\log \chi(G^{\boxtimes n})}{n} := R(G).$$

We can define a similar quantity

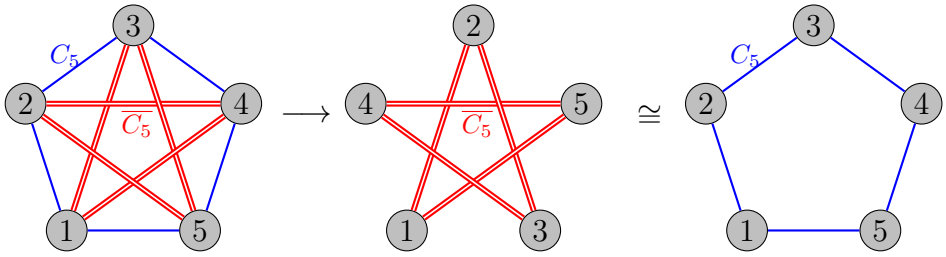
$$r(N) = \lim_{n\to\infty} \sqrt[n]{\chi^*(N^{(n)})} = \lim_{n\to\infty} \sqrt[n]{\chi(G^{\boxtimes n})} := r(G).$$

How do we estimate these quantities? (In general these are very hard to compute.)

Definition 6.13. Given $G = (V, E)$, the *complement of* G, denoted $\overline{G} = (V, \overline{E})$, is given by $(v, w) \in \overline{E} \Leftrightarrow \big(w \neq v \text{ and } (v, w) \notin E\big)$.

Geometrically, the complement is what you might expect–if a graph is missing a possible edge, that edge is in the complement, and every edge in a graph is missing from the complement.

Example 6.14 (Complement of C_5)**.** In doubled red lines, we see $\overline{C_5}$ "complementing" C_5 to make a complete graph.



By renumbering the vertices of $\overline{C_5}$, we see that, interestingly enough, $\overline{C_5}$ is a cycle, and thus C_5 is self-complementary.

Theorem 6.15. (Lovasz Sandwich Theorem) *Let G be a graph.*

(1) $\alpha(G) \leq \theta(G) \leq \chi(\overline{G})$.
(2) $\theta(\overline{G^{\boxtimes n}}) = \theta(\overline{G})^n$.

In particular, $\chi(G^{\boxtimes n}) \geq \theta(\overline{G^{\boxtimes n}}) = \theta(\overline{G})^n$, *and* $r(G) = \lim_{n\to\infty} \sqrt[n]{\chi(G^{\boxtimes n})} \geq \theta(\overline{G})$.

Our aim is to use these general ideas in the quantum world (with modifications, since the quantum model is different from the classical model).

7. One Shot Zero Error Capacity of Quantum Channels

The following material is based on two papers: [**?**] (by Duan-Severini-Winter) and [**?**](by Jop Briet, Harry Buhrman, Monique Laurent, Teresa Piovesan, and Giannicola Scarpa et al).

Now, we assume that $\mathcal{N}$ is a quantum channel; i.e. a CPTP map $\mathcal{N} : M_n \to M_m$ with $\mathcal{N}(X) = \sum_i E_i X E_i^*$, where $\sum_i E_i^* E_i = I$. To model a notion of one-shot zero error capacity, we want to find the maximum number of perfectly distinguishable states $v_1, ..., v_d \in \mathbb{C}^n$ such that the outputs $\mathcal{N}(v_i v_i^*)$ are perfectly distinguishable. Recall that if $\{W_1, ..., W_k\}$ is a measurement system and P is the density matrix of some ensemble, then $p_i = \mathrm{Tr}(W_i P W_i^*)$ is the probability of getting outcome i.

Definition 7.1. We say that $\{P_1, ..., P_d\} \subseteq M_n$ are *perfectly distinguishable* if there exists a measurement system $\{W_1, ..., W_k\}$ with $k \geq d$ such that $\mathrm{Tr}(W_i P_j W_i^*) = \delta_{ij}$ for all $1 \leq i, j \leq d$.

Lemma 7.2. *If $P \in M_m$, with $P \geq 0$ and $Tr(P) = 0$, then $P = 0$.*

Proof. If $\lambda_1, ..., \lambda_m$ are the eigenvalues of P, then since $P \geq 0$, each $\lambda_i \geq 0$. But $0 = \mathrm{Tr}(P) = \sum_{i=1}^{m} \lambda_i = 0$ so each $\lambda_i = 0$. Since P is diagonalizable with entries equal to the eigenvalues, we have $P = 0$. $\square$

Note that $M_m \cong \mathbb{C}^{m^2}$, so it is a Hilbert space equipped with the inner product

$$\langle Y, X \rangle = \sum_{i,j} \bar{y}_{i,j} x_{i,j} = Tr(Y^* X)\,.$$

Lemma 7.3. *Let $P, Q \in M_m$ with $P \geq 0$ and $Q \geq 0$. The following are equivalent.*

(1) $\langle P, Q \rangle = 0$.
(2) $PQ = 0$.
(3) *ran* $(P) \perp$ *ran* (Q).

Proof. $(3 \implies 2)$:
Note that ran $(Q) \subseteq$ ran $(P)^{\perp} = \ker(P^*) = \ker(P)$. Hence if $x \in \mathbb{C}^m$, we have $PQx = 0$, since $Qx \in \ker(P)$. Therefore, $PQ = 0$.
$(2 \implies 1)$:
$\langle P, Q \rangle = \mathrm{Tr}(P^*Q) = \mathrm{Tr}(PQ) = \mathrm{Tr}(0) = 0$.
$(1 \implies 3)$:
Because $P \geq 0$, there is a unitary U such that $U^*PU = D := \mathrm{diag}(d_1, ..., d_k, 0, 0, ..., 0)$, where each $d_k > 0$. Then let $\widetilde{Q} = U^*QU$. Note that

$$\begin{aligned} \langle D, \widetilde{Q} \rangle &= \mathrm{Tr}(D\widetilde{Q}) \\ &= \mathrm{Tr}((U^*PU)(U^*QU)) \\ &= \mathrm{Tr}(PQUU^*) \\ &= \mathrm{Tr}(PQ) = \langle P, Q \rangle = 0. \end{aligned}$$

Moreover, we have

$$0 = \langle D, \widetilde{Q} \rangle = \mathrm{Tr}(D\widetilde{Q}) = \sum_{i=1}^{k} d_i \widetilde{q}_{ii}.$$

Hence, each $\widetilde{q}_{ii} = 0$ for $1 \leq i \leq k$. We can write

$$\widetilde{Q} = \begin{pmatrix} \widetilde{Q}_{11} & \widetilde{Q}_{12} \\ \widetilde{Q}_{21} & \widetilde{Q}_{22} \end{pmatrix},$$

where $\widetilde{Q}_{11} \in M_k$. Since $\widetilde{Q} \geq 0$, we also have $\widetilde{Q}_{11} \geq 0$. But $\mathrm{Tr}(\widetilde{Q}_{11}) = 0$ so $\widetilde{Q}_{11} = 0$. Since $\widetilde{Q} \geq 0$, we can write

$$\widetilde{Q} = \begin{pmatrix} 0 & \widetilde{Q}_{12} \\ (\widetilde{Q}_{12})^* & \widetilde{Q}_{22} \end{pmatrix}.$$

We now compute, for $v \in \mathbb{C}^k$ and $h \in \mathbb{C}^{n-k}$,

$$\left\langle \begin{pmatrix} v \\ h \end{pmatrix}, \widetilde{Q} \begin{pmatrix} v \\ h \end{pmatrix} \right\rangle = \left\langle \begin{pmatrix} v \\ h \end{pmatrix}, \begin{pmatrix} \widetilde{Q}_{12} h \\ (\widetilde{Q}_{12})^* v + \widetilde{Q}_{22} v \end{pmatrix} \right\rangle$$
$$= \langle v, \widetilde{Q}_{12} h \rangle + \langle h, (\widetilde{Q}_{12})^* v + \widetilde{Q}_{22} h \rangle$$

Pick $v = -r\widetilde{Q}_{12}h$ for some $r > 0$. The inner product then becomes

$$\left\langle \begin{pmatrix} v \\ h \end{pmatrix}, \widetilde{Q} \begin{pmatrix} v \\ h \end{pmatrix} \right\rangle = -r\|\widetilde{Q}_{12}h\|^2 - r\langle h, (\widetilde{Q}_{12})^* \widetilde{Q}_{12} h \rangle + \langle h, \widetilde{Q}_{21} h \rangle.$$

Let $r \to \infty$. Since this inner product must always be positive, we must have $\|\widetilde{Q}_{12}h\|^2 = 0$ so that $\widetilde{Q}_{12} = 0$. Hence, $\widetilde{Q} = \begin{pmatrix} 0 & 0 \\ 0 & \widetilde{Q}_{22} \end{pmatrix}$. Hence, ran $(D) \subseteq$ span $\{e_1, ..., e_k\} \perp$ span $\{e_{k+1}, ..., e_m\} \supseteq$ ran $(\widetilde{Q})$. Thus, ran $(D) \perp$ ran $(\widetilde{Q})$. Now, since $U^*PU = D$ and $U^*QU = \widetilde{Q}$, we have

$$U^*\text{ran }(P) = \text{ran }(U^*PU) \perp \text{ran }(U^*QU) = U^*\text{ran }(Q).$$

Since unitaries preserve inner products, we have ran $(P) \perp$ ran (Q). □

Proposition 7.4. *Let $\{P_1, \ldots, P_d\} \subseteq M_m$ be density matrices. Then $\{P_1, \ldots, P_d\}$ are perfectly distinguishable if and only if $P_iP_j = 0$ for all $i \neq j$.*

Proof. ($\Longleftarrow$)**:**
By the above lemma, $P_iP_j = 0 \implies$ ran $(P_i) \perp$ ran (P_j). Let W_i be the orthogonal projection onto ran (P_i). Then $\sum_{i=1}^{d} W_i$ is a projection, so $W_{d+1} = I - \sum_{i=1}^{d} W_i$ is a projection. Then $\sum_{i=1}^{d+1} W_i^* W_i = \sum_{i=1}^{d+1} W_i = I$, so $\{W_1, ..., W_{d+1}\}$ is a measurement system. Note that

$$\text{Tr}(W_iP_jW_i^*) = \text{Tr}(W_iP_j) = \begin{cases} 0 \text{ if } i \neq j \\ \text{Tr}(P_j) = 1 \text{ if } i = j. \end{cases}$$

($\Longrightarrow$)**:**
Suppose that $\{P_1, ..., P_d\}$ is perfectly distinguishable, with $\{W_1, ..., W_k\}$ a measurement system such that $\text{Tr}(W_iP_jW_i^*) = \delta_{ij}$. If $i \neq j$, then $\text{Tr}(W_iP_jW_i^*) = 0 = \text{Tr}(P_jW_i^*W_i)$. Hence, $P_jW_i^*W_i = 0$ by the lemma. We also have $\text{Tr}(W_iP_iW_i^*) = 1 = \text{Tr}(P_j) = \text{Tr}(P_jW_i^*W_i)$, since P_j is a density matrix. Hence, $0 = \text{Tr}(P_i(I - W_i^*W_i))$, but $I - W_i^*W_i \geq 0$ and $P_i \geq 0$. By the lemma,

$P_i(I - W_i^*W_i) = 0$, so that $P_i = P_iW_i^*W_i$. Hence, for $i \neq j$, we can write

$$P_iP_j = P_i(W_i^*W_iP_j) = P_i \cdot 0 = 0,$$

as desired.

□

Remark 7.5. If $v_1, ..., v_n$ are states, then $v_i \perp v_j$ for $i \neq j$ if and only if $(v_jv_j^*)(v_iv_i^*) = 0$. Indeed, we see that $v_j(v_j^*v_i)v_i^* = v_j(\langle v_j|v_i\rangle)v_i^* = \langle v_j|v_i\rangle v_jv_i^*$.

Corollary 7.6. *Let $\mathcal{N} : M_n \to M_m$ be a CPTP map. Then*

$$\alpha(\mathcal{N}) = \max\{d : \exists v_1, ..., v_d \in \mathbb{C}^n \text{ orthonormal with } \mathcal{N}(v_iv_i^*)\mathcal{N}(v_jv_j^*) = 0, \forall i \neq j\}.$$

The maximum number d of vectors such that $\eta(v_iv_i^*) \cdot \eta(v_jv_j^*) = 0$ for all $i \neq j$ is called the **one-shot zero-error capacity** of η, and is denoted by $\alpha(\eta)$.

Proposition 7.7. *Let $\mathcal{N} : M_n \to M_m$ be a CPTP map. Write $\mathcal{N}(X) = \sum_{i=1}^K E_iXE_i^*$, and let $\{v_1, ..., v_d\} \subseteq \mathbb{C}^n$ be orthonormal. For $i \neq j$, we have $\mathcal{N}(v_iv_i^*) \cdot \mathcal{N}(v_jv_j^*) = 0$ if and only if $E_kv_i \perp E_\ell v_j$ for all k, ℓ.*

Proof. Since each $\mathcal{N}(v_iv_i^*)$ is positive, we note that $\mathcal{N}(v_iv_i^*)\mathcal{N}(v_jv_j^*) = 0$ if and only if $\mathrm{Tr}(\mathcal{N}(v_iv_i^*)\mathcal{N}(v_jv_j^*)) = 0$. Substituting the expression for $\mathcal{N}$ gives

$$0 = \mathrm{Tr}\left(\left(\sum_{\ell=1}^K E_\ell v_iv_i^*E_\ell^*\right)\left(\sum_{k=1}^K E_kv_jv_j^*E_k^*\right)\right).$$

Now, whenever $P \geq 0$ and $Q \geq 0$, one has $\mathrm{Tr}(PQ) \geq 0$ (just diagonalize P). The above expression forces $\mathrm{Tr}((E_\ell v_iv_i^*E_\ell^*)(E_kv_jv_j^*E_k^*)) = 0$ for all k, ℓ. Hence, $\mathcal{N}(v_iv_i^*)\mathcal{N}(v_jv_j^*) = 0$ if and only if

$$\mathrm{Tr}(E_\ell v_i(v_i^*E_\ell^*E_kv_j)v_j^*E_k^*) = 0, \forall k, \ell.$$

We can write

$$\begin{aligned}
\mathrm{Tr}(E_\ell v_i(v_i^*E_\ell^*E_kv_j)v_j^*E_k^*) &= \mathrm{Tr}(E_\ell v_i(E_\ell v_i)^*(E_kv_j)v_j^*E_k^*) \\
&= \langle E_\ell v_i|E_kv_j\rangle \mathrm{Tr}(E_\ell v_iv_j^*E_k^*) \\
&= \langle E_\ell v_i|E_kv_j\rangle \mathrm{Tr}(v_j^*E_k^*E_\ell v_i) \\
&= \langle E_\ell v_i|E_kv_j\rangle\langle E_kv_j|E_\ell v_i\rangle \\
&= |\langle E_\ell v_i|E_kv_j\rangle|^2.
\end{aligned}$$

It follows that $\mathcal{N}(v_iv_i^*)\mathcal{N}(v_jv_j^*) = 0$ if and only if $E_\ell v_i \perp E_kv_j$.

□

Next, we need to show that the one-shot zero-error capacity of $\mathcal{N}$ is independent of any particular Choi-Kraus representation of $\mathcal{N}$. In order to do this we first need the notion of an operator system.

Definition 7.8 (Operator System)**.** A subspace $\mathcal{S} \subset B(\mathcal{H})$ is called an operator system provided that:

(1) $I \in \mathcal{S}$
(2) If $X \in \mathcal{S}$, then $X^* \in \mathcal{S}$

Proposition 7.9. *Let $\mathcal{N} : M_n \to M_m$ be a CPTP map defined by $\mathcal{N}(X) = \sum_{i=1}^K E_i X E_i^* = \sum_{\ell=1}^r Y_\ell X Y_\ell^*$. Then*

$$span\{E_i^* E_j : 1 \leq i, j \leq K\} = span\{Y_\ell^* Y_k : 1 \leq \ell, k \leq r\} \subseteq M_n$$

and this is an operator system.

Proof. From **??** we already know that $\text{span}\{E_i : 1 \leq i \leq K\} = \text{span}\ \{Y_\ell : 1 \leq \ell \leq r\}$. Hence, we may write $E_i = \sum_{\ell=1}^r \alpha_{i\ell} Y_\ell$. It follows that

$$E_i^* E_j = \sum_{\ell,k=1}^r \overline{\alpha}_{i\ell}\alpha_{jk} Y_\ell^* Y_k \in \text{span}\ \{Y_\ell^* Y_k : 1 \leq \ell, k \leq r\}.$$

Similarly, one can show that $Y_k^* Y_\ell \in \text{span}\{E_i^* E_j : 1 \leq i, j \leq k\}$.
Since $\mathcal{N}$ is trace-preserving, we must have $I = \sum_{i=1}^K E_i^* E_i \in \text{span}\{E_i^* E_j\}_{i,j=1}^K$.
Since the adjoint is conjugate-linear, we need only check that $(E_i^* E_j)^* \in \text{span}\{E_i^* E_j\}_{i,j=1}^K$ for each choice of i, j. But this is immediate since $(E_i^* E_j)^* = E_j^* E_i$. Hence, $\text{span}\{E_i^* E_j\}_{i,j=1}^K$ is an operator system. □

Definition 7.10. Given a CPTP map $\mathcal{N} : M_n \to M_m$ with $\mathcal{N}(X) = \sum_{i=1}^K E_i X E_i^*$. Then the *operator system of* $\mathcal{N}$ is the space

$$\mathcal{S}_\mathcal{N} = \text{span}\{E_i^* E_j : 1 \leq i, j \leq K\}.$$

Remark 7.11. Notice that because of **??**, this operator system is independent of the particular Choi-Kraus representation of $\mathcal{N}$.

Definition 7.12. Given an operator system $\mathcal{S} \subseteq M_n$, define

$$\alpha(\mathcal{S}) = \max\{d \in \mathbb{N} : \exists v_1, ..., v_d \in \mathbb{C}^n \text{ orthonormal with } v_i v_j^* \perp \mathcal{S}, \forall i \neq j\}.$$

(Here, we mean that $\text{Tr}(X^* v_i v_j^*) = 0$ for all $X \in \mathcal{S}$.) We call $\alpha(\mathcal{S})$ the *independence number* or the *one-shot capacity* of $\mathcal{S}$.

Proposition 7.13. *Let $\mathcal{N} : M_n \to M_m$ be CPTP, and write $\mathcal{N}(X) = \sum_{i=1}^K E_i X E_i^*$. Let $\{v_1, ..., v_d\} \subseteq \mathbb{C}^n$ be an orthonormal set. Then $\mathcal{N}(v_i v_i^*)\mathcal{N}(v_j v_j^*) = 0$ for all $i \neq j$ if and only if $v_i v_j^* \perp \mathcal{S}_\mathcal{N}$ for all $i \neq j$.*

Proof. Let $i, j \in \{1, ..., d\}$. One has $v_i v_j^* \perp \mathcal{S}_\eta$ if and only if $v_i v_j^* \perp E_k^* E_\ell$ for all k, ℓ, which occurs if and only if $\mathrm{Tr}((v_i v_j^*)^* E_k^* E_\ell) = 0$ for all k, ℓ. One may rewrite this as

$$\mathrm{Tr}(v_j v_i^* E_k^* E_\ell) = \mathrm{Tr}((E_k v_i)^* (E_\ell v_j)),$$

so that $v_i v_j^* \perp \mathcal{S}_\eta$ if and only if $E_k v_i \perp E_\ell v_j$ for all k, ℓ, and this occurs if and only if $\mathcal{N}(v_i v_i^*)\mathcal{N}(v_j v_j^*) = 0$, by the proposition.

□

Theorem 7.14 ([?])**.** *Let $\eta : M_n \to M_m$ be a CPTP map, then*

$$\alpha(\mathcal{S}_\mathcal{N}) = \alpha(\mathcal{N})$$

Corollary 7.15. *Let $\mathcal{N}_i : M_n \to M_{m_i}$ be CPTP maps for $i = 1, 2$. If $\mathcal{S}_{\mathcal{N}_1} = \mathcal{S}_{\mathcal{N}_2}$, then $\alpha(\mathcal{N}_1) = \alpha(\mathcal{N}_2)$.*

Remark 7.16. The quantity $\alpha(\mathcal{S})$ really depends on the containment $\mathcal{S} \subseteq M_n$ and is not merely an instrinsic property of $\mathcal{S}$. For example, let $\mathcal{S} \subseteq M_n$ be an operator system, and define

$$\widetilde{\mathcal{S}} = \left\{ \begin{pmatrix} X & 0 \\ 0 & X \end{pmatrix} : X \in \mathcal{S} \right\} = I_2 \otimes \mathcal{S} \subseteq M_{2n}.$$

Clearly $\widetilde{\mathcal{S}}$ is an operator system. The map $x \mapsto \begin{pmatrix} x & 0 \\ 0 & x \end{pmatrix}$ markes $\mathcal{S}$ and $\tilde{\mathcal{S}}$ the "same." But if $\alpha(\mathcal{S}) = d$, then we can choose $v_1, ..., v_d \in \mathbb{C}^n$ orthonormal such that $v_i v_j^* \perp \mathcal{S}$ for all $i \neq j$. Consider the set

$$\{e_1 \otimes v_1, ..., e_1 \otimes v_d, e_2 \otimes v_1, ..., e_2 \otimes v_d\} \subseteq \mathbb{C}^2 \otimes \mathbb{C}^n \simeq \mathbb{C}^{2n}.$$

These are of the form

$$(v_1, 0), ..., (v_d, 0), (0, v_1), ..., (0, v_d).$$

One computes

$$(e_1 \otimes v_i)(e_1 \otimes v_j)^* = \begin{pmatrix} v_i v_j^* & 0 \\ 0 & 0 \end{pmatrix}$$

$$(e_1 \otimes v_i)(e_2 \otimes v_j)^* = \begin{pmatrix} 0 & v_i v_j^* \\ 0 & 0 \end{pmatrix}$$

$$(e_2 \otimes v_i)(e_1 \otimes v_j)^* = \begin{pmatrix} 0 & 0 \\ v_i v_j^* & 0 \end{pmatrix}$$

$$(e_2 \otimes v_i)(e_2 \otimes v_j)^* = \begin{pmatrix} 0 & 0 \\ 0 & v_i v_j^* \end{pmatrix}.$$

It follows that $\alpha(\widetilde{\mathcal{S}}) \geq 2\alpha(\mathcal{S})$, so that the independence number really depends on the embedding of the operator system in the space of matrices.

Remark 7.17. One can check that whenever $\mathcal{S}_1 \subseteq \mathcal{S}_2$ as operator systems with $\mathcal{S}_1, \mathcal{S}_2 \subseteq M_n$, then $\alpha(\mathcal{S}_1) \geq \alpha(\mathcal{S}_2)$.

Proposition 7.18. *Let $\mathcal{S} \subseteq M_n$ be an operator system. Then there is $m \in \mathbb{N}$ and a CPTP map $\mathcal{N} : M_n \to M_m$ such that $\mathcal{S} = \mathcal{S}_{\mathcal{N}}$.*

Proof. Let $\mathcal{S} = \text{span}\{X_1, ..., X_k\}$. Let

$$H = \begin{pmatrix} 0 & X_1 & & & \\ X_1^* & 0 & X_2 & & \\ & X_2^* & 0 & & \\ & & \ddots & \ddots & X_k \\ & & & X_k^* & 0 \end{pmatrix} \in M_{k+1}(M_n) = M_{n(k+1)}.$$

Observe that $H = H^*$. Pick $r > 0$ with $rI + H \geq 0$. Then let

$$P = (P_{ij}) = \frac{1}{r(k+1)}(rI + H) \geq 0.$$

One can factor $P = (P_{ij}) = C^*C$ where $C = (C_{ij}) \in M_{k+1}(M_n) = M_{n(k+1)}$. Write

$$C = \left[\overbrace{C_1}^{n} \ \vdots \ \cdots \ \vdots \ \overbrace{C_{k+1}}^{n} \right],$$

where each C_ℓ is of size $n(k+1) \times n$. Define $\Phi : M_n \to M_{n(k+1)}$ by

$$\Phi(X) = \sum_{\ell=1}^{k+1} C_\ell X C_\ell^*.$$

One can see that Φ is completely positive. To see that Φ is trace-preserving, one notes that $C_p^* C_q = P_{p,q}$ for all $1 \leq p, q \leq k+1$. Then $C_p^* C_p = P_{p,p}$ for each p, while $P_{p,p} = \frac{1}{r(k+1)}(rI_n) = \frac{1}{k+1} I_n$. Since $\sum_{p=1}^{k+1} P_{p,p} = I_n$, we see that $\sum_{\ell=1}^{k+1} C_\ell^* C_\ell = I_n$, so Φ is trace-preserving.

Since Φ is CPTP, its corresponding operator system is

$$\mathcal{S}_\Phi = \text{span}\{C_p^* C_q : 1 \leq p, q \leq k+1\}.$$

But one has $C_p^* C_q = P_{p,q}$, so

$$\mathcal{S}_\Phi = \text{span}\{P_{ij} : 1 \leq i, j \leq k+1\} = \text{span}\{I, X_1, ..., X_k, X_1^*, ..., X_k^*\} = \mathcal{S}.$$

□

7.1. **Analogue of Shannon Capacity.** We now develop a quantum analogue of the Shannon capacity. Using the channel $\mathcal{N} : M_n \to M_m$ k times corresponds to the k-fold tensor product

$$\underbrace{\mathcal{N} \otimes \cdots \otimes \mathcal{N}}_{k \text{ times}} : M_n^{\otimes k} \to M_m^{\otimes k},$$

or $\mathcal{N}^{\otimes k} : M_{n^k} \to M_{m^k}$. In this way, we may define

$$\Theta_2(\mathcal{N}) = \limsup_{k\to\infty} \sqrt[k]{\alpha(\mathcal{N}^{\otimes k})}.$$

It is interesting to see what happens from the operator system viewpoint. If $\mathcal{N}(X) = \sum_{i=1}^K E_i X E_i^*$, then

$$\eta\otimes\mathcal{N}(X\otimes Y) = \mathcal{N}(X)\otimes\mathcal{N}(Y) = \left(\sum_{i=1}^K E_i X E_i^*\right)\otimes\left(\sum_{j=1}^K E_j Y E_j^*\right) = \sum_{i,j=1}^K (E_i\otimes E_j)(X\otimes Y)(E_i\otimes E_j)$$

Therefore, we see that

$$\mathcal{S}_{\mathcal{N}\otimes\mathcal{N}} = \text{span}\{(E_i\otimes E_j)^*(E_k\otimes E_\ell)\}_{i,j,k,\ell=1}^K = \text{span}\{E_i^* E_k\otimes E_j^* E_\ell\}_{i,j,k,\ell=1}^K = \mathcal{S}_\mathcal{N}\otimes\mathcal{S}_\mathcal{N}.$$

We conclude that

$$\mathcal{S}_{\mathcal{N}^{\otimes k}} = \underbrace{\mathcal{S}_\mathcal{N} \otimes \cdots \otimes \mathcal{S}_\mathcal{N}}_{k \text{ times}} \subseteq M_{n^k}.$$

Hence, it follows that $\alpha(\mathcal{N}^{\otimes k}) = \alpha(\mathcal{S}_\mathcal{N}^{\otimes k})$.

Definition 7.19. Given an operator system $\mathcal{S} \subseteq M_n$. Then the **the Shannon Capacity of** $\mathcal{S}$ is defined by

$$\Theta_2(\mathcal{S}) = \limsup_{k\to\infty} \sqrt[k]{\alpha(\mathcal{S}^{\otimes k})}.$$

From the remarks above we have that $\Theta_2(\mathcal{N}) = \Theta_2(\mathcal{S}_\mathcal{N})$.

7.2. **Operator Systems of Graphs.** Let $G = (V, E)$ be a graph on n vertices and let $V = \{1, ..., n\}$. The *operator system of* G is the subspace of M_n defined by

$$\mathcal{S}_G := \text{span}(\{E_{11}, ..., E_{nn}\} \cup \{E_{ij} : (i, j) \in E\}) \subseteq M_n.$$

Note that $I = \sum_{i=1}^n E_i i$ and $\mathcal{S}_G$ is self-adjoint by the symmetry of the set $\{E_{11}, ..., E_{nn}\} \cup \{E_{ij} : (i, j) \in E\}$, hence $\mathcal{S}_G$ is indeed an operator system.

Definition 7.20 (Hamming Distance)**.** Let $a = (a_1, ..., a_n) \in \mathbb{Z}_2^n$ and $b = (b_1, ..., b_n) \in \mathbb{Z}_2^n$ be binary n-tuples. The *Hamming distance* between a and b is defined as

$$d_H(a, b) := |\{i : a_i \neq b_i\}| = \sum_{i=1}^n |a_i - b_i|.$$

Problem 7.21. *Let $X = \mathbb{Z}_2^3$ and $Y = \mathbb{Z}_2^9$. Consider the encoding given by*

$$x = (1,0,1) \xrightarrow{e} (111,000,111) \xrightarrow[0,1,\ or\ 2\ bit\ flips]{noise} y$$

(so the only events of non-zero probability are zero, one or two bit flips). Formally, we define

$$p(y|e(x)) = \begin{cases} 0 \text{ if } d_H(y,e(x)) > 2 \\ \text{non-zero if } d_H(y,e(x)) \leq 2. \end{cases}$$

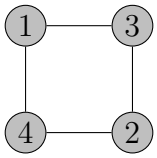
Find the confusability graph G. Show that $x \simeq x'$ (adjacent to or equal) if and only if $d_H(x,x') \leq 1$. Find $\alpha(N) = \alpha(G)$ and $\chi^*(N) = \chi(G)$ (just give a convincing argument).

Proof. **Homework problem 9**; due 11th February, Thursday. □

Problem 7.22. *Let* $\mathcal{S} = \left\{ \begin{pmatrix} a & b \\ b & a \end{pmatrix} : a, b \in \mathbb{C} \right\} \subseteq M_2$. *Prove that* $\alpha(\mathcal{S}) = 2$.

Proof. **Homework problem 10**; due 11th February, Thursday. □

Example 7.23. Let G be the four-cycle graph:



G

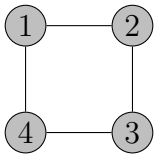
Then we have,

$$\mathcal{S}_G = \text{span}\{E_{11}, E_{22}, E_{33}, E_{44}, E_{13}, E_{31}, E_{14}, E_{41}, E_{23}, E_{32}, E_{24}, E_{42}\},$$

that is,

$$\mathcal{S}_G = \left\{ \begin{pmatrix} a & 0 & b & c \\ 0 & d & e & f \\ g & h & k & 0 \\ \ell & m & 0 & n \end{pmatrix} : a, ..., n \in \mathbb{C} \right\}.$$

But this depends on the labelling, because if we have it labelled as G':



G'

Then we have,

$$\mathcal{S}'_G = \text{span}\{E_{11}, E_{22}, E_{33}, E_{44}, E_{12}, E_{21}, E_{14}, E_{41}, E_{23}, E_{32}, E_{34}, E_{43}\},$$

that is,

$$\mathcal{S}'_G = \left\{ \begin{pmatrix} a & b & 0 & c \\ d & e & f & 0 \\ 0 & g & h & k \\ \ell & 0 & m & n \end{pmatrix} : a, ..., n \in \mathbb{C} \right\}.$$

However this is just a permutation: $\begin{matrix} 1 \mapsto 1 \\ 2 \mapsto 3 \\ 3 \mapsto 2 \\ 4 \mapsto 4 \end{matrix}$, $U = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$.

So that $U\mathcal{S}_G U^* = \mathcal{S}'_G$.

7.3. **Including Classical in Quantum.** Suppose that our alphabets are $X = \{x_1, ..., x_n\}$ and $Y = \{y_1, ..., y_m\}$, with noisy channel N (given by a matrix of probabilities as before) by which we send letters from X and recieve letters from Y. We can think of $x_1, ..., x_n$ as pure states. In this context, a *mixed state* would be of the form $(p_1, ..., p_m)$ where each p_i is the probability of being in state x_i, $p_i \geq 0$ and $p_1 + \cdots + p_m = 1$. This is the classical analogue of a density matrix.

The analogue of M_n is $\mathbb{C}^n \simeq \ell^\infty(X) = \ell^\infty_n$ (functions on n points), with basis given by the elements

$$\delta_x(x') = \begin{cases} 1 \text{ if } x' = x \\ 0 \text{ if } x' \neq x. \end{cases}$$

(We can also think of δ_x as e_i.) Then N defines a linear map from $\ell^\infty_n \to \ell^\infty_m$ (by sending letters from X to letters in Y). Indeed, $N : \ell^\infty_n \to \ell^\infty_m$ acts via matrix multiplication to give

$$N \begin{pmatrix} p_1 \\ \vdots \\ p_n \end{pmatrix} = \begin{pmatrix} q_1 \\ \vdots \\ q_m \end{pmatrix}.$$

Since the p_i's add to 1 and each of the column sums is 1, it follows that each $q_i \geq 0$ and $q_1 + \cdots + q_m = 1$. Hence, mixed states get sent to mixed states.

We identify ℓ^∞_n as an algebra with pointwise multiplication. In this way, we can identify ℓ^∞_n with $\mathcal{D}_n \subseteq M_n$, the algebra of diagonal matrices, by sending $\delta_{x_i} \mapsto E_{ii}$. We define, for each $1 \leq i \leq m$ and $1 \leq j \leq m$, the matrix $B_{ji} = \sqrt{p(y_j|x_i)} E_{ji} \in M_{m,n}$. Define a map

$\mathcal{N} : M_n \to M_m$ by

$$\mathcal{N}(X) = \sum_{i,j} B_{ji} X B_{ji}^*.$$

Then

$$\mathcal{N}(E_{ii}) = \sum_{\ell,j} p(y_j|x_\ell) E_{j\ell} E_{ii} E_{\ell j} = \sum_j p(y_j|x_i) E_{jj}.$$

If $i \neq j$, then none of the products of the matrix units in the expression for $\mathcal{N}(E_{ij})$ will remain, so that $\mathcal{N}(E_{ij}) = 0$. Now, by construction, $\mathcal{N}$ is a CP map. To see that it is also TP, we need to check that $\sum_{i,j} B_{ji}^* B_{ji} = I$. We compute

$$\begin{aligned}
\sum_{i,j} B_{ji}^* B_{ji} &= \sum_{i,j} p(y_j|x_i) E_{ij} E_{ji} \\
&= \sum_i \left(\sum_j p(y_j|x_i) \right) E_{ii} \\
&= \sum_i E_{ii} = I_n.
\end{aligned}$$

Therefore, η is CPTP. The associated operator system with $\mathcal{N}$ is given by

$$\begin{aligned}
\mathcal{S}_{\mathcal{N}} &= \text{span}\{B_{k\ell}^* B_{ji}\} \\
&= \text{span}\{\sqrt{p(y_k|x_\ell)p(y_j|x_i)} E_{\ell k} E_{ji}\} \\
&= \text{span}\{\sqrt{p(y_j|x_\ell)p(y_j|x_i)} E_{\ell i}\} \\
&= \text{span}(\{E_{ii}\}_i \cup \{E_{\ell i} : \exists j \text{ with } p(y_j|x_\ell)p(y_j|x_i) \neq 0\}). \\
&= \text{span}(\{E_{ii}\} \cup \{E_{\ell i} : x_\ell \sim x_i\}). \\
&= \mathcal{S}_G.
\end{aligned}$$

Theorem 7.24. *Let G be a graph on n vertices. Let $\mathcal{S}_G \subseteq M_n$ be the operator system of the graph. Then $\alpha(\mathcal{S}_G) = \alpha(G)$.*

Proof. First we show that $\alpha(G) \leq \alpha(\mathcal{S}_G)$. If $\alpha(G) = k$ and $\{i_1, ..., i_k\}$ is a set of independent vertices, then $i_a \not\sim i_b$ for $a \neq b$. Then $E_{i_a, i_b} \notin \mathcal{S}_G$, so that $E_{i_a,i_b} = e_{i_a} e_{i_b}^* \perp \mathcal{S}_G$. This is for all $a \neq b$ in $\{i_1, ..., i_k\}$, so that $k \leq \alpha(\mathcal{S}_G)$.

To see the other inequality, we first note the following: for a given $v = (v_1, ..., v_n)^t \in \mathbb{C}^n$, let $\text{supp}(v) = \{i : v_i \neq 0\}$. If $v, w \in \mathbb{C}^n$ are such that $vw^* = (v_i \overline{w}_j) \perp \mathcal{S}_G$, then $(v_i \overline{w}_j) \perp E_{kk}$ for all k, so that $v_k \overline{w}_k = 0$

for all k. Hence, $\text{supp}(v) \cap \text{supp}(w) = \emptyset$. Since $(v_i \overline{w}_j) \perp E_{k\ell}$ whenever $k \sim \ell$, we also have $v_k w_\ell = 0$.

Now, suppose that $\alpha(\mathcal{S}_G) = d$; we will show that $\alpha(G) \geq d$. Let $v_1, ..., v_d \in \mathbb{C}^n$ be an orthonormal set such that $v_i v_j^* \perp \mathcal{S}_G$. Then for $i \neq j$, we have $\text{supp}(v_i) \cap \text{supp}(v_j) = \emptyset$. Pick $i_k \in \text{supp}(v_k)$. These are distinct elements $\{i_1, ..., i_d\} \subseteq \{1, ..., n\}$. Since $v_k v_\ell^* \perp \mathcal{S}_G$ for $k \neq \ell$, we have $v_k v_\ell^*$ as 0 in the (p, q) entry whenever $p \sim q$ in G. Now, $v_k v_\ell^*$ is non-zero in the (i_k, i_ℓ) entry, so that

$$\langle v_k v_\ell^*, E_{i_k, i_\ell} \rangle \neq 0.$$

This shows that $E_{i_k, i_\ell} \notin \mathcal{S}_G$. Equivalently, $i_k \not\sim i_\ell$. Therefore, $\{i_1, ..., i_d\}$ is an independent set in G. Therefore, $d = \alpha(\mathcal{S}_G) \leq \alpha(G)$, and this shows that $\alpha(\mathcal{S}_G) = \alpha(G)$. □

Corollary 7.25. *If $\mathcal{N} : M_n \to M_m$ is a CPTP map and $\mathcal{S}_\mathcal{N} = \mathcal{S}_G$ for a graph G, then $\alpha(\mathcal{N}) = \alpha(G)$.*

We now move towards the analogue of chromatic number for an operator system. In the classical case, we had $x_j \mapsto y_i$, and wanted k extra pieces of information sent through a perfect channel, to eliminate any confusion over what was sent. That is, in the classical case, we want $f : X \to \{1, ..., k\}$ such that knowing y_i and $f(x_j)$ uniquely determines what x_j was sent. We saw that the smallest such k was $\chi(G)$, the chromatic number of the confusability graph G.

In the quantum setting, the perfect channel should be $\text{id}_k : M_k \to M_k$, and the noisy channel is some CPTP map $\mathcal{N} : M_n \to M_m$. Since in the classical case we were sending elements from the cartesian product of X and $f(X)$, here the combined channel becomes $\mathcal{N} \otimes \text{id}_k : M_n \otimes M_k \to M_m \otimes M_k$.

We want an orthonormal basis $v_1, ..., v_n \in \mathbb{C}^n$ and a function $f : \{1, ..., n\} \to \{1, ..., k\}$ such that $\{v_j \otimes e_{f(j)}\}_{j=1}^n$ are perfectly distinguishable when they pass through $\mathcal{N} \otimes \text{id}_k$. The least k for which we can do this is denoted by $\chi^*(\mathcal{N})$.

To motivate the next definition, recall that $\{v_i \otimes e_{f(i)}\}_{i=1}^n$ are perfectly distinguishable after passing through $\mathcal{N} \otimes \text{id}_k$ if and only if $(\mathcal{N} \otimes \text{id}_k)((v_i \otimes e_{f(i)})(v_i \otimes e_{f(i)})^*)$ are mutually orthogonal in $M_m \otimes M_k = M_k(M_m)$. Partition $v_1, ..., v_n$ into the sets $S_j = \{v_i : f(i) = j\}$, for $1 \leq j \leq k$. If $\mathcal{N}(X) = \sum E_\ell X E_\ell^*$, then

$$(\mathcal{N} \otimes \text{id}_k)(X \otimes Y) = \mathcal{N}(X) \otimes Y = \sum_\ell (E_\ell \otimes I_k)(X \otimes Y)(E_\ell \otimes I_k)^*.$$

Since this is true for all simple tensors, we see that for all $Z \in M_n \otimes M_k$,

$$\mathcal{N} \otimes \mathrm{id}_k(Z) = \sum_\ell (E_\ell \otimes I_k) Z (E_\ell \otimes E_k)^*.$$

The associated operator system for $\mathcal{N} \otimes \mathrm{id}_k$ is

$$\mathcal{S}_{\mathcal{N}\otimes \mathrm{id}_k} = \mathrm{span}\{(E_\ell \otimes I_k)^*(E_r \otimes I_k)\}_{\ell,r} = \mathrm{span}\{E_\ell^* E_r \otimes I_k\}_{\ell,r} \subseteq M_n \otimes M_k = M_k(M_n).$$

Recall that the identificatoin $M_n \otimes M_k = M_k(M_n)$ sends

$$E_\ell^* E_r \otimes I_k \mapsto \begin{pmatrix} E_\ell^* E_r & & \\ & \ddots & \\ & & E_\ell^* E_r \end{pmatrix},$$

where this is a block matrix of size k. Then we really have

$$\mathcal{S}_{\mathcal{N}\otimes \mathrm{id}_k} = \left\{ \begin{pmatrix} A & & \\ & \ddots & \\ & & A \end{pmatrix} \in M_k(M_m) : A \in \mathcal{S}_\mathcal{N} \right\} := \mathcal{S}_\mathcal{N}^{(k)}.$$

So $\{v_i \otimes e_{f(i)} : 1 \leq i \leq n\}$ is perfectly distinguishable after applying $\mathcal{N} \otimes \mathrm{id}_k$ if and only if $(v_i \otimes e_{f(i)})(v_j \otimes e_{f(j)})^* \perp \mathcal{S}_{\eta \otimes \mathrm{id}_k}$, while $(v_i \otimes e_{f(i)})(v_j \otimes e_{f(j)})^* = v_i v_j^* \otimes E_{f(i),f(j)}$. We want $v_i v_j^* \otimes E_{f(i),f(j)} \perp \mathcal{S}_\mathcal{N}^{(k)}$, so we consider two cases: either $f(i) \neq f(j)$, in which case we have $v_i v_j^* \otimes E_{f(i),f(j)} \perp \mathcal{S}_\eta^{(k)}$ automatically by the above definition of $\mathcal{S}_\mathcal{N}^{(k)}$, or $f(i) = f(j) = \ell$, in which case $v_i v_j^* \otimes E_{\ell\ell} \perp \mathcal{S}_\mathcal{N}^{(k)}$ if and only if $v_i v_j^* \perp \mathcal{S}_\mathcal{N}$.

We conclude that $\chi^*(\mathcal{N})$ is the least k such that there is an orthonormal basis $v_1, ..., v_n$ of $\mathbb{C}^n$ and a partition of the basis into k subsets $S_1, ..., S_k$ such that when $v_i, v_j \in \mathcal{S}_\ell$, we have $v_i v_j^* \perp \mathcal{S}_\mathcal{N}$.

Definition 7.26. Let $\mathcal{S} \subseteq M_n$ be an operator system. Then the *chromatic number* of $\mathcal{S}$, denoted by $\chi^*(\mathcal{S})$, is the least k such that there is an orthonormal basis $v_1, ..., v_n$ for $\mathbb{C}^n$ and a partition of $\{1, ..., n\}$ into k subsets $S_1, ..., S_k$, such that if $i, j \in S_\ell, i \neq j$, then $v_i v_j^* \perp \mathcal{S}$.

Theorem 7.27. *Let G be a graph on n vertices, and $\mathcal{S}_G \subseteq M_n$ be the corresponding operator system. Then $\chi^*(\mathcal{S}_G) = \chi(G)$.*

Homework problem 11; due 25th February, Thursday. (Hint: this is similar to the proof for $\alpha(G)$ above.)

Here's our proof. First a lemma.

Lemma 7.28. *Let $v_1, ..., v_n \subseteq \mathbb{C}^n$ be a linearly independent set. Then there is a re-ordering of these vectors such that after the re-ordering, the i-th component of v_i is non-zero for all $1 \leq i \leq n$.*

Proof. Let $A = (a_{i,j}) = [v_1 \vdots \cdots \vdots v_n]$ be the matrix whose columns are the vectors $v_1, ..., v_n$, so that $a_{i,j} = \langle e_j, v_i \rangle$. Since these vectors are linearly independent A is invertible, hence $det(A) \neq 0$. But

$$det(A) = \sum_{\sigma} (-1)^{sgn(\sigma)} \prod_{j=1}^{n} a_{\sigma(j),j},$$

and so there must be at least one permutation for which $\prod_{j=1}^{n} a_{\sigma(j),j} \neq 0$.

Re-order the vectors so that $\hat{v}_j = v_{\sigma(j)}$ and the result follows. □

Proof. We now prove the theorem. Assume that $\chi(G) = K$, so that there is a (disjoint) partition of the vertices into K subsets, $S_1 \cup \cdots \cup S_K = \{1, \ldots, n\}$ corresponding to the coloring. If $i, j \in S_l$, then $(i, j) \notin E(G)$ and hence, $e_i e_j^* \perp \mathcal{S}_G$. Thus, we have a partition of the standard o.n.b. $\{e_1, ..., e_n\}$ into K subsets satisfying the conditions in the definition of $\chi^*(\mathcal{S}_G)$ and, $\chi^*(\mathcal{S}_G) \leq \chi(G)$.

Conversely, assume that $\chi^*(\mathcal{S}_G) = K$. Then we have an o.n.b. $\{v_1, ..., v_n\}$ for $\mathbb{C}^n$ and a partition $S_1 \cup \cdots \cup S_K = \{1, ..., n\}$ such that if $i, j \in S_l$ then $v_i v_j^* \perp \mathcal{S}_G$. After re-ordering we may assume that $\langle e_i, v_i \rangle \neq 0$ for all i. This means that the matrix $v_i v_j^*$ is non-zero in the (i, j)-entry. This implies that $E_{i,j} \notin \mathcal{S}_G$. Hence, $(i, j) \notin E(G)$. Thus, if for all l, we color all the vertices in S_l with color l, then whenever two vertices ahve the same color they are not adjacent. Thus, $\chi^*(\mathcal{S}_G) \leq \chi(G)$. □

Problem 7.29. *Let* $\mathcal{S} = span\{I_n, E_{ij} : i \neq j\} \subseteq M_n$, *so that* $\dim(\mathcal{S}) = n^2 - n + 1$. *Prove that* $\chi^*(\mathcal{S}) = n$.

Proof. **Homework problem 12**; due 25th February, Thursday. □

8. Composite Systems, Entanglement and Joint Probabilities

Let $\mathcal{H}_A$ be the state space for Alice's lab and $\{X_k\}_k$, $\sum_k X_k^* X_k = I_{\mathcal{H}_A}$ be a measurement system on $\mathcal{H}_A$ and let $\mathcal{H}_B$ be the state space for Bob's lab and $\{Y_\ell\}_\ell$, $\sum_\ell Y_\ell^* Y_\ell = I_{\mathcal{H}_B}$ be a measurement system on $\mathcal{H}_B$. Suppose $\psi \in \mathcal{H}_A \otimes \mathcal{H}_B$ such that $\|\psi\| = 1$. If $p_A(k)$ and $p_B(\ell)$ respectively denote the probability that Alice gets outcome k in the combined lab and the probability that Bob gets outcome ℓ in the combined lab, then

$$p_A(k) = \|(X_k \otimes I)\psi\|^2 \text{ and } p_B(\ell) = \|(I \otimes Y_\ell)\psi\|^2.$$

If Alice's outcome is k, then the state becomes

$$\frac{(X_k \otimes I)\psi}{\|(X_k \otimes I)\psi\|}.$$

Similarly if Bob's outcome is ℓ, then the state becomes

$$\frac{(I \otimes Y_\ell)\psi}{\|(I \otimes Y_\ell)\psi\|}.$$

The *joint probability* of getting outcome k for Alice and outcome ℓ for Bob, denoted by $p_{A,B}(k,\ell)$, is given by

$$p_{A,B}(k,\ell) = \|(X_k \otimes Y_\ell)\psi\|^2.$$

We can also use the notion of conditional probabilities in the quantum setting. The *conditional probability* that Bob gets outcome ℓ, given that Alice got outcome k is given by

$$p(B = \ell | A = k) = \frac{\|(I \otimes Y_\ell)(X_k \otimes I)\psi\|^2}{\|(X_k \otimes I)\psi\|^2} = \frac{p(B = \ell, A = k)}{p(A = k)},$$

since if Alice has already got the outcome k, that is, $A = k$ then the state is $\frac{(X_k \otimes I)\psi}{\|(X_k \otimes I)\psi\|}$, so that the probability of getting outcome ℓ for Bob given that $A = k$ is computed as in the usual definition of conditional probability.

Note: We use $A = k$ to mean "A gets the outcome k."

Definition 8.1. A state $\psi \in \mathcal{H}_A \otimes \mathcal{H}_B$ is said to be *separable* if it is of the form $\psi = \gamma \otimes \phi$ for some $\gamma \in \mathcal{H}_A$ and $\phi \in \mathcal{H}_B$; that is to say, ψ is an elementary tensor (without loss of generality, by scaling, each of γ and ϕ are norm 1). If ψ is not of this form, we say that ψ is *entangled*.

It is worth noticing how separable states behave in a combined lab. Indeed, if $\psi = \gamma \otimes \phi$ is separable with $\|\gamma\| = \|\phi\| = 1$, then one has

$$p_A(k) = \|(X_k \otimes I)(\gamma \otimes \phi)\|^2 = \|X_k\gamma \otimes \phi\|^2 = \|X_k\gamma\|^2\|\phi\|^2 = \|X_k\gamma\|^2,$$

while the joint probability becomes

$$\begin{aligned} p(B = \ell, A = k) &= \|(X_k \otimes Y_\ell)(\gamma \otimes \phi)\|^2 \\ &= \|X_k\gamma \otimes Y_\ell\phi\|^2 = \|X_k\gamma\|^2\|Y_\ell\phi\|^2 \\ &= p_A(k) \cdot p_B(\ell). \end{aligned}$$

Recall that in probability, events E_1, E_2 are *independent* if $\mathrm{Prob}(E_1 \cap E_2) = \mathrm{Prob}(E_1) \cdot \mathrm{Prob}(E_2)$, so we infer that $A = k$ and $B = \ell$ are

independent then and only then

$$\begin{aligned} p(B=\ell|A=k) &= \frac{\|(X_k\otimes Y_\ell)(\gamma\otimes\phi)\|^2}{\|(X_k\otimes I)(\gamma\otimes\phi)\|^2} \\ &= \frac{\|X_k\gamma\|^2\|Y_\ell\phi\|^2}{\|X_k\gamma\|^2} \\ &= \|Y_\ell\phi\|^2 \\ &= p(B=\ell). \end{aligned}$$

Thus in case of separable states, the quantum probabilities exactly reflect independent classical probabilities.

Definition 8.2. The state $\psi = \frac{1}{\sqrt{2}}(e_0\otimes e_0 + e_1\otimes e_1) \in \mathbb{C}^2\otimes\mathbb{C}^2$ is called the *Einstein-Poldosky-Rosen (EPR) state*.

Example 8.3 (Quantum Teleportation)**.** Let E_{00}, E_{11} be the diagonal matrix units in M_2. Then $E_{00}^*E_{00} + E_{11}^*E_{11} = E_{00}^2 + E_{11}^2 = I_{\mathcal{H}_A}$ where $\mathcal{H}_A = \mathbb{C}^2$, so this is a measurement system. Suppose Bob has the same measurement system (in a different lab) so that $\mathcal{H}_A = \mathcal{H}_B = \mathbb{C}^2$. In the combined lab, the probabilities are as follows:

$$\begin{aligned} p_A(0) = p_A(A=0) &= \|(E_{00}\otimes I)\psi\|^2 \\ &= \|(E_{00}\otimes I)(\frac{1}{\sqrt{2}})(e_0\otimes e_0 + e_1\otimes e_1)\|^2 \\ &= \frac{1}{2}\|e_0\otimes e_0\|^2 = \frac{1}{2}. \end{aligned}$$

Similarly, one can check that $p_A(1) = \frac{1}{2}$, while $p_B(0) = p_B(1) = \frac{1}{2}$.

If A observes the outcome 0, then the state changes to

$$\frac{(E_{00}\otimes I)(\frac{1}{\sqrt{2}})(e_0\otimes e_0 + e_1\otimes e_1)}{\|(E_{00}\otimes I)(\frac{1}{\sqrt{2}})(e_0\otimes e_0 + e_1\otimes e_1)\|} = e_0\otimes e_0.$$

Now suppose B performs a measurement. B has measurement operators $I\otimes E_{00}$ and $I\otimes E_{11}$. We know that $(I\otimes E_{11})(e_0\otimes e_0) = 0$, so B cannot possibly measure the outcome 1. Therefore, if A measures 0, then B must measure 0 with probability 1. The same analysis works if A measures 1. This demonstrates that entangled systems, to some degree, behave like dependent events. We confirm this with the computations below.

$$p(B=0|A=0) = \|(I\otimes E_{00})(e_0\otimes e_0)\|^2 = \|e_0\otimes e_0\|^2 = 1 \neq p_B(0) = \frac{1}{2},$$

$$p(B=1|A=0) = \|(I\otimes E_{11})(e_0\otimes e_0)\|^2 = 0 \neq p_B(1) = \frac{1}{2}.$$

This shows that there is a large amount of dependence here. This is the basis for "spooky action at a distance", or "quantum teleportation".

Example 8.4 (Super Dense Coding)**.** The idea is the following: If Alice has states in $\mathbb{C}^2$, we know that we can only make two states perfectly distinguishable. Consider the EPR state $\psi = \frac{1}{\sqrt{2}}(e_0 \otimes e_0 + e_1 \otimes e_1)$. Consider the matrices that were used to form the 1-Pauli's, given by

$$I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, Y = \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix}, \text{ and } Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Then Alice, after applying these operations on the EPR state, has

$$\begin{aligned} I\psi &= \frac{1}{\sqrt{2}}(e_0 \otimes e_0 + e_1 \otimes e_1) \\ X\psi &= \frac{1}{\sqrt{2}}(e_1 \otimes e_0 + e_0 \otimes e_1) \\ Y\psi &= \frac{1}{\sqrt{2}}(e_0 \otimes e_1 - e_1 \otimes e_0) \\ Z\psi &= \frac{1}{\sqrt{2}}(e_0 \otimes e_0 - e_1 \otimes e_1) \end{aligned}$$

four outcomes that are orthonormal and hence perfectly distinguishable!

Discussion: Suppose Alice and Bob start with the EPR state $\psi = \frac{1}{\sqrt{2}}(e_0 \otimes e_0 + e_1 \otimes e_1)$ and Alice performs one of the four above mentioned operations on the state of the photon in its lab and sends that single photon to Bob via a quantum channel. Bob now has access to both the photons (actually the access to two states, one EPR and one which he received from Alice, since the above four states are perfectly distinguishable and hence there exists a measurement system that tells Bob which state he received.) This allows B to know precisely which operation A performed. Moral: Alice only needed to send one photon to communicate four pieces of information. Similarly, in $\mathbb{C}^d$, with basis $e_0, \ldots, e_{d-1}$, consider the EPR state $\psi = \frac{1}{\sqrt{d}}(e_0 \otimes e_0 + \cdots + e_{d-1} \otimes e_{d-1})$. Then there exist d^2 unitaries, U_i, in $\mathbb{C}^d$ such that $(U_i \otimes I)\psi$ is orthogonal to $(U_j \otimes I)\psi$ for any $i \neq j$. Again, if B keeps half of the photons, then A can communicate d^2 pieces of information.This example shows the existence of a way (entanglement) to boost the capacity of the quantum channel in question.

Our next goal is to determine, given a noisy channel $\mathcal{E} : M_n \to M_m$ which is CPTP, if there is a way to use entanglement between sender and receiver to boost capacity. It turns out we can and this is known

as entanglement-enhanced capacity. Before exploring this in detail, we discuss entanglement in more detail.

Definition 8.5. An ensemble of states $\{\psi_i, p_i\}$, $\psi_i \in \mathcal{H}_A \otimes \mathcal{H}_B$, $p_i > 0$ and $\sum_i p_i = 1$ is called *separable* if each $\psi_i \in \mathcal{H}_A \otimes \mathcal{H}_B$ is separable. A density matrix $P \in \mathcal{L}(\mathcal{H}_A \otimes \mathcal{H}_B)$ is called *separable* if it is the density matrix of a separable ensemble.

Proposition 8.6. *Let $P \in \mathcal{L}(\mathcal{H}_A \otimes \mathcal{H}_B)$ be a density matrix. Then the following are equivalent.*

(1) *P is separable.*
(2) *There exist density matrices $R_i \in \mathcal{L}(\mathcal{H}_A)$ and $Q_i \in \mathcal{L}(\mathcal{H}_B)$ and $p_i > 0$ with $\sum_i p_i = 1$ such that $P = \sum_i p_i R_i \otimes Q_i$.*
(3) *There exist $E_i \in \mathcal{L}(\mathcal{H}_A)$ and $F_i \in \mathcal{L}(\mathcal{H}_B)$ that are rank one projections, along with $p_i > 0$ with $\sum_i p_i = 1$, such that $P = \sum_i p_i E_i \otimes F_i$.*

Proof. **Homework problem 13**; due 25th February, Thursday. □

8.1. **Partial Traces and Distinguishability.** Suppose we are given $Q, P \in \mathcal{L}(\mathcal{H}_A \otimes \mathcal{H}_B)$ as density matrices. Alice can only do measurements of the form $X \otimes I_B$ in the mixed lab. When can she distinguish Q from P?

Recall that if $X \in M_n$ and $Y \in M_p$, then

- $M_n \otimes M_p \simeq M_n(M_p) \simeq M_{np}$,
- $X \otimes Y = (x_{ij}Y) \in M_n \otimes M_p \simeq M_{np}$,
- $\mathrm{Tr}_{np}(X \otimes Y) = x_{11}\mathrm{Tr}_p(Y) + \cdots + x_{nn}\mathrm{Tr}_p(Y) = \mathrm{Tr}_n(X)\mathrm{Tr}_p(Y) = (\mathrm{Tr}_n \otimes \mathrm{Tr}_p)(X \otimes Y)$,
- $\mathcal{L}(\mathcal{H}_A) \otimes \mathcal{L}(\mathcal{H}_B) \simeq \mathcal{L}(\mathcal{H}_A \otimes \mathcal{H}_B)$ when $dim(\mathcal{H}_A) < \infty$ and $dim(\mathcal{H}_B) < \infty$. So if $Z \in \mathcal{L}(\mathcal{H}_A \otimes \mathcal{H}_B)$, then we can write $Z = \sum_i X_i \otimes Y_i$ with $X_i \in \mathcal{L}(\mathcal{H}_A)$ and $Y_i \in \mathcal{L}(\mathcal{H}_B)$ for all i.

We use the above calculations to motivate the notion of partial traces. Given a linear map

$$f : \mathcal{L}(\mathcal{H}_B) \longrightarrow \mathbb{C},$$

we have a linear map

$$I_{\mathcal{L}(\mathcal{H}_A)} \otimes f : \mathcal{L}(\mathcal{H}_A) \otimes \mathcal{L}(\mathcal{H}_B) \longrightarrow \mathcal{L}(\mathcal{H}_A) \otimes \mathbb{C} \simeq \mathcal{L}(\mathcal{H}_A)$$

given by

$$(I_{\mathcal{L}(\mathcal{H}_A)} \otimes f)(X \otimes Y) = X \cdot f(Y) \in \mathcal{L}(\mathcal{H}_A).$$

When we let

$$f = \mathrm{Tr} : \mathcal{L}(\mathcal{H}_B) \longrightarrow \mathbb{C},$$

we get a linear map

$$\mathrm{Tr}_B := I_{\mathcal{L}(\mathcal{H}_A)} \otimes \mathrm{Tr} : \mathcal{L}(\mathcal{H}_A \otimes \mathcal{H}_B) \to \mathcal{L}(\mathcal{H}_A),$$

called the *partial trace* with respect to the space $\mathcal{H}_B$. Similarly,

$$\mathrm{Tr}_A := \mathrm{Tr} \otimes I_{\mathcal{L}(\mathcal{H}_B)} : \mathcal{L}(\mathcal{H}_A \otimes \mathcal{H}_B) \to \mathcal{L}(\mathcal{H}_B)$$

is the partial trace with respect to the space $\mathcal{H}_A$.

More concretely, suppose that $\mathcal{H}_A = \mathbb{C}^n$ and $\mathcal{H}_B = \mathbb{C}^p$. Then $\mathcal{L}(\mathcal{H}_A) = M_n$ and $\mathcal{L}(\mathcal{H}_B) = M_p$. So,

$$\mathrm{Tr}_B(X \otimes Y) = X\mathrm{Tr}(Y) \text{ and } \mathrm{Tr}_A(X \otimes Y) = \mathrm{Tr}(X)Y.$$

Writing this out using block matrices, we obtain

$$\mathrm{Tr}_A(X \otimes Y) = \mathrm{Tr}_A((x_{ij}Y)) = \left(\sum_{i=1}^{n} x_{ii}\right) Y = \sum_i (x_{ii}Y).$$

If $Z = (Z_{ij}) \in M_n(M_p)$ is in block form (where $1 \leq i, j \leq n$), then

$$\mathrm{Tr}_A((Z_{ij})) = \sum_{i=1}^{n} Z_{ii} \in M_p.$$

We also see that $\mathrm{Tr}_B((x_{ij}Y)) = X \cdot \mathrm{Tr}(Y) = (x_{ij}\mathrm{Tr}(Y))$, so that $\mathrm{Tr}_B((Z_{ij})) = (\mathrm{Tr}(Z_{ij})) \in M_n$.

Lemma 8.7. *Let $P \in \mathcal{L}(\mathcal{H}_A \otimes \mathcal{H}_B)$ be a density matrix. Let $P_B := Tr_A(P) \in \mathcal{L}(\mathcal{H}_B)$ and $P_A := Tr_B(P) \in \mathcal{L}(\mathcal{H}_A)$ are density matrices.*

Proof. We first show that $P_B = \mathrm{Tr}_A(P)$ is a density matrix. Write $P = \sum_i X_i \otimes Y_i$, so that $P_B = \sum_i \mathrm{Tr}(X_i)Y_i$. Now, since P is a density matrix, we have

$$1 = \mathrm{Tr}(P) = \mathrm{Tr}\left(\sum_i X_i \otimes Y_i\right) = \sum_i \mathrm{Tr}(X_i \otimes Y_i) = \sum_i \mathrm{Tr}(X_i)\mathrm{Tr}(Y_i),$$

and the latter quantity is equal to $\mathrm{Tr}\left(\sum_i \mathrm{Tr}(X_i)Y_i\right)$ by linearity. By definition, this is none other than $\mathrm{Tr}(P_B)$ so that $\mathrm{Tr}(P_B) = 1$. We must show that $P_B \geq 0$. Let $v \in \mathcal{H}_B$, and let $\{e_k\}$ be an orthonormal basis for $\mathcal{H}_A$. Then $\mathrm{Tr}(X) = \sum_k \langle e_k | X e_k \rangle$ for any $X \in M_n$. Let

$w_k = e_k \otimes v \in \mathcal{H}_A \otimes \mathcal{H}_B$. Then since P is positive, we obtain

$$\begin{aligned}
0 &\leq \sum_k \langle e_k \otimes v | P(e_k \otimes v) \rangle \\
&= \sum_k \left\langle e_k \otimes v \middle| \sum_i (X_i \otimes Y_i) e_k \otimes v \right\rangle \\
&= \sum_i \sum_k \langle e_k \otimes v | X_i e_k \otimes Y_i v \rangle \\
&= \sum_i \left(\sum_k \langle e_k | X_i e_k \rangle \right) \langle v | Y_i v \rangle \\
&= \sum_i \mathrm{Tr}(X_i) \langle v | Y_i v \rangle \\
&= \langle v | P_B v \rangle.
\end{aligned}$$

Since v is arbitrary, it follows that $P_B \geq 0$. The proof is similar for P_A, so both are density matrices. □

Proposition 8.8. *Let $P \in \mathcal{L}(\mathcal{H}_A \otimes \mathcal{H}_B)$ be a density matrix. Suppose that Alice has a measurement system $\{M_k\}_k$ and Bob has a measurement system $\{W_\ell\}_\ell$. Let P_A, P_B be the partial traces of P as above. Then*

$$p(A = k) = Tr(M_k P_A M_k^*) \text{ and } p(B = \ell) = Tr(W_\ell P_B W_\ell^*).$$

Proof. As before, we need only prove the result for one of the above probabilities. Write $P = \sum_i X_i \otimes Y_i$ as before. Then

$$\begin{aligned}
p(A = k) &= \mathrm{Tr}((M_k \otimes I_{\mathcal{H}_B}) P (M_k \otimes I_{\mathcal{H}_B})^*) \\
&= \sum_i \mathrm{Tr}(M_k X_i M_k^* \otimes Y_i) \\
&= \sum_i \mathrm{Tr}(M_k X_i M_i^*) \cdot \mathrm{Tr}(Y_i) \\
&= \mathrm{Tr}\left(M_k \left(\sum_i X_i \mathrm{Tr}(Y_i) \right) M_k^* \right) \\
&= \mathrm{Tr}(M_k P_A M_k^*), \text{ as desired.}
\end{aligned}$$

Similarly for B. □

It is helpful to think of the above in the discrete case. Suppose Alice has events $\{x_1, ..., x_n\}$ and Bob has events $\{y_1, ..., y_p\}$. Consider a matrix with (i, j)-entry given by $p_{ij} = P(A = x_i, B = y_j)$. Then

$\sum_{i,j} p_{ij} = 1$, since this is the sum of all the possible joint probabilities. In this setting, we have

$$p(A = x_i) = \sum_j p(A = x_i, B = y_j) = \sum_j p_{ij},$$

while

$$p(B = y_j) = \sum_i p(A = x_i, B = y_j).$$

So in this case, a row sum gives a probability for Alice and a column sum gives a probability for Bob. This is like in the quantum setting, except you are "summing out" one of the state spaces (by taking the partial trace).

8.2. **Entanglement-Assisted One Shot Zero Error Capacity.** Suppose that Alice is sending a discrete number of messages (aka states) through a noisy channel to Bob, which corresponds to a CPTP map $\mathcal{N} : \mathcal{L}(\mathcal{H}_A) \to \mathcal{L}(\mathcal{H}_B)$. The idea is the following: suppose that Alice has experiments that she can conduct on a possibly different state space, given by experiments $\mathcal{E}_1, ..., \mathcal{E}_M : \mathcal{L}(\mathcal{H}_{A_0}) \to \mathcal{L}(\mathcal{H}_A)$. She wants Bob to know which experiment was conducted. We will need some sort of resource space corresponding to a perfect channel, and we assume that we have an entangled state $w \in \mathcal{H}_{A_0} \otimes \mathcal{H}_R$, where $\mathcal{H}_R$ corresponds to the resource space. So, initially we had the mappings

$$\mathcal{L}(\mathcal{H}_{A_0}) \xrightarrow{\mathcal{E}_m} \mathcal{L}(\mathcal{H}_A) \xrightarrow{\mathcal{N}} \mathcal{L}(\mathcal{H}_B),$$

and we obtain

$$\mathcal{L}(\mathcal{H}_{A_0} \otimes \mathcal{H}_R) \xrightarrow{\mathcal{E}_m \otimes \mathrm{id}_R} \mathcal{L}(\mathcal{H}_A \otimes \mathcal{H}_R) \xrightarrow{\mathcal{N} \otimes \mathrm{id}_R} \mathcal{L}(\mathcal{H}_B \otimes \mathcal{H}_R).$$

We denote $\rho_m = (\mathcal{N} \otimes \mathrm{id}_R) \circ (\mathcal{E}_m \otimes \mathrm{id}_R)(ww^*) \in \mathcal{L}(\mathcal{H}_B \otimes \mathcal{H}_R)$. We desire to have $\rho_1, ..., \rho_M$ perfectly distinguishable (then Bob can use a certain measurement system to determine what was sent).

Definition 8.9. Given a CPTP map $\mathcal{N} : \mathcal{L}(\mathcal{H}_A) \to \mathcal{L}(\mathcal{H}_B)$, the *entanglement-assisted one-shot zero error capacity* of $\mathcal{N}$, denoted by $\widetilde{\alpha}(\mathcal{N})$, is the largest M for which the following is true:

- there exists a finite-dimensional Hilbert space $\mathcal{H}_R$, an (entangled) state $w \in \mathcal{H}_{A_0} \otimes \mathcal{H}_R$, and CPTP maps $\mathcal{E}_1, ..., \mathcal{E}_M : \mathcal{L}(\mathcal{H}_{A_0}) \to \mathcal{L}(\mathcal{H}_A)$ such that the set $\{\rho_1, ..., \rho_M\}$ is perfectly distinguishable, where $\rho_m = (\eta \otimes \mathrm{id}_R) \circ (\mathcal{E}_m \otimes \mathrm{id}_R)(ww^*)$ for every $m \in \{1, ..., M\}$.

Definition 8.10. Given an operator system $\mathcal{S} \subseteq M_n$, the *entanglement-assisted independence number* of $\mathcal{S}$, denoted by $\tilde{\alpha}(\mathcal{S})$, is the largest

M for which there exists $d \in \mathbb{N}$, a state $\varphi \in \mathbb{C}^n \otimes \mathbb{C}^d$ and unitaries $U_1, ..., U_M \in M_{nd}$ such that for each $m \neq n$,

$$U_n \varphi\varphi^* U_m^* \perp M_d(\mathcal{S}) = \{(s_{ij}) \in M_d(M_n) : s_{ij} \in \mathcal{S}, \forall 1 \leq i, j \leq n\}.$$

The following remarks are essential for the next theorem.

Remark 8.11. If $V : \mathbb{C}^\ell \to \mathbb{C}^n$ is an isometry (so $\ell \leq n$ and $V \in M_{n,\ell}$), then there is $W \in M_{n,n-\ell}$ such that $U = \begin{pmatrix} V & \vdots & W \end{pmatrix}_{n \times n}$ is unitary. Indeed, since $\dim(V(\mathbb{C}^\ell)) = \ell$, we have $\dim(V(\mathbb{C}^\ell)^\perp) = n - \ell$, so we may cefine an isometry $W : \mathbb{C}^{n-\ell} \to V(\mathbb{C}^\ell)^\perp \subseteq \mathbb{C}^n$. One can check that W works.

Remark 8.12. If $\mathcal{E} : M_\ell \to M_n$ is a CPTP map, and $\mathcal{E}(X) = \sum_{i=1}^d A_i X A_i^*$ where $A_i \in M_{n,\ell}$ with $\sum_{i=1}^d A_i^* A_i = I_\ell$, then $V = \begin{pmatrix} A_1 \\ \vdots \\ A_d \end{pmatrix}$ is such that $V^*V = I_\ell$, so that V is an isometry. Now choose $d' \geq d$ such that $\ell | nd'$. Then let $\tilde{V} = \begin{pmatrix} A_1 \\ \vdots \\ A_d \\ 0 \\ \vdots \\ 0 \end{pmatrix}$ be an $nd' \times \ell$ matrix. Now we can find other block matrices such that $U = \begin{pmatrix} \tilde{V} & \ddots & \tilde{W} \end{pmatrix} \in M_{d'}(M_{n,\ell})$ is unitary.

Theorem 8.13. (DSW) *Let $\mathcal{N} : \mathcal{L}(\mathcal{H}_A) \to \mathcal{L}(\mathcal{H}_B)$ be a CPTP map. Write $\mathcal{N}(X) = \sum_i E_i X E_i^*$, so that $\mathcal{S}_\mathcal{N} = span\{E_i^* E_j\}_{i,j} \subseteq \mathcal{L}(\mathcal{H}_A)$ is the corresponding operator system. Then $\tilde{\alpha}(\mathcal{N}) = \tilde{\alpha}(\mathcal{S}_\mathcal{N})$.*

Proof. This proof is quite long, so we will only show that $\tilde{\alpha}(\mathcal{N}) \leq \tilde{\alpha}(\mathcal{S}_\mathcal{N})$. Suppose that $\tilde{\alpha}(\mathcal{N}) = M$ and $\mathcal{N}(X) = \sum_k E_k X E_k^*$ for $X \in \mathcal{L}(\mathcal{H}_A)$. Write $\mathcal{E}_1, ..., \mathcal{E}_M : \mathcal{L}(\mathcal{H}_{A_0}) \to \mathcal{L}(\mathcal{H}_A)$ to be CPTP maps as in the definition of $\tilde{\alpha}(\mathcal{N})$. We may write $\mathcal{E}_m(Y) = \sum_{\alpha=1}^d A_{m,\alpha} Y A_{m,\alpha}^*$ (we may choose d to be the same for all the $\mathcal{E}_m$'s by adding enough zeros). We may assume that $\mathcal{H}_R \simeq \mathbb{C}^r$. As in the definition of $\tilde{\alpha}(\mathcal{N})$, we choose $w \in \mathcal{H}_{A_0} \otimes \mathbb{C}^k$, say $w = \begin{pmatrix} w_1 \\ \vdots \\ w_r \end{pmatrix}$ where $w_i \in \mathcal{H}_{A_0}$. Then

$$\rho_m = (\mathcal{N} \otimes \mathrm{id}_r) \circ (\mathcal{E}_m \otimes \mathrm{id}_r)(ww^*),$$

and $\rho_1, ..., \rho_m$ are perfectly distinguishable. This is equivalent to having

$$\rho_n \cdot \rho_m = 0 \text{ for } n \neq m.$$

We may write

$$\begin{aligned}\rho_m &= \sum_{k,\alpha} (E_k \otimes I_r)(A_{m,\alpha} \otimes I_r)(ww^*)(A_{m,\alpha} \otimes I_r)^*(E_k \otimes I_r)^* \\ &= \sum_{k,\alpha} (E_k A_{m,\alpha} \otimes I_r)(ww^*)(A^*_{m,\alpha} E^*_k \otimes I_r).\end{aligned}$$

Each term in the sum is positive, so having $\rho_n \cdot \rho_m = 0$ for $n \neq m$ implies that each term in the product of their sums must be 0. Therefore,

$$(E_k A_{n,\alpha} \otimes I_r)(ww^*)(A^*_{n,\alpha} E^*_k \otimes I_r)(E_\ell A_{m,\beta} \otimes I_r)(ww^*)(A^*_{m,\beta} E^*_\ell \otimes I_r) = 0.$$

Starting at the first occurrence of w^* and ending at the last occurence of w gives an inner product. Hence, the above is equivalent to having

$$[(E_k A_{n,\alpha} \otimes I_r)w] \langle (E_k A_{n,\alpha} \otimes I_r)w | (E_\ell A_{m,\beta} \otimes I_r)w \rangle [w^*(A^*_{m,\beta} E^*_\ell \otimes I_r)] = 0.$$

This is equivalent (by taking trace) to having

$$\langle (E_k A_{n,\alpha} \otimes I_r)w | (E_\ell A_{m,\beta} \otimes I_r)w \rangle \cdot \langle (E_\ell A_{m,\beta} \otimes I_r)w | (E_k A_{n,\alpha} \otimes I_r)w \rangle = 0,$$

which happens if and only if $(E_k A_{n,\alpha} \otimes I_r)w \perp (E_\ell A_{m,\beta} \otimes I_r)w$ for all k, α, ℓ, β, and for all $n \neq m$. We recall that $w = \begin{pmatrix} w_1 \\ \vdots \\ w_r \end{pmatrix}$, so the above condition can be written as

$$\begin{aligned}0 &= \left\langle (E_k A_{n,\alpha} \otimes I_r) \begin{pmatrix} w_1 \\ \vdots \\ w_r \end{pmatrix}, (E_\ell A_{m,\ell} \otimes I_r) \begin{pmatrix} w_1 \\ \vdots \\ w_r \end{pmatrix} \right\rangle \\ &= \sum_{t=1}^{r} \langle E_k A_{n,\alpha} w_t | E_\ell A_{m,\beta} w_t \rangle \\ &= \sum_{t=1}^{r} \operatorname{Tr}\left(E_\ell A_{m,\beta} w_t w_t^* A^*_{n,\alpha} E^*_k\right).\end{aligned}$$

Let $\chi = \sum_{t=1}^{r} w_t w_t^*$. The above equation becomes

$$\begin{aligned}0 &= \operatorname{Tr}(E_\ell A_{m,\beta} \chi A^*_{n,\alpha} E^*_k) \\ &= \operatorname{Tr}(E^*_k E_\ell (A_{m,\beta} \chi A^*_{n,\alpha}))\end{aligned}$$

This implies that $A_{m,\beta} \chi A^*_{n,\alpha} \perp \mathcal{S}_\mathcal{N}$ for all α, β and for all $m \neq n$. We may use the previous remark to build unitaries $U_m = \begin{pmatrix} \tilde{V}_m & \vdots & \tilde{W}_m \end{pmatrix}$

where $\tilde{V}_m = \begin{pmatrix} A_{m,1} \\ \vdots \\ A_{m,d} \\ 0 \\ \vdots \\ 0 \end{pmatrix}$ is of size $nd' \times \ell$ for some $d' \geq d$. Moreover, we may ensure that all of the U_m's are of the same size. Then $U_m \underbrace{\begin{pmatrix} \chi & & & \\ & 0 & & \\ & & \ddots & \\ & & & 0 \end{pmatrix}}_{\tilde{\chi}} U_n^* = (A_{m,\beta}\chi A_{n,\alpha}^*)_{1\leq\beta,\alpha\leq d'}$. Each one of these block entries is perpendicular to $\mathcal{S}_\mathcal{N}$, so we have $U_m\tilde{\chi}U_n = (A_{m,\beta}\chi A_{n,\alpha}^*) \perp M_{d'}(\mathcal{S}_\mathcal{N})$, for all $1 \leq m, n \leq M$ and $m \neq n$. Therefore, $M \leq \tilde{\alpha}(\mathcal{S}_\mathcal{N})$, so that $\tilde{\alpha}(\mathcal{N}) \leq \tilde{\alpha}(\mathcal{S}_\mathcal{N})$. □

Corollary 8.14. *If $\mathcal{S}_{\mathcal{N}_1} = \mathcal{S}_{\mathcal{N}_2}$ then $\tilde{\alpha}(\mathcal{N}_1) = \tilde{\alpha}(\mathcal{N}_2)$.*

DSW[**?**] next introduces the concept of quantum Lovasz function, a quantum analogue of the Lovasz function. Recall that if G is a graph on n vertices, then $\theta(G) = \sup\{\lambda_{max}(I_n + H) : H = H^*, h_{ij} = 0 \ \forall (i,j) \in E(G), h_{i,i} = 0 \forall i, I + H \geq 0\}$. Note that, in the above definition, "$h_{ij} = 0 \, \forall (i,j) \in E(G)$ and each $h_{i,i} = 0$" happens if an d only if H precisely has 0's where the operator system of G, $\mathcal{S}_G$ is having something non-zero. Thus H must be orthogonal to $\mathcal{S}_G$. Consequently, having H as in the supremum above is equivalent to having $H = H^*$, $I + H \geq 0$ and $H \perp \mathcal{S}_G$, that is,

$$\theta(G) = \sup\{\lambda_{max}(I_n + H) : H = H^*, \, I + H \geq 0, \ H \perp \mathcal{S}_G\}.$$

Recall that $\theta(G) \geq \alpha(G)$, the one-shot zero error capacity of the classical channel with confusability graph G.

Definition 8.15. Given an operator system $\mathcal{S} \subseteq M_n$, the *Quantum Lovasz function* of $\mathcal{S}$ is defined by

$$\theta(\mathcal{S}) = \sup\{\lambda_{max}(I_n + H) : H = H^*, \, I + H \geq 0, \, H \perp \mathcal{S}\}.$$

We also define, for $d \geq 1$, the function

$$\theta_d(\mathcal{S}) = \sup\{\lambda_{max}(I_{nd} + H) : H = H^*, \, I_{nd} + H \geq 0, \, H \perp M_d(\mathcal{S})\},$$

and define $\tilde{\theta}(\mathcal{S}) = \sup_{d\in\mathbb{N}} \theta_d(\mathcal{S})$.

Theorem 8.16 (DSW)**.**

(1) *If G is a and $\mathcal{S}_G \subseteq M_n$ (where G is a graph on n vertices) is the associated operator system, then $\theta(G) = \theta(\mathcal{S}_G) = \tilde{\theta}(\mathcal{S}_G)$.*
(2) *$\theta(\mathcal{S}) \geq \alpha(\mathcal{S})$ and $\tilde{\theta}(\mathcal{S}) \geq \tilde{\alpha}(\mathcal{S})$ whenever $\mathcal{S}$ is an operator system.*
(3) *If G, H are graphs, then $\mathcal{S}_G \otimes \mathcal{S}_H = \mathcal{S}_{G \boxtimes H}$.*
(4) *If $\mathcal{S}, \mathcal{T}$ are operator systems, then $\tilde{\theta}(\mathcal{S} \otimes \mathcal{T}) = \tilde{\theta}(\mathcal{S})\tilde{\theta}(\mathcal{T})$.*

Corollary 8.17. *Let $\mathcal{N} : \mathcal{L}(\mathcal{H}_A) \to \mathcal{L}(\mathcal{H}_B)$ be a CPTP map. Then*

$$\tilde{\theta}(\mathcal{S}_{\mathcal{N}}) \geq \sup_{m \in \mathbb{N}} \sqrt[m]{\tilde{\alpha}(\mathcal{N}^{\otimes m})}.$$

Proof. By (2) of the previous theorem, $\tilde{\theta}(\mathcal{S}_{\mathcal{N}^{\otimes m}}) \geq \tilde{\alpha}(\mathcal{N}^{\otimes m})$. But $\mathcal{S}_{\mathcal{N}^{\otimes m}} = (\mathcal{S}_{\mathcal{N}})^{\otimes m}$, so we have $\tilde{\theta}(\mathcal{S}_{\mathcal{N}^{\otimes m}}) = \tilde{\theta}(\mathcal{S}_{\mathcal{N}})^m$. Taking m-th roots on both sides yields the required result. □

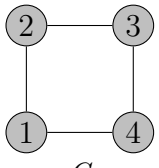
A few observations are in order. Using (1) and (2), we obtain $\theta(G) \geq \alpha(G)$ which is Lovasz's inequality. This is a classic inequality in graph theory which follows easily from their work. Moreover, combining (1), (3) and (4), we have $\theta(G \boxtimes H) = \theta(G)\theta(H)$ which is the multiplicativity result of Lovasz from before. So, these functions are very good generalizations of what was done before.

Recall that $\theta(G) = \max\{\lambda_{\max}(I + K) : K = K^*, K \perp \mathcal{S}_G, I + K \geq 0\}$. Lovasz proves that

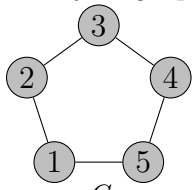
$$\theta(G) = \min\{\lambda_{\max}(A) : A = A^*, a_{ij} = 1 \forall i \nsim j\}.$$

This is useful for:

Problem 8.18. *Find $\theta(C_4)$ and $\theta(C_5)$ where C_4 is the 4-cycle graph*



C_4



C_5

and C_5 is the 5-cycle graph.

Proof. **Homework problem 14**; due 3rd March, Thursday. □

Problem 8.19. *Let $\mathcal{S} = span\{I_n, E_{ij} : i \neq j\}$. Compute $\theta(\mathcal{S})$ and $\tilde{\theta}(\mathcal{S})$.*

Proof. **Homework problem 15**; due 3rd March, Thursday. □

9. Dilations: State Purification, POVM's vs. PVM's

The idea of dilation is to make things simpler by representing them on a larger space.

Example 9.1. Suppose we want to find a formula for $\cos(\alpha+\beta)$. The easiest way is to think of $e^{i\alpha} = \cos(\alpha) + i\cos\beta$ so that $\cos(\theta) = \mathrm{Re}(e^{i\theta})$. (Here, we are in a sense dilating to a form of $\mathbb{R}^2$, namely $\mathbb{C}$.) Now the problem is very easy, since

$$\begin{aligned}\cos(\alpha+\beta) &= \mathrm{Re}(e^{i(\alpha+\beta)}) \\ &= \mathrm{Re}(e^{i\alpha}e^{i\beta}) \\ &= \mathrm{Re}((\cos\alpha + i\sin\alpha)(\cos\beta + i\sin\beta)) \\ &= \cos\alpha\cos\beta - \sin\alpha\sin\beta.\end{aligned}$$

Example 9.2. Consider the Fibonacci numbers given by $f_1 = f_2 = 1$ and $f_{n+2} = f_{n+1} + f_n$. Then these numbers are really satisfying the equation

$$\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} f_n \\ f_{n+1}\end{pmatrix} = \begin{pmatrix} f_{n+1} \\ f_{n+2}\end{pmatrix}.$$

Let $R = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$. Then $R^n\begin{pmatrix} 1 \\ 1\end{pmatrix} = \begin{pmatrix} f_{n+1} \\ f_{n+2}\end{pmatrix}$. Now, $R = R^*$ is diagonalizable with eigenvalues $\frac{1\pm\sqrt{5}}{2}$. Hence, for some unitary U, we have $R = U^*\begin{pmatrix} \frac{1+\sqrt{5}}{2} & 0 \\ 0 & \frac{1-\sqrt{5}}{2}\end{pmatrix}U$. It follows that

$$\begin{pmatrix} f_{n+1} \\ f_{n+2}\end{pmatrix} = U^*\begin{pmatrix} \left(\frac{1+\sqrt{5}}{2}\right)^n & 0 \\ 0 & \left(\frac{1-\sqrt{5}}{2}\right)^n\end{pmatrix}U\begin{pmatrix} 1 \\ 1\end{pmatrix}.$$

This gives a nice explicit formula!

Example 9.3. *(Halmos Dilation)* Let $C \in \mathcal{L}(\mathcal{H})$ with $\|C\| \leq 1$. Consider the operator matrix $\begin{pmatrix} C & \sqrt{I-CC^*} \\ \sqrt{I-C^*C} & -C^*\end{pmatrix} \in \mathcal{L}(\mathcal{H}\oplus\mathcal{H})$. One can check that this is a unitary, while C is just a corner of the matrix.

9.1. **State Purification.** In the quantum setting, this idea of dilation is essentially what Physicists refer to as state purification. To be more precise, consider a measurement system $\{X_i\}_{i=1}^m$ with each $X_i : \mathcal{H}_A \to \mathcal{K}$, so that $\sum_{i=1}^m X_i^*X_i = I_A$. Consider an ensemble $\{v_k, p_k\}_{k=1}^r$ where $v_k \in \mathcal{H}_A$ with $\|v_k\| = 1$ and $\sum_{k=1}^r p_k = 1$. Let $\rho = \sum_{k=1}^r p_k v_k v_k^*$ be the density matrix of the ensemble. Recall that the probability of getting outcome i is

$$\sum_{k=1}^r p_k\|X_iv_k\|^2 = \mathrm{Tr}(X_i\rho X_i^*).$$

Suppose we are only interested in the probabilities. Then we can replace ρ by a pure state on a larger space. We will see this in two ways.

The first way to see the above is by letting

$$v = \begin{pmatrix} \sqrt{p_1}v_1 \\ \vdots \\ \sqrt{p_r}v_r \end{pmatrix} \in \mathcal{H}_A \otimes \mathbb{C}^r \text{ where } r = \text{rank } (\rho).$$

One can see that v is a unit vector. Now replace X_i by $\tilde{X}_i : \mathcal{H}_A \otimes \mathbb{C}^r \to \mathcal{K} \otimes \mathbb{C}^r$, with

$$\tilde{X}_i = \begin{pmatrix} X_i & & \\ & \ddots & \\ & & X_i \end{pmatrix} = X_i \otimes I_r.$$

Then it follows that

$$\sum_{i=1}^{m} \tilde{X}_i^* \tilde{X}_i = I_{\mathcal{H}_A} \otimes I_r = \begin{pmatrix} I_{\mathcal{H}_A} & & \\ & \ddots & \\ & & I_{\mathcal{H}_A} \end{pmatrix}.$$

Hence, $\{\tilde{X}_i\}_{i=1}^m$ is a measurement system on $\mathcal{H}_A \otimes \mathbb{C}^r$. Moreover, it is readily checked that

$$\|\tilde{X}_i v\|^2 = \sum_{k=1}^{r} p_k \|X_i v_k\|^2,$$

which is the same probability as before (for getting outcome i). Hence, if we only care about probabilities, then we may replace our ensemble with a pure state.

The second way to obtain the above is more canonical. Let $\mathcal{H}_A = \mathbb{C}^n$ and $\mathcal{L}(\mathcal{H}_A) = M_n \simeq \mathbb{C}^{n^2}$ which is a Hilbert space. Given $X_i \in \mathcal{L}(\mathcal{H}_A) = M_n$, we may obtain a "new map" $\tilde{X}_i : M_n \to M_n$, defined by

$$\tilde{X}_i(Y) = X_i Y = \begin{pmatrix} X_i y_1 & \vdots & \cdots & \vdots & X_i y_n \end{pmatrix},$$

where $Y = \begin{pmatrix} y_1 & \vdots & \cdots & \vdots & y_n \end{pmatrix} \in M_n$. Then

$$\tilde{X}_i Y \simeq \begin{pmatrix} X_i & & \\ & \ddots & \\ & & X_i \end{pmatrix} \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}.$$

Given a density matrix $\rho \in M_n$, the probability of obtaining outcome i is

$$\text{Tr}(X_i \rho X_i^*) = \text{Tr}((X_i \rho^{\frac{1}{2}})(\rho^{\frac{1}{2}} X_i^*)) = \text{Tr}((\rho^{\frac{1}{2}} X_i^*)(X_i \rho^{\frac{1}{2}})) = \langle \tilde{X}_i(\rho^{\frac{1}{2}}) | \tilde{X}_i(\rho^{\frac{1}{2}}) \rangle_{M_n}.$$

Since $\rho^{\frac{1}{2}}$ is a vector in M_n, we obtain

$$\|\rho^{\frac{1}{2}}\|^2_{M_n} = \langle \rho^{\frac{1}{2}} | \rho^{\frac{1}{2}} \rangle_{M_n} = \mathrm{Tr}((\rho^{\frac{1}{2}})^*(\rho^{\frac{1}{2}})) = \mathrm{Tr}(\rho) = 1.$$

It follows that $\rho^{\frac{1}{2}}$ is a unit vector, hence a pure state in the Hilbert space M_n. (So, when physicists replace ρ by $\sqrt{\rho}\sqrt{\rho}$, they actually consider $\sqrt{\rho}$ to be a pure state in space M_n.)

9.2. **POVM's and PVM's.** With this in mind, we will now talk about positive operator-valued measures (POVM's). If we are only interested in probabilities in the context of a measurement system $\{X_i\}_i$ and a state h, then the probability of obtaining outcome i is $\|X_i h\|^2 = \langle X_i h | X_i h \rangle = \langle h | X_i^* X_i h \rangle$. The probability, then only really depends on $R_i = X_i^* X_i \geq 0$, while $I = \sum_{i=1}^m X_i^* X_i = \sum_{i=1}^m R_i$.

Definition 9.4 (POVM)**.** An *m-outcome positive operator-valued measure* (POVM) on a Hilbert space $\mathcal{H}_A$ is a set $\{R_i\}_{i=1}^m \subseteq \mathcal{L}(\mathcal{H}_A)$ with $R_i \geq 0$ and $\sum_{i=1}^m R_i = I$.

Definition 9.5 (PVM)**.** An *m-outcome projection-valued measure* (PVM) on a Hilbert space $\mathcal{H}_A$ is a set of projections $\{P_i\}_{i=1}^m$ in $\mathcal{L}(\mathcal{H}_A)$ such that $\sum_{i=1}^m P_i = I_{\mathcal{H}_A}$.

Obviously, every PVM is a POVM. Note that each PVM corresponds to decomposing the Hilbert space $\mathcal{H}$ into a direct sum of its subspaces. Namely, $\mathcal{H} = \mathcal{H}_1 + \cdots + \mathcal{H}_n$, where $\mathcal{H}_i$ is a subspace of $\mathcal{H}$ and $\mathcal{H}_i \perp \mathcal{H}_j$, for $i \neq j$.

Proposition 9.6. *Let $\{R_i\}_{i=1}^m$ be a POVM on $\mathcal{H}$. Then there is a PVM $\{P_i\}_{i=1}^m$ on $\mathcal{H} \otimes \mathbb{C}^m$ and an isometry $V : \mathcal{H} \to \mathcal{H} \otimes \mathbb{C}^m$ such that $R_i = V^* P_i V$ for all $1 \leq i \leq m$.*

Proof. As usual, we identify

$$\mathcal{H} \otimes \mathbb{C}^m \cong \underbrace{\mathcal{H} \oplus \cdots \oplus \mathcal{H}}_{m \text{ times}} = \mathcal{H}^{\otimes m}.$$

Let $P_i = I_{\mathcal{H}} \otimes E_{ii}$, the $m \times m$ matrix over $\mathcal{L}(\mathcal{H})$ with $I_{\mathcal{H}}$ in the (i,i)-position and 0 everywhere else. That is, each P_i is the projection onto the i-th copy of $\mathcal{H}$ in $\mathcal{H} \otimes \mathbb{C}^m$ which means

$$P_i \left(\begin{bmatrix} h_1 \\ \vdots \\ h_n \end{bmatrix} \right) = \begin{bmatrix} 0 \\ \vdots \\ h_i \\ \vdots \\ 0 \end{bmatrix}.$$

Then clearly $P_i = P_i^* = P_i^2$ and $\sum_{i=1}^m P_i = I_{\mathcal{H}} \otimes I_m = I_{\mathcal{H}\otimes\mathbb{C}^m}$. Hence, $\{P_i\}_{i=1}^m$ is a PVM. Define $V : \mathcal{H} \to \mathcal{H}\otimes\mathbb{C}^m$ by $Vh = \sum_{i=1}^m (R_i^{1/2}h)\otimes e_i = (R_1^{\frac{1}{2}}h, ..., R_m^{\frac{1}{2}}h)$. This is linear, with

$$\|Vh\|^2 = \sum_{i=1}^m \langle R_i^{\frac{1}{2}}h | R_i^{\frac{1}{2}}h\rangle = \sum_{i=1}^m \langle h|R_i h\rangle = \langle h, h\rangle = \|h\|^2.$$

Therefore, V is an isometry. Finally, we see that

$$\begin{aligned}\langle h|V^*P_iVh\rangle &= \langle Vh|(0, ..., 0, \underbrace{R_i^{\frac{1}{2}}h}_{i\text{-th slot}}, 0, ..., 0)\rangle \\ &= \langle (R_1^{\frac{1}{2}}h, ..., R_i^{\frac{1}{2}}h, ..., R_m^{\frac{1}{2}}h)|(0, ..., 0, R_i^{\frac{1}{2}}h, 0, ..., 0)\rangle \\ &= \langle R_i^{\frac{1}{2}}h|R_i^{\frac{1}{2}}h\rangle \\ &= \langle h|R_ih\rangle,\end{aligned}$$

which implies that $R_i = V^*P_iV$. □

Regarding the previous proposition, a few notes are in order:

(1) If $h \in \mathcal{H}$, then

$$\langle h|R_ih\rangle_{\mathcal{H}} = \langle h|V^*P_iVh\rangle_{\mathcal{H}} = \langle Vh|P_i(Vh)\rangle_{\mathcal{H}\otimes\mathbb{C}^m},$$

so with $\tilde{h} = Vh$, we have

$$\|\tilde{h}\| = \|h\| \text{ and } \text{prob}(i) = \langle h|R_ih\rangle = \langle \tilde{h}|P_i\tilde{h}\rangle.$$

So, to summarize, when we are given a POVM $\{R_i\}_{i=1}^n$ on $\mathcal{H}$ and a state $h \in \mathcal{H}$, we may regard the pair as the PVM $\{P_i\}_{i=1}^n$ on a larger Hilbert space $\mathcal{H}\otimes\mathbb{C}^n$ together with the state $\tilde{h} = Vh \in \mathcal{H}\otimes\mathbb{C}^n$ with the same probabilities of outcomes.

(2) Note that R_i, R_j, R_i^2 can be anything and there is no reason to have $R_iR_j = R_jR_i$. But, for a PVM, the fact that the projections add to the identity implies that their ranges are pairwise orthogonal, so that $P_iP_j = 0$ for $i \neq j$. Hence, the computations with a PVM are *much better* than the computations with a POVM.

Remark 9.7. This proposition actually is a special case of the Stinespring's dilation theorem. Let $l_n^\infty = C(\{1, \ldots, n\})$ be the abelian C*-algebra generated by the functions δ_i given by $\delta_i(j) = \delta_{ij}$. Thus a complex-valued function f on $\mathbb{C}^n$ can be regarded as $(\lambda_1, \ldots, \lambda_n)$, where $\lambda_i = f(i)$, and we identify $f = \sum_{i=1}^n \lambda_i\delta_i$. Hence each δ_i corresponds to the basis vector e_i for l_n^∞.

We can then define a completely positive map $\Phi\colon l_n^\infty \to \mathcal{L}(\mathcal{H})$ by $\Phi(\delta_i) = R_i$, so $\Phi((\lambda_1, \ldots, \lambda_n)) = \sum_{i=1}^n \lambda_iR_i$. There is also a map

$\pi\colon l_n^\infty \to \mathcal{L}(\mathcal{H}\otimes\mathbb{C}^n)$ by $\pi(\delta_i) = P_i$, $\pi((\lambda_1,\dots,\lambda_n)) = \sum_{i=1}^n \lambda_i P_i$. Since the P_i's are orthgonal projections,

$$\begin{aligned}\pi((\lambda_1,\dots,\lambda_n)\cdot(\mu_1,\dots,\mu_n)) &= \sum_{i=1}^n \lambda_i\mu_i P_i = (\sum_{i=1}^n \lambda_i P_i)(\sum_{j=1}^n \mu_j P_j)\\ &= \pi((\lambda_1,\dots,\lambda_n))\pi((\mu_1,\dots,\mu_n)).\end{aligned}$$

Readers can check that π preserves the unit and adjoints. Hence π is indeed a unital $*$-homomorphism. Moreover,

$$\begin{aligned}V^*\pi((\lambda_1,\dots,\lambda_n))V = V^*(\sum_{i=1}^n \lambda_i P_i)V &= \sum_{i=1}^n \lambda_i V^* P_i V\\ &= \sum_{i=1}^n \lambda_i R_i = \Phi((\lambda_1,\dots,\lambda_n)).\end{aligned}$$

9.3. **POVM's and PVM's in multiexperiment settings.** Now what if we have more than one measurement system? Suppose that Alice's state space is $\mathcal{H}_A$. Suppose that she has a whole family of experiments from which she could choose. Each one is represented by a POVM $\{R_{t,i}\}_{i=1}^m$ where $t \in T$ and T is the set of experiments.

(Note that in complicated situation each experiment in the family may have different number of outcomes. However, we can get rid of that by adding extra outcomes when necessary with 0 probabilities and assuming that all experiments have same number of outcomes; the number being that of the highest-outcome experiment.)

The next theorem tells us that we can again dilate this family of POVM's into a family of PVM's simultaneously; that is, using only one isometry V that works for all $v \in \mathcal{V}$.

Theorem 9.8. *Let $\{R_{t,i}\}_{i=1}^m$ be a family of POVM's on $\mathcal{H}$, indexed by $t \in T$ with $|T| = n$ (we only consider the case where $|T|$ is finite). Then there is a Hilbert space $\mathcal{K}$ and a family of PVM's $\{P_{t,i}\}_{i=1}^m$ on $\mathcal{K}$ for $t \in T$, and an isometry $V : \mathcal{H} \to \mathcal{K}$ such that $V^*P_{t,i}V = R_{t,i}$ for all i, t. Moreover, if $\dim(\mathcal{H}_A) < \infty$, then $\dim(\mathcal{K}) < \infty$.*

Proof. We just proved the case for $n = 1$, so we proceed by induction. Assume that we can do this for $|T| = n$. Now suppose we have $n + 1$ experiments. We know that there exists a Hilbert space $\mathcal{K}_1$, an isometry $V_1 : \mathcal{H} \to \mathcal{K}_1$ and PVM's $\{P_{t,i}\}_{i=1}^m$ for $1 \le t \le n$ such that $V_1^*P_{t,i}V_1 = R_{t,i}$ for all i and for all $1 \le t \le n$. Let $\tilde{R}_{n+1,i} = V_1 R_{n+1,i} V_1^* \in \mathcal{L}(\mathcal{K}_1)$.

Then $\tilde{R}_{n+1,i} \geq 0$ and

$$\sum_{i=1}^{m} \tilde{R}_{n+1,i} = V_1 \left(\sum_{i=1}^{m} R_{n+1,i} \right) V_1^* = V_1 V_1^*,$$

which is a projection. Adjust $\tilde{R}_{n+1,1}$ by setting $\tilde{R}_{n+1,1} = V_1 R_{n+1,1} V^* + (I - V_1 V_1^*)$, so that $\{\tilde{R}_{n+1,i}\}_{i=1}^m$ are a POVM on $\mathcal{K}_1$.

On $\mathcal{K}_1$, we have PVM's $\{P_{t,i}\}_{i=1}^m$ for $1 \leq n$ and a POVM $\{\tilde{R}_{n+1,i}\}_{i=1}^m$. Let $\mathcal{K} = \mathcal{K}_1 \otimes \mathbb{C}^m$, and define $V_2 : \mathcal{K}_1 \to \mathcal{K}$ by

$$V_2 k = ((\tilde{R}_{n+1,1})^{\frac{1}{2}} k, ..., (\tilde{R}_{n+1,m})^{\frac{1}{2}} k).$$

Then V_2 is an isometry. Set $P_{n+1,i} = I_{\mathcal{K}_1} \otimes E_{ii}$, for $1 \leq i \leq m$. Then $\{P_{n+1,i}\}_{i=1}^m$ is a PVM and $V_2^* P_{n+1,i} V_2 = \tilde{R}_{n+1,i}$. Now set $Q_{t,j} = V_2 P_{t,j} V_2^* \in \mathcal{L}(\mathcal{K})$ for $2 \leq j \leq m$, and set $Q_{t,1} = V_2 P_{t,1} V_2^* + (I - V_2 V_2^*)$, for all $1 \leq t \leq n$. It is easy to see that $V_2^* Q_{t,j} V_2 = P_{t,j}$. We need to see that $\{Q_{t,i}\}_{i=1}^m$ for $1 \leq t \leq n$ are PVM's and not just POVM's. Note that $\{P_{n+1,i}\}_{i=1}^m$ is a PVM. For the other ones, we will see that $Q_{t,j}^2 = Q_{t,j}$ and hence they must be PVM's (for $i \geq 2$). Note that

$$Q_{t,j}^2 = (V_2 P_{t,j} V_2^*)(V_2 P_{t,j} V_2^*) = V_2 P_{t,j}^2 V_2 = V_2 P_{t,j} V_2 = Q_{t,j}.$$

Finally,

$$\begin{aligned} Q_{t,1}^2 &= [V_2 P_{t,1} V_2^* + (I - V_2 V_2^*)][V_2 P_{t,1} V_2^* + (I - V_2 V_2^*)] \\ &= V_2 P_{t,1} V_2^* + (I - V_2 V_2^*) = Q_{t,1}. \end{aligned}$$

□

To motivate the concept of quantum probabilities, we will talk about finite input-output games in the next subsection.

9.4. **Finite Input-Output Games.** Consider a two-person game in which there are two players Alice and Bob and a referee R. Let I_A, I_B be finite input sets and $\mathcal{O}_A, \mathcal{O}_B$ be finite output sets. Every game has got some rules. The "rules" for this game can be described by a function

$$\lambda : I_A \times I_B \times \mathcal{O}_A \times \mathcal{O}_B \to \{0, 1\} \text{ where,}$$

$$\lambda(a, b, x, y) = \begin{cases} 0 \text{ means that the move is disallowed} \\ 1 \text{ means that the move is allowed.} \end{cases}$$

Alice, Bob and the referee are all aware of the "rules" (and hence the function λ) of this game. Now, Alice and Bob are going to play the game together, against the referee, but they are not allowed to communicate during the game. They are allowed to collaborate to decide any kind of strategy before the game begins. However, once the game begins, they are not allowed to communicate with each other.

One Round The game begins when the referee R gives Alice an element $a \in I_A$ and Bob an element $b \in I_B$ (Alice does not know what Bob has been given, and Bob does not know what Alice has been given) and each of the players (Alice and Bob) then "independently" produces an output, say $x \in \mathcal{O}_A$ and $y \in \mathcal{O}_B$. They "win" the game if $\lambda(a, b, x, y) = 1$ and loose if $\lambda(a, b, x, y) = 0$.

Multiple Round The referee R selects $a \in I_A$ and $b \in I_B$ and gives these to Alice and Bob respectively and they independently produce outputs $x \in \mathcal{O}_A$ and $y \in \mathcal{O}_B$ respectively. Alice and Bob "win" the game if in every round $\lambda(a, b, x, y) = 1$ and loose otherwise.

Recall that Alice and Bob are allowed to communicate before the beginning of the game and can come up with some sort of strategy to win the game. A strategy can be deterministic (so that they always win the game) or probabilistic (so that they win the game with certain probability).

Definition 9.9. A *deterministic strategy* is a pair of functions $f : I_A \to \mathcal{O}_A$ and $g : I_B \to \mathcal{O}_B$ such that for all $a \in I_A$ and $b \in I_B$, we have $\lambda(a, b, f(a), g(b)) = 1$ (so they can always win).

Of course, not all games have a deterministic strategy. So, more generally, suppose we have a probability density $\pi : I_A \times I_B \to [0, 1]$ (that is, $\pi(a, b) \geq 0$ and $\sum_{a,b} \pi(a, b) = 1$). For a pair of functions $f : I_A \to \mathcal{O}_A$ and $g : I_B \to \mathcal{O}_B$ and the probability density π, the *value* of (f, g) is given by the expression $\sum_{a,b} \pi(a, b)\lambda(a, b, f(a), g(b))$. It is easy to see that,

$$\text{the value of } (f, g) \text{ with respect to } \pi = \sum_{a,b} \pi(a, b)\lambda(a, b, f(a), g(b)) \leq 1.$$

Note that;

(1) The equality holds (that is, the value of (f, g) with respect to π is 1) $\iff$ for all a, b we have $\lambda(a, b, f(a), g(b)) = 1$ whenever $\pi(a, b) > 0$.
(2) Suppose $\pi(a, b) > 0$ for all a, b. Then the equality holds $\iff$ $\lambda(a, b, f(a), g(b)) = 1$ for all a, b. This is equivalent to say that (f, g) is a deterministic strategy.

Remark 9.10. A game $\mathcal{G}$ can be determined by $\mathcal{G} := (I_A, I_B, \mathcal{O}_A, \mathcal{O}_B, \lambda)$. Mostly we will be dealing with games where $I_A = I_B$ and $\mathcal{O}_A = \mathcal{O}_B$.

Example 9.11 (Graph Colouring Game)**.** Given a graph $G = (V, E)$, we define $I_A = I_B = V$ and $\mathcal{O}_A = \mathcal{O}_B$ to be a set of colours. In graph colouring game, two players, Alice and Bob, try to convince a referee

R that they have a colouring of the graph G. So, the rule of the game can be defined as follows:

$$\lambda(v,w,a,b) = \begin{cases} 1 \text{ if } a \neq b \\ 0 \text{ if } a = b \end{cases}, \text{ if } v \sim w;$$

$$\lambda(v,v,a,b) = \begin{cases} 1 \text{ if } a = b \\ 0 \text{ if } a \neq b \end{cases}, \text{ and}$$

$$\lambda(v,w,a,b) = 1 \text{ if } v \not\sim w \text{ and } v \neq w.$$

Notice that the last equation asserts that if the vertices are neither identical nor adjacent, then Alice and Bob can respond with any colours and hence in such case any move is allowed. For this game, if $|\mathcal{O}_A| \geq \chi(G)$, then before the game, Alice and Bob can find a colouring of the graph. This gives a function $f = g : V \to \mathcal{O}_A = \mathcal{O}_B$ such that for each $v, w \in V = \mathcal{O}_A$, we have $\lambda(v, w, f(v), g(w)) = 1$. Hence, we have a deterministic strategy.

Definition 9.12 (Probabilistic Strategy)**.** Given a finite input-output game as above, a *probabilistic strategy* is a conditional probability density $p(x, y|a, b)$, the probability that Alice and Bob produce outcome x and y respectively, given that Alice receives a and Bob receives b.

Clearly then $p(x, y|a, b) \geq 0$ and for all $a \in I_A$ and $b \in I_B$, we have

$$\sum_{x \in \mathcal{O}_A,\, y \in \mathcal{O}_B} p(x, y|a, b) = 1.$$

Definition 9.13 (Perfect/Winning Strategy)**.** We call $p(x, y|a, b)$ a *perfect* or *winning strategy* provided that whenever $\lambda(a, b, x, y) = 0$, we have $p(x, y|a, b) = 0$. Given a strategy $p(x, y|a, b)$ and a density on inputs $\pi : I_A \times I_B \to [0, 1]$, the *value* of the strategy is

$$\sum_{x,y,a,b} \pi(a, b)\lambda(a, b, x, y)p(x, y|a, b).$$

Note that $\sum_{x,y,a,b} \pi(a, b)p(x, y|a, b) = \sum_{a,b} \pi(a, b) = 1$, so the value of the strategy is always at most 1. If $\pi(a, b) > 0$ for all a, b, then the value of the strategy is 1 if and only if $p(x, y|a, b)$ is perfect.

There are two natural questions we can ask. Given some conditions on the allowable $p(x, y|a, b)$,

(1) Decide whether there exists a perfect strategy.
(2) If not, find the maximal value, or the supremum of the values of strategies over all allowed probabilities $p(x, y|a, b)$.

We have the following goals:

(1) What are some "natural" families of probability densities?
(2) Show that "classical" densities are not as good as "quantum densities".
(3) Consider different mathematical models for what are "quantum probability densities".

9.5. **Classical Densities.** We suppose that Alice and Bob share some probability space (Ω, μ) and that for each input $a \in I_A$, Alice has a function $f_a : \Omega \to \mathcal{O}_A$ such that $\mu(\{w : f_a(w) = x\})$ is the probability that Alice produces output x, given that she received input a. Similarly, we suppose that for each $b \in I_B$ Bob has a function $g_b : \Omega \to \mathcal{O}_B$ such that $\mu(\{w : g_b(w) = y\}$ is the probability that Bob produces output y, given that he received input b. In this case, we have

$$p(x, y|a, b) = \mu(\{w : f_a(w) = x,\ g_b(w) = y\}).$$

The set of all such $p(x, y|a, b)$ is called the set of all *local densities*. When $I_A = I_B, \mathcal{O}_A = \mathcal{O}_B, |I_A| = n$ and $|\mathcal{O}_B| = k$, then these are n^2k^2-tuples of real numbers. We denote this set by $LOC(n, k) = C_{\text{loc}}(n, k)$.

9.6. **Quantum Densities.** The idea is that for each input, Alice has a different experiment with $|\mathcal{O}_A|$ outcomes. There is a state space $\mathcal{H}_A(\dim \mathcal{H}_A < \infty)$, and for each $a \in I_A$, there is a POVM, $\{E_{a,x}\}_{x\in\mathcal{O}_A}$ (so $E_{a,x} \geq 0$ and $\sum_x E_{a,x} = I$), such that if we are in state $\psi \in \mathcal{H}_A$, then $\langle\psi|E_{a,x}\psi\rangle = p_A(x|a)$. Similarly, for each input $b \in I_B$, Bob also has a quantum experiment with $|\mathcal{O}_B|$ outcomes. These correspond to POVM's $\{F_{b,y}\}_{y\in\mathcal{O}_B}$ (where $F_{b,y} \geq 0$ and $\sum_y F_{b,y} = I$) on $\mathcal{H}_B(\dim \mathcal{H}_B < \infty)$.

The strategy is the following: Pick a state $\psi \in \mathcal{H}_A \otimes \mathcal{H}_B$ (often this will be entangled) such that $p(x, y|a, b) = \langle\psi|E_{a,x} \otimes F_{y,b}(\psi)\rangle$. The set of all such tuples when $I_A = I_B, |I_A| = n, \mathcal{O}_A = \mathcal{O}_B$, and $|\mathcal{O}_B| = k$ is denoted by $Q(n, k) = C_q(n, k)$.

We show that

(1) $C_{\text{loc}}(n, k) \subsetneq C_q(n, k)$ (both are contained in $\mathbb{R}^{n^2k^2}$).
(2) Many games that do not have perfect loc strategies do have perfect quantum strategies.

Other Versions of Quantum Densities There is more than one definition of a quantum density. Other versions include (but may not be limited to):

(1) The same as above except drop the requirement of $\mathcal{H}_A$ and $\mathcal{H}_B$ being finite-dimensional, that is, drop the condition $\dim(\mathcal{H}_A), \dim(\mathcal{H}_B) < \infty$. The resulting set is denoted by $C_{qs}(nk)$.

(2) Another axiomatic description is that there is a universal state space $\mathcal{H}$ and all of the $E_{a,x}, F_{b,y}$ act on this space and commute with each other, for all choices of a, x, b, y. (This reflects the fact that the labs are different. Indeed, in the case of different labs, the operators would look like $E_{a,x} \otimes I$ and $I \otimes F_{b,y}$ and these always commute.) In this case, there is a state $\psi \in \mathcal{H}$ such that $p(x, y|a, b) = \langle \psi | E_{a,x} F_{b,y} \psi \rangle$. The set of all such quantum densities will be denoted by $C_{qc}(n, k)$.

It is clear that

$$C_{\text{loc}}(n, k) \subseteq C_q(n, k) \subseteq C_{qs}(n, k) \subseteq C_{qc}(n, k) \subseteq \mathbb{R}^{n^2 k^2}.$$

Strong Tsirelson Conjecture The *Strong (Bivariate) Tsirelson Conjecture* is that

$$C_q(n, k) = C_{qc}(n, k), \text{ for all } n, k.$$

This problem is still open.

Since it is not known whether $C_q(n, k)$ is closed or not, let us define

$$C_{qa}(n, k) := \overline{C_q(n, k)}.$$

Next, suppose we had the commuting model with all the $E_{a,x}$'s and $F_{b,y}$'s being projections, and $\psi \in \mathcal{H}$ with $\|\psi\| = 1$, then the vectors $v_{a,x} = E_{a,x}\psi$ satisfy

$$v_{a,x} \perp v_{a,x'} \text{ for } x \neq x', \text{ while } \sum_x v_{a,x} = \left(\sum_x E_{a,x} \right) \psi = \psi.$$

Similarly, with $w_{b,y} = F_{b,y}\psi$ we have

$$w_{b,y} \perp w_{b,y'} \text{ for } y \neq y', \text{ while } \sum_y w_{b,y} = \psi.$$

Moreover, since $E_{a,x}$'s and $F_{b,y}$'s commute, we have,

$$\langle v_{a,x} | w_{b,y} \rangle = \langle E_{a,x}\psi | F_{b,y}\psi \rangle = \langle \psi | E_{a,x} F_{b,y} \psi \rangle = p(x, y|a, b) \geq 0.$$

Definition 9.14. The set $C_{\text{vect}}(n, k)$ is the set of all probability densities $p(x, y|a, b)$ of the form $p(x, y|a, b) = \langle v_{a,x} | w_{b,y} \rangle$ for some set of vectors $\{v_{a,x}, w_{b,y}\}$ satisfying

$$\sum_x v_{a,x} = \sum_y w_{b,y} := \psi, \, \forall a, b$$

where

- ψ is some vector in $\mathcal{H}$ with $\|\psi\| = 1$,
- $v_{a,x} \perp v_{a,x'}$ for $x \neq x'$,
- $w_{b,y} \perp w_{b,y'}$ for $y \neq y'$, and
- $\langle v_{a,x}, w_{b,y} \rangle \geq 0$.

Definition 9.15. We define $NSB(n,k) = C_{nsb}(n,k)$ as the set of all n^2k^2-tuples $p(x,y|a,b)$ such that

(1) $p(x,y|a,b) \geq 0$ for all x, y, a, b.
(2) $\sum_{x,y} p(x,y|a,b) = 1$.
(3) $\sum_y p(x,y|a,b) = \sum_y p(x,y|a,b')$, for all b, b'. (We denote this common value by $P_A(x|a)$.)
(4) $\sum_x p(x,y|a,b) = \sum_x p(x,y|a',b)$ for all a, a'. (We denote this common value by $P_B(y|b)$.

Probability densities with these properties are called *non-signalling boxes*. Axiom 3 and 4 are non-signalling conditions. Indeed, suppose we had $\sum_y p(x,y|a,b) \neq \sum_y p(x,y|a,b')$ for some x, b, b'. This means that if Alice runs experiment a and computed the probability of getting the outcome x, then she would get different probabilities depending on whether Bob runs experiment b or b'. This means that Bob could "signal" which experiment he ran, to Alice. (In most models, the axioms don't allow this to happen)

Theorem 9.16. *For each $k, n \in \mathbb{N}$, we have the sequence of inclusions*

$$C_{loc}(n,k) \subseteq C_q(n,k) \subseteq C_{qs}(n,k) \subseteq C_{qa}(n,k) \subseteq C_{qc}(n,k) \subseteq C_{vect}(n,k) \subseteq C_{nsb}(n,k) \subseteq \mathbb{R}^{n^2k^2}.$$

Next suppose we are given a finite input-output game $\mathcal{G} = (I_A, I_B, \mathcal{O}_A, \mathcal{O}_B, \lambda)$ where $|I_A| = |I_B| = n$, $|\mathcal{O}_A| = |\mathcal{O}_B| = k$, $\lambda : I_A \times I_B \times \mathcal{O}_A \times \mathcal{O}_B \to \{0,1\}$ and $t \in \{\text{loc}, q, qs, qa, qc, \text{vect}, nsb\}$, we say that $\mathcal{G}$ has a *perfect t-strategy* if there exists $p \in C_t(n,k)$ such that $\lambda(a,b,x,y) = 0 \implies p(x,y|a,b) = 0$.

Given a probability density $\pi : I_A \times I_B \to [0,1]$ and t as above, the *t-value* of $\mathcal{G}$ is defined as

$$\omega_t(\mathcal{G}, \pi) = \sup \left\{ \sum_{x,y,a,b} \pi(a,b) p(x,y|a,b) \lambda(a,b,x,y) : p \in C_t(n,k) \right\}.$$

The idea is that we could distinguish among the sets $C_t(n,k)$ by either finding games with perfect strategies for one t but not another t, or we

could show that the value of the game depends (non-trivially) on the choice of t.

Theorem 9.17 (Ozawa)**.** *Connes Embedding Conjecture (from operator algebras) is true if and only if $C_{qa}(n,k) = C_{qc}(n,k)$ for all $k, n \in \mathbb{N}$.*

Remark 9.18. It is not known if $C_q(n,k) = C_{qa}(n,k)$ (that is, we are not sure if $C_q(n,k)$ is closed). It is known that $C_{qc}(n,k) \neq C_{\text{vect}}(n,k)$.

We now will move towards showing that $C_{\text{loc}}(2,2) \neq C_q(2,2)$ (so these sets are different even in the small cases). We first need the Clauser-Horne-Shimony-Holt inequality.

Let $I_A = \{a, a'\}$ and $I_B = \{b, b'\}$. Let $\mathcal{O}_A = \mathcal{O}_B = \{+, -\}$, and let $p \in C_{\text{loc}}(2,2) \subseteq \mathbb{R}^{2^2 \cdot 2^2} = \mathbb{R}^{16}$, say $p(\pm, \pm | r, s)$ for $r = a, a'$ and $s = b, b'$. Set $\sigma_{a,b}(p) = p(+,+|a,b) + p(-,-|a,b) - p(+,-|a,b) - p(-,+|a,b)$. Given $f_a, f_{a'}, g_b, g_{b'} : \Omega \to \{-1, +1\}$ (where (Ω, μ) is a probability space) such that $p(+,+|a,b) = \mu(\{t : f_a(t) = 1,\, g_b(t) = -1\})$ and $p(-,-|a,b) = \mu(\{t : f_a(t) = -1,\, g_b(t) = -1\})$, then

$$\begin{aligned}\int_\Omega f_a(t) g_b(t)\, d\mu(t) &= p(+,+|a,b) + p(-,-|a,b) - p(+,-|a,b) - p(-,+|a,b) \\ &= \sigma_{a,b}(p).\end{aligned}$$

Theorem 9.19 (CHSH-inequality)**.** *Let $p \in C_{loc}(2,2)$. Then*

$$-2 \leq \sigma_{a,b}(p) + \sigma_{a,b'}(p) + \sigma_{a',b}(p) - \sigma_{a',b'}(p) \leq 2.$$

Proof. By the above work, letting $M = \sigma_{a,b}(p) + \sigma_{a,b'}(p) + \sigma_{a',b}(p) - \sigma_{a',b'}(p)$, we have

$$M = \int_\Omega (f_a(t)[g_b(t) + g_{b'}(t)] + f_{a'}(t)[g_b(t) - g_{b'}(t)])\, d\mu(t).$$

Note that $g_b(t) + g_{b'}(t) \in \{-2, 2\}$ if and only if $g_b(t) - g_{b'}(t) = 0$ (both take values in $\{-1, 1\}$), while $g_{b'}(t) - g_{b'}(t) \in \{-2, 2\}$ if and only if $g_b(t) + g_{b'}(t) = 0$. Looking the integrand and noting that $f_a(t), f_{a'}(t) \in \{-1, 1\}$, it follows that

$$-2 \leq f_a(t)[g_b(t) + g_{b'}(t)] + f_{a'}(t)[g_b(t) - g_{b'}(t)] \leq 2,\ \forall t.$$

Since μ is a probability measure, integrating gives $-2 \leq M \leq 2$, as desired.

□

Now let us look at some special elements of $C_q(2,2)$. Let $\psi = \frac{1}{\sqrt{2}}(e_0 \otimes e_0 + e_1 \otimes e_1) \in \mathbb{C}^2 \otimes \mathbb{C}^2$. Given an angle θ, let $v_\theta = (\cos\theta, \sin\theta)^t$. Then define

$$P_\theta = v_\theta v_\theta^* = \begin{pmatrix} \cos^2(\theta) & \cos\theta \sin\theta \\ \cos\theta \sin\theta & \sin^2(\theta) \end{pmatrix}.$$

It follows that $I - P_\theta = P_\theta^\perp = \begin{pmatrix} \sin^2(\theta) & -\cos(\theta)\sin(\theta) \\ -\cos(\theta)\sin(\theta) & \cos^2(\theta) \end{pmatrix}$. Given $\theta_a, \theta_b, \theta_{a'}, \theta_{b'}$, define $E_{a,+} = P_{\theta_a}$, $E_{a,-} = P_{\theta_a}^\perp$, $E_{a',+} = P_{\theta_{a'}}$, $E_{a',-} = P_{\theta_{a'}}^\perp$, $F_{b,+} = P_{\theta_b}$, $F_{b,-} = P_{\theta_b}^\perp$, $F_{b',+} = P_{\theta_{b'}}$, $F_{b',-} = P_{\theta_{b'}}^\perp$. Then the probability arising from this for $+, +, a, b$ is

$$\begin{aligned} p(+,+|a,b) &= \langle\psi|(E_{a,+} \otimes F_{b,+}\psi\rangle \\ &= \frac{1}{2}\langle e_0 \otimes e_0 + e_1 \otimes e_1 | (\cos^2\theta_a e_0 + \cos\theta_a \sin\theta_a e_1) \otimes (\cos^2\theta_b e_0 + \cos\theta_b \sin\theta_b e_1) \\ &\quad + (\cos\theta_a \sin\theta_a e_0 + \sin^2(\theta_a) e_1) \otimes (\cos\theta_b \sin\theta_b e_0 + \sin^2\theta_b e_1)\rangle \\ &= \frac{1}{2}[\cos^2\theta_a \cos^2\theta_b + 2\cos\theta_a \sin\theta_a \cos\theta_b \sin\theta_b + \sin^2\theta_a \sin^2\theta_b] \\ &= \frac{1}{2}(\cos\theta_a \cos\theta_b + \sin\theta_a \sin\theta_b)^2 \\ &= \frac{1}{2}\cos^2(\theta_a - \theta_b). \end{aligned}$$

Through a similar calculation we have $p(-,-|a,b) = \frac{1}{2}\cos^2(\theta_a - \theta_b)$, while

$$p(+,-|a,b) = p(-,+|a,b) = \frac{1}{2}\sin^2(\theta_a - \theta_b).$$

It follows that

$$\sigma_{a,b}(p) = \cos^2(\theta_a - \theta_b) - \sin^2(\theta_a - \theta_b) = \cos(2(\theta_a - \theta_b)).$$

Hence, we have

$$\begin{aligned} \sigma_{a,b}(p) + \sigma_{a,b'}(p) + \sigma_{a',b}(p) + \sigma_{a',b} - \sigma_{a',b'}(p) &= \cos(2\theta_a - 2\theta_b) + \cos(2\theta_a - 2\theta_{b'}) \\ &\quad + \cos(2\theta_{a'} - 2\theta_b) - \cos(2\theta_{a'} - 2\theta_b). \end{aligned}$$

Set $\theta_{a'} = \frac{\pi}{4}$, $\theta_{b'} = 0$ and $\theta_a = \theta_b = \frac{\pi}{8}$. Then the above expression becomes

$$\cos(0) + \cos\left(\frac{\pi}{4}\right) + \cos\left(\frac{\pi}{4}\right) - \cos\left(\frac{\pi}{2}\right) = 1 + \sqrt{2} > 2.$$

This violates the CHSH inequality, so that $p \notin C_{\text{loc}}(2,2)$. It follows that $C_{\text{loc}}(2,2) \subsetneq C_q(2,2)$.

Problem 9.20. *Find and justify*

$$\max\{\cos(\theta_a - \theta_b) + \cos(\theta_a - \theta_{b'}) + \cos(\theta_{a'} - \theta_b) - \cos(\theta_{a'} - \theta_{b'}) : \theta_a, \theta_b, \theta_{a'}, \theta_{b'}\}.$$

Proof. **Homework problem 16**; due 15th March, Tuesday. □

Next recall that $p(x,y|a,b) \in C_{nsb}$ if and only if $\sum_y p(x,y|a,b) = \sum_y p(x,y|a,b') := p_A(x|a)$ and $\sum_x p(x,y|a,b) = \sum_x p(x,y|a',b) = p_B(y|b)$.

Let $p \in C_{nsb}(2,2)$ with outcomes $+,-$ be given by the probabilities below:

$$\left[\begin{array}{ccccccccccc}
 & & & & & b & & & & b' & \\
 & & & & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
 & & & & + & & - & & + & & - \\
 & & & & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
 & \vdots & + & \vdots & 1/2 & & 0 & \vdots & 1/2 & & 0 \\
a & \vdots & & \vdots & & & & \vdots & & & \\
 & \vdots & - & \vdots & 0 & & 1/2 & \vdots & 0 & & 1/2 \\
 & \vdots & & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
 & \vdots & + & \vdots & 1/2 & & 0 & \vdots & 0 & & 1/2 \\
a' & \vdots & & \vdots & & & & \vdots & & & \\
 & \vdots & - & \vdots & 0 & & 1/2 & \vdots & 1/2 & & 0
\end{array}\right]$$

This is an element of $C_{nsb}(2,2)$. One can see that

$$\sigma_{a,b} = p(+,+|a,b) + p(-,-|a,b) - p(+,-|a,b) - p(+,-|a,b) = 1,$$

while $\sigma_{a,b'} = 1$, $\sigma_{a',b} = 1$ and $\sigma_{a',b'} = -1$. Examining the quantity in the CHSH inequality, we have

$$\sigma_{a,b} + \sigma_{a,b'} + \sigma_{a',b} - \sigma_{a',b'} = 4.$$

Problem 9.21. *Prove that if $p \in C_{vect}(2,2)$, then*

$$\sigma_{a,b} + \sigma_{a,b'} + \sigma_{a',b} - \sigma_{a',b'} < 4,$$

and conclude that $C_{nsb}(2,2) \neq C_{vect}(2,2)$. (Hint: Prove by the method of contradiction, i.e., assume that one of them has quantity 4 above and arrive at a contradiction.)

Proof. **Homework problem 17**; due 15th March, Tuesday. □

Problem 9.22. *Prove that $C_{vect}(n,k) \subseteq C_{nsb}(n,k)$ for every n,k.*

Proof. **Homework problem 18**; due 15th March, Tuesday. □

We stated in Theorem **??** that for each $k,n \in \mathbb{N}$, we have the sequence of inclusions

$$\begin{aligned} C_{\text{loc}}(n,k) \subseteq C_q(n,k) \subseteq C_{qs}(n,k) \subseteq C_{qa}(n,k) \\ \subseteq C_{qc}(n,k) \subseteq C_{\text{vect}}(n,k) \subseteq C_{nsb}(n,k) \subseteq \mathbb{R}^{n^2k^2}. \end{aligned}$$

We will try and prove some of these inclusions. But before that let's recall and state the key conjectures as problems:

(1) **(strong) Tsirelson's Problem**: Is $C_q(n,k) = C_{qc}(n,k)$ for all n, k?
(2) **Connes Problem**: Is $C_{qa}(n,k) = C_{qc}(n,k)$ for all n, k?
(3) **Closure Problem**: Is $C_q(n,k) = C_{qa}(n,k)$ for all n, k?
(4) In an unpublished preprint, Werner-Scholze claimed that for all n, k; $C_{qs}(n,k) = C_{qa}(n,k)$. However, the proof is not established to be correct. So, we have:
Werner-Scholze Problem: Is $C_{qs}(n,k) = C_{qa}(n,k)$ for all n, k?

Proposition 9.23. *For any $n, k \in \mathbb{N}$, $C_{loc}(n,k) \subseteq C_{qc}(n,k)$.*

Proof. Let $f_a, g_b : \Omega \to \{1, ..., k\}$ for $1 \leq a, b \leq m$ be such that $p(x,y|a,b) = \mu(\{t \in \Omega : f_a(t) = x,\, g_b(t) = y\})$ for all x, y. Let us define

$$\Omega_{a,i} = \{t : f_a(t) = i\}$$

and

$$\Omega'_{b,j} = \{t : g_b(t) = j\}.$$

We observe that

$$\bigcup_{i=1}^{k} \Omega_{a,i} = \bigcup_{j=1}^{k} \Omega'_{b,j} = \Omega.$$

We further define

$$\chi_{a,i}(t) = \begin{cases} 1 \text{ if } t \in \Omega_{a,i} \\ 0 \text{ if } t \notin \Omega_{a,i}. \end{cases}$$

Similarly, we define

$$\chi'_{b,j}(t) = \begin{cases} 1 \text{ if } t \in \Omega'_{b,j} \\ 0 \text{ if } t \notin \Omega'_{b,j}. \end{cases}$$

Note that

$$p(i,j|a,b) = \mu(\Omega_{a,i} \cap \Omega'_{b,j}) = \int_{\Omega_{a,i} \cap \Omega'_{b,j}} 1 \, d\mu = \int_{\Omega} \chi_{a,i} \chi'_{b,j} \, d\mu.$$

We then define $\mathcal{H} = L^2(\Omega, \mu)$. For this Hilbert space let $E_{a,i} = M_{\chi_{a,i}}$, which denotes multiplication by $\chi_{a,i}$ on $L^2(\Omega, \mu)$. Then $\sum_{i=1}^{k} E_{a,i}$ is multiplication by $\sum_{i=1}^{k} \chi_{a,i} = 1$, so that $\sum_{i=1}^{k} E_{a,i} = M_{\sum_{i=1}^{k} E_{a,i}} =$

$M_1 = I$. Similarly, if $F_{b,j} = M_{\chi'_{b,j}}$, then $\sum_{j=1}^{k} F_{b,j} = I$. Whenever $h \in \mathcal{H}$, we have

$$\langle h|E_{a,i}h\rangle = \int_\Omega \overline{h}(E_{a,i}h)\,d\mu = \int_{\Omega_{a,i}} |h|^2\,d\mu \geq 0,$$

so that $E_{a,i} \geq 0$. It is easy to see that $E_{a,i}^2 = M_{\chi_{a,i}^2} = M_{\chi_{a,i}} = E_{a,i}$ so each $E_{a,i}$ is a projection (similarly, the $F_{b,j}$'s are projections). All of these operators ($E_{a,i}$'s, $F_{b,j}$'s), being multiplication operators, commute with each other (and hence each $E_{a,i}$ commutes with each $F_{b,j}$). Finally, we set $\psi = 1 \in \mathcal{H}$. Then $\|\psi\|^2 = \int_\Omega 1\,d\mu = \mu(\Omega) = 1$ so ψ is a state. We observe that for each a, b, i, j, we have

$$\begin{aligned}\langle \psi|E_{a,i}F_{b,j}\psi\rangle &= \int_\Omega \overline{1}(M_{\chi_{a,i}}M_{\chi'_{b,j}} \cdot 1)\,d\mu \\ &= \int_{\Omega_{a,i}\cap\Omega'_{b,j}} 1\,d\mu \\ &= \mu(\Omega_{a,i} \cap \Omega'_{b,j}) = p(i,j|a,b).\end{aligned}$$

It follows that $C_{\text{loc}}(n,k) \subseteq C_{qc}(n,k)$. □

9.7. **Disambiguation Theorems.** In literature, some authors define the sets $C_q(n,k)$, $C_{qc}(n,k)$ using POVM's; others define using PVM's. In this subsection, we aim to show that no matter what is used in defining these, we end up getting the same sets both ways.

For the sake of proof, let us write $C_q(n,k)$ and $C_{qc}(n,k)$ to be the sets when we use POVM's in the definition, and let $\tilde{C}_q(n,k)$ and $\tilde{C}_{qc}(n,k)$ be the sets when we use PVM's in the definition. Since every PVM is also a POVM, we have $\tilde{C}_q(n,k) \subseteq C_q(n,k)$ and $\tilde{C}_{qc}(n,k) \subseteq C_{qc}(n,k)$.

Proposition 9.24. $\tilde{C}_q(n,k) = C_q(n,k)$.

Proof. Let $p(x,y|a,b) \in C_q(n,k)$. Then there exist $E_{a,x} \geq 0$ on $\mathcal{H}_A$ with $\dim(\mathcal{H}_A) < \infty$ and $\sum_x E_{a,x} = I$, and $F_{b,y} \geq 0$ on $\mathcal{H}_B$ with $\sum_y F_{b,y} = I$ (and $\dim(\mathcal{H}_B) < \infty$) and $\psi \in \mathcal{H}_A \otimes \mathcal{H}_B$ such that $p(x,y|a,b) = \langle\psi|(E_{x,a} \otimes F_{y,b})\psi\rangle$ for each x, y, a, b.

By the dilation theorem, there is a Hilbert space $\tilde{\mathcal{H}}_A$ that is finite-dimensional and projections $P_{a,x}$ on $\tilde{\mathcal{H}}_A$ such that $\sum_x P_{a,x} = I$ and an isometry $V_A : \mathcal{H}_A \to \tilde{\mathcal{H}}_A$ such that $E_{x,a} = V_A^* P_{x,a} V_A$ for all x, a. One can do the same for Bob and obtain a Hilbert space $\tilde{\mathcal{H}}_B$ of finite dimension, along with projections $Q_{b,y}$ such that $\sum_y Q_{b,y} = I$ and an isometry $V_B : \mathcal{H}_B \to \tilde{\mathcal{H}}_B$ such that $V_B^* Q_{b,y} V_B = F_{b,y}$ for all b, y. Now

define $\varphi = (V_A \otimes V_B)(\psi) \in \tilde{\mathcal{H}}_A \otimes \tilde{\mathcal{H}}_B$. Then $\langle \varphi, P_{a,x} \otimes Q_{b,y} \varphi \rangle \in \tilde{C}_q(n,k)$, while

$$\begin{aligned} \langle \varphi, P_{a,x} \otimes Q_{b,y} \varphi \rangle &= \langle (V_A \otimes V_B)\psi, (P_{a,x} \otimes Q_{b,y})(V_A \otimes V_B)\psi \rangle \\ &= \langle \psi | (V_A^* P_{a,x} V_A) \otimes (V_B^* Q_{b,y} V_B) \psi \rangle \\ &= p(x,y|a,b). \end{aligned}$$

Therefore, $C_q(n,k) \subseteq \tilde{C}_q(n,k)$. □

Proposition 9.25. $\tilde{C}_{qs}(n,k) = C_{qs}(n,k)$.

Proof. The proof is the same as above, except that we may not have $\mathcal{H}_A$ and $\mathcal{H}_B$ finite-dimensional. We hence assume that $\mathcal{H}_A$ and $\mathcal{H}_B$ are not necessarily finite dimensional and mimick the proof. Interestingly, the spaces $\tilde{\mathcal{H}}_A$ and $\tilde{\mathcal{H}}_B$ will still exist, and the rest of the proof will be the same. □

Theorem 9.26 (Paulsen-Todorov)**.** $\tilde{C}_{qc}(n,k) = C_{qc}(n,k)$

Proof. We don't prove this theorem as it uses Boca's theorem (about completely positive maps on free products of C^*-algebras), along with stuff about group C^*-algebras of free products. □

This leads us to,

Corollary 9.27. $C_{qc}(n,k) \subseteq C_{vect}(n,k)$.

Proof. Let $p(x,y|a,b) \in C_{qc}(n,k)$. Then Theorem **??** guarantees the existence of projections $P_{a,x}$ and $Q_{b,y}$ on a Hilbert space $\mathcal{H}$ with $\sum_x P_{a,x} = \sum_y Q_{b,y} = I$ such that $P_{a,x}Q_{b,y} = Q_{b,y}P_{a,x}$ on $\mathcal{H}$ for all a,b,x,y. Moreover, there is $\psi \in \mathcal{H}$ with $\|\psi\| = 1$ such that $p(x,y|a,b) = \langle \psi | P_{a,x}Q_{b,y}\psi \rangle$. Set $h_{a,x} = P_{a,x}\psi$ and $k_{b,y} = Q_{b,y}\psi$. Then $\sum_x h_{a,x} = \psi$ for all a and $\sum_y k_{b,y} = \psi$ for all b. Since $P_{a,x}P_{a,x'} = 0$ (because they are orthogonal to each other as projections), we have $h_{a,x} \perp h_{a,x'}$ for $x \neq x'$, and similarly $k_{b,y} \perp k_{b,y'}$ for $y \neq y'$.

Finally, $\langle h_{a,x} | k_{b,y} \rangle \in C_{\text{vect}}(n,k)$ but $\langle h_{a,x} | k_{b,y} \rangle = \langle P_{a,x}\psi | Q_{b,y}\psi \rangle$, but this is equal to $\langle \psi | P_{a,x}Q_{b,y}\psi \rangle = p(x,y|a,b)$. Therefore, $C_{qc}(n,k) \subseteq C_{\text{vect}}(n,k)$. □

Notice that in the proof above, the product $P_{a,x}P_{a,x'}$ had no reason to vanish if we have had merely positive operators. This is where we made use of the fact that these are actually orthogonal projections.

We end this subsection by proving an important theorem for finite input-output games.

Theorem 9.28. *Let $\mathcal{G} = (I_A, I_B, \mathcal{O}_A, \mathcal{O}_B, \lambda)$ where λ is the rule function as before. Then $\mathcal{G}$ has a perfect loc-strategy if and only if $\mathcal{G}$ has a perfect deterministic strategy.*

Proof. If $\mathcal{G}$ has a perfect deterministic strategy, then this means that there are functions $f : I_A \to \mathcal{O}_A$ and $g : I_B \to \mathcal{O}_B$ such that $\lambda(a, b, f(a), g(b))$ never violates the rule, which implies that $\lambda(a, b, f(a), g(b)) = 1$ for all a, b. (The reason we are not able to see the probability space is because it has got only one point!) Let $\Omega = \{t_0\}$ and $f_a : \Omega \to \mathcal{O}_A$ and $g_b : \Omega \to \mathcal{O}_B$ be given by $f_a(t_0) = f(a)$ and $g_b(t_0) = g(b)$. Suppose that $\lambda(a, b, x, y) = 0$. We must show that $\mu(\{t : f_a(t) = x,\ g_b(t) = y\}) = 0$. Note that for t_0 we have $\lambda(a, b, f_a(t_0), g_b(t_0)) = \lambda(a, b, f(a), g(b)) = 1$ so that $\{t : f_a(t) = x,\ g_b(t) = y\} = \emptyset$, and has measure zero. It follows that $\mathcal{G}$ has a perfect loc-strategy.

Conversely, suppose that $\mathcal{G}$ has a perfect loc-strategy. Then there is a probability space (Ω, μ) and functions $f_a : \Omega \to \mathcal{O}_A$ and $g_b : \Omega \to \mathcal{O}_B$ such that whenever $\lambda(a, b, x, y) = 0$, we have $\mu(\{t : f_a(t) = x,\ g_b(t) = y\}) = 0$. Set $\Omega_{a,x} = \{t : f_a(t) = x\}$. These sets are pairwise disjoint with union equal to Ω. Similarly, if $\Omega'_{b,y} = \{t : g_b(t) = y\}$, then the sets $\{\Omega'_{b,y}\}$ are pairwise disjoint with union equal to Ω. Then $\mu(\Omega_{a,x} \cap \Omega'_{b,y}) = 0$. Define

$$\mathcal{N} = \bigcup_{\substack{a,b,x,y \\ \lambda(a,b,x,y)}} (\Omega_{a,x} \cap \Omega'_{b,y}),$$

which is a finite union of sets of measure zero, so that $\mu(\mathcal{N}) = 0$. Therefore, $\mathcal{N} \neq \Omega$. Pick $t_0 \in \Omega \setminus \mathcal{N}$. Define $f : I_A \to \mathcal{O}_A$ and $g : I_B \to \mathcal{O}_B$ by $f(a) = f_a(t_0)$ and $g(b) = g_b(t_0)$. By definition, since $t_0 \notin \mathcal{N}$, we must have $t_0 \notin \Omega_{a,x} \cap \Omega'_{b,y}$ for all a, x, b, y. Therefore, if we had $\lambda(a, b, f(a), g(b)) = 0$ for some a, b, then with $x = f_a(t_0)$ and $y = g_b(t_0)$, we would have $\lambda(a, b, x, y) = 0$. Hence, $t_0 \in \Omega_{a,x} \cap \Omega'_{b,y}$ which is a contradiction. Hence, $\lambda(a, b, f(a), g(b)) = 1$ for all a, b. Therefore, $\mathcal{G}$ has a perfect deterministic strategy. □

9.8. **The Graph Colouring Game.** Let $G = (V, E)$ be a graph with $|V| = n$. Recall that a *c-colouring* is a function $f : V \to \{1, ..., c\}$ such that whenever $v \sim w$, we have $f(v) \neq f(w)$. We saw that

$$\chi(G) = \min\{c \in \mathbb{N} : \text{ there exists a } c\text{-colouring of } G\}.$$

The graph colouring game, as a finite input-output game, is given by input sets $I_A = I_B = V$, output sets $\mathcal{O}_A = \mathcal{O}_B = \{1, ..., c\}$, and the

rule function λ defined as follows:

$$\text{for each vertex } v, \text{we have } \lambda(v,v,i,j) = \begin{cases} 1 \text{ if } i = j \\ 0 \text{ if } i \neq j \end{cases};$$

$$\text{for every } (v,w) \in E, \text{we have } \lambda(v,w,i,j) = \begin{cases} 1 \text{ if } i \neq j \\ 0 \text{ if } i = j \end{cases}.$$

Definition 9.29. For $t \in \{\text{loc}, q, qs, qa, qc, \text{vect}, nsb\}$, we set

$$\chi_t(G) = \min\{c \in \mathbb{N} : \exists p \in C_t(n,c) \text{ that is a perfect strategy}\}.$$

Now, in the case of the graph colouring game as described above, if $p \in C_t(n,c)$ is a perfect strategy, then using the rule function for this game, we must have $p(i,j|v,v) = 0$ for all $i \neq j$, while for each $(v,w) \in E$, we must have $p(i,i|v,w) = 0$. Note that since

$$C_{\text{loc}}(n,c) \subseteq C_q(n,c) \subseteq ... \subseteq C_{nsb}(n,k),$$

we must have, in the light of the above definition,

$$\chi_{\text{loc}}(G) \geq \chi_q(G) \geq ... \geq \chi_{nsb}(G).$$

Observe that if the Tsirelson Conjecture is true, then we must have $\chi_q(G) = \chi_{qc}(G)$ for every graph G. Similarly, if the Connes Embedding Conjecture is true, then $\chi_{qa}(G) = \chi_{qc}(G)$ for all graphs G. If the Closure Conjecture is true, then $\chi_q(G) = \chi_{qa}(G)$ for every graph. And finally, if the Werner-Scholze statement is true, then $\chi_{qs}(G) = \chi_{qa}(G)$ for all graphs G.

The goal, then, is to find methods to calculate these parameters for different graphs.

Definition 9.30 (Hadamard Graph)**.** Let $N \in \mathbb{N}$. The Hadamard graph, $\Omega_N = (V,E)$ is the graph where V is the set of all N-tuples with entries ± 1 (so Ω_N has 2^N vertices), with $(v,w) \in E(\Omega_N)$ if and only if $\langle v,w \rangle = 0$.

Remark 9.31. Notice that for any $N \in \mathbb{N}$, there exists a Hadamard graph Ω_N. Also, note that if N is odd then $v \cdot w \neq 0$ for every v, w, and hence Ω_N is the empty graph on 2^N vertices. So we focus on the case when N is even.

Proposition 9.32. *Let $G = (V,E)$ be a graph on $|V| = n$ vertices. Then $\chi(G)\alpha(G) \geq n$.*

Proof. Let $\chi(G) = c$. Then there exists a c-colouring $f : V \to \{1, ..., c\}$ for G. This allows us to partition the set V of vertices of G into c mutually disjoint sets as following: define $V_i = \{v \in V : f(v) = i\}; 1 \leq$

$i \leq c$ so that $V = \bigcup_{i=1}^{c} V_i$ and this union is a disjoint union. If $v, w \in V_i$ then v, w are not adjacent which implies that V_i is an independent set. Hence, $|V_i| \leq \alpha(G)$. It follows that

$$n = |V| = \sum_{i=1}^{c} |V_i| \leq c \cdot \alpha(G) = \chi(G)\alpha(G).$$

□

Theorem 9.33 (Frankl, 1986)**.** [**?**, Theorem 5.1] *If $N = 4p^{2\ell+1}$ where p is an odd prime, then $\left(\frac{27}{16}\right)^{\frac{N}{4}} \leq \chi(\Omega_N)$.*

Proof. From Proposition **??**, we obtain $\left(\frac{27}{16}\right)^{\frac{N}{4}} \leq \frac{2^N}{\alpha(\Omega_N)} \leq \chi(\Omega_N)$. □

Theorem 9.34. (Frankl-Rodl, 1987) *There is $\varepsilon_0 > 0$ such that for large enough N with N even, we have $\alpha(\Omega_N) \leq (2 - \varepsilon_0)^N$. Using the proposition above gives $\chi(\Omega_N) \geq \left(\frac{2}{2-\varepsilon_0}\right)^N$.*

Theorem 9.35 (DeKlerk-Pasechnik, 2005)**.** $\chi(\Omega_{16}) \geq 29$ *and* $\alpha(\Omega_{16}) = 2304$.

The above theorem is another example of $\chi(\Omega_N) \geq N$

It follows from results of Broussard-Cleve-Tapp, 1999, that $\chi_q(\Omega_N) \leq N$ whenever $N = 2^k$, $k \in \mathbb{N}$.

The first big result that got people interested in the quantities $\chi_t(G)$ is the following (stated here without proof).

Theorem 9.36 (Avis-Hagasawa-Kikuchi-Sasaki, 2006)**.** *If Ω_N is the Hadamard graph and N is even, then $\chi_q(\Omega_N) \leq N$.*

From the above theorem, it is clear that we need far fewer colours than the classical colouring number to fool a referee into believing that we have a colouring for our graph.

Proposition 9.37. *For every graph G, $\chi_{loc}(G) = \chi(G)$.*

Proof. We saw that there is $p \in C_{\text{loc}}(n, c)$ that is a perfect strategy for a game if and only if the game has a deterministic perfect strategy. In this case, there will be functions $f, g : V \to \{1, ..., c\}$ such that $\lambda(v, w, f(v), g(w)) = 1$ for all $v, w \in V$. In particular, we must have $\lambda(v, v, f(v), g(v)) = 1$. Using the first rule for the graph colouring game, we must have $f(v) = g(v)$, so that $f = g$. Now in the case when $v \sim w$, we have $\lambda(v, w, f(v), f(w)) = 1$ which forces $f(v) \neq f(w)$ by the second rule of the game. Therefore, f is a colouring of G by c colours, so $\chi_{\text{loc}}(G) \leq \chi(G)$. The other inequality is similar. □

Corollary 9.38 (Corollary to AHKS)**.** *For large N, $C_{loc}(2^N, N) \neq C_q(2^N, N)$.*

Like the situation where loc strategies don't fool the referee in any way (since the colouring number stays the same), if we allow any non-signalling box strategy, then the situation becomes much less interesting.

Proposition 9.39 (Paulsen-Todorov)**.** *Let G be a graph on $n \geq 2$ vertices. Then $\chi_{nsb}(G) = 2$.*

Proof. For $1 \leq i, j \leq 2$, we set

$$p(i,j|v,v) = \begin{cases} \frac{1}{2} \text{ if } i = j \\ 0 \text{ if } i \neq j \end{cases} .$$

For $v \neq w$, set

$$p(i,j|v,w) = \begin{cases} 0 \text{ if } i = j \\ \frac{1}{2} \text{ if } i \neq j. \end{cases}$$

We claim that $p \in C_{nsb}(n, 2)$ and p is a perfect nsb-strategy. Indeed, if $i \neq j$ then $p(i,j|v,v) = 0$ so the first rule is satisfied. For the second rule, whenever $v \sim w$ and $i = j$ we have $p(i,j|v,w) = 0$ so the second rule is satisfied. Therefore, p is a perfect strategy. It is easy to check that

$$\sum_{i,j} p(i,j|v,w) = 2 \cdot \frac{1}{2} = 1;$$

$$\sum_{j=1}^{2} p(i,j|v,w) = \frac{1}{2} \text{ for all } w, \text{ and}$$

$$\sum_{i=1}^{2} p(i,j|v,w) = \frac{1}{2} \text{ for all } v.$$

Hence, $p \in C_{nsb}(n, 2)$. This completes the proof. □

The next theorem implies the above result of [AHKS].

Theorem 9.40 (P-Todorov)**.** *Let $V \subseteq \mathbb{C}^N$ with $|V| = n$, where each $v \in V$ is such that $v = (v(0), ..., v(N-1)) \in \mathbb{C}^N$ with $|v(j)| = 1$ for all j. Define $E = \{(v, w) : v, w \in V, v \perp w\}$ and let $G = (V, E)$. Then $\chi_q(G) \leq N$ (and hence $\chi_q(\Omega_N) \leq N$).*

Proof. Let $\omega = e^{\frac{2\pi i}{N}}$. For each $v \in V$, let $D_v = \text{diag}(v(0), ..., v(N-1))$ which is unitary. For each $0 \leq k \leq N-1$, let $R_k = \frac{1}{N}(\omega^{(\ell-j)k})_{\ell,j=0}^{N-1} \in M_N$. Set $h_k = \frac{1}{\sqrt{N}} \sum_{\ell=0}^{N-1} \omega^{\ell k} e_\ell$, so that $R_k = h_k h_k^* \geq 0$. Recall that

$\sum_{k=0}^{N-1}(\omega^\ell)^k = 0$ for all $1 \le \ell \le N-1$ (it is equal to N when $\ell = 0, N$). Hence,

$$\sum_{k=0}^{N-1} R_k = \frac{1}{N}\left(\sum_{k=0}^{N-1}(\omega^{\ell-j})^k\right) = I.$$

Let $P_{v,k} = D_v^* R_k D_v$, so that $P_{v,k} \ge 0$ and $\sum_{k=0}^{N-1} P_{v,k} = I$. These are Alice's POVM's. Define $Q_{w,k} = D_w R_k^t D_w^*$. Since the transpose map is positive, $Q_{w,k} \ge 0$ and $\sum_k Q_{w,k} = I$. These $Q_{w,k}$'s are Bob's POVM's. Now let $\xi = \frac{1}{\sqrt{N}} \sum_{p=0}^{N-1} e_p \otimes e_p \in \mathbb{C}^N \otimes \mathbb{C}^N$. Then we define

$$p(k, m|v, w) = \langle \xi, P_{v,k} \otimes Q_{w,m}\xi \rangle \in C_q(n, N).$$

We need to show that

(1) $p(k, m|v, v) = 0$ for $k \ne m$, and

(2) $v \perp w \implies p(k, k|v, w) = 0$.

To see (1), notice that

$$\begin{aligned}
\langle \xi, P_{v,k} \otimes Q_{v,m}\xi \rangle &= \frac{1}{N}\sum_{p,q=1}^{N-1} \langle e_q \otimes e_q, (P_{v,k} \otimes Q_{v,m}) e_p \otimes e_p \rangle \\
&= \frac{1}{N}\sum_{p,q} \langle e_q, D_v^* R_k D_v e_p \rangle \langle e_q, D_v R_m^t D_v^* e_p \rangle \\
&= \frac{1}{N^3}\sum_{p,q} \overline{v(q)} v(p) \omega^{(p-q)k} \overline{v(p)} v(q) \omega^{(q-p)m} \\
&= \frac{1}{N^3}\sum_{p,q} \omega^{(p-q)(k-m)} \\
&= \frac{1}{N^3}\left(\sum_p (\omega^{k-m})^p\right)\left(\sum_q (\omega^{k-m})^q\right) \\
&= 0 \text{ if } m \ne \ell.
\end{aligned}$$

Now suppose that $v \perp w$. Then we consider

$$\begin{aligned}\langle \xi, (P_{v,k} \otimes Q_{w,k})\xi \rangle &= \frac{1}{N} \sum_{p,q=0}^{N-1} \langle e_q, P_{v,k} e_p \rangle \langle e_q, Q_{w,k} e_p \rangle \\ &= \frac{1}{N^3} \sum_{p,q} \overline{v(q)} v(p) \omega^{(p-q)k} \overline{w(p)w(q)} \omega^{(q-p)k} \\ &= \frac{1}{N^3} \sum_{p,q} (\overline{v(q)} w(q))(v(p)\overline{w(p)}) \\ &= \frac{1}{N^3} \langle v, w \rangle \cdot \langle w, v \rangle = 0.\end{aligned}$$

Hence, (2) is proved and it follows that p is a perfect q-strategy. This yields $\chi_q(G) \leq N$. □

Definition 9.41 (Synchronous)**.** For $t \in \{\text{loc}, q, qs, qa, qc, \text{vect}, nsb\}$, a correlattion (or, a conditional probability) $p(i,j|a,b) \in C_t(n,k)$ is said to be *synchronous* if $p(i,j|a,a) = 0$ for all $i \neq j$ and for all $a \in I$, where I is the common input set.

Proposition 9.42. *Suppose $p(i,j|a,b) \in C_{vect}(n,k)$ and that $\{v_{a,i}\}_{a,i}$ and $\{w_{b,j}\}_{b,j}$ are vectors from some Hilbert space $\mathcal{H}$ satisfying the definition of $p \in C_{vect}(n,k)$. If p is synchronous, then $v_{b,j} = w_{b,j}$.*

Proof. Note that

$$\begin{aligned}1 = \sum_{i,j=1}^{k} p(i,j|a,a) &= \sum_{i=1}^{k} p(i,i|a,a) \\ &= \sum_{i=1}^{k} \langle v_{a,i} | w_{a,i} \rangle = \left\langle \begin{pmatrix} v_{a,1} \\ \vdots \\ v_{a,k} \end{pmatrix}, \begin{pmatrix} w_{a,1} \\ \vdots \\ w_{a,k} \end{pmatrix} \right\rangle.\end{aligned}$$

Since there is some $\psi \in \mathcal{H}$ with $\sum_{i=1}^{k} v_{a,i} = \psi = \sum_{i=1}^{k} w_{a,i}$, $v_{a,i} \perp v_{a,j}$ and $w_{a,i} \perp w_{a,j}$ for $i \neq j$, we have

$$1 = \|\psi\|^2 = \sum_{i=1}^{k} \|v_{a,i}\|^2 = \sum_{i=1}^{k} \|w_{a,i}\|^2.$$

Using Cauchy-Schwarz on the inner product before forces $h_{a,i} = k_{a,i}$ for each i. □

Definition 9.43. For a graph G on n vertices, we define $\theta'(G)$ as:

$$\theta'(G) := \sup\{\lambda_{\max}(I_n + K) : I_n + K \geq 0,\ k_{ii} = 0,\ k_{ij} = 0\, \forall (i,j) \in E(G),\ k_{ij} \geq 0\}.$$

Remark 9.44. It is easy to see that $\theta'(G) \leq \theta(G)$.

Proposition 9.45. *Let $P, X \in \mathcal{B}(\mathcal{H})$.*

(1) *If $\begin{pmatrix} P & X \\ X^* & P \end{pmatrix} \geq 0$, then $\|X\| \leq \|P\|$.*

(2) *We have $\begin{pmatrix} I & X \\ X^* & I \end{pmatrix} \geq 0$ if and only if $\|X\| \leq 1$.*

Proof. Let $h \in \mathcal{H}, \|h\| = 1$ and $t \in \mathbb{R}$. Then

$$\begin{aligned} 0 &\leq \left\langle \begin{pmatrix} -Xh \\ th \end{pmatrix}, \begin{pmatrix} P & X \\ X^* & P \end{pmatrix} \begin{pmatrix} -Xh \\ th \end{pmatrix} \right\rangle \\ &= \left\langle \begin{pmatrix} -Xh \\ th \end{pmatrix}, \begin{pmatrix} -PXh + tXh \\ -X^*Xh + tPh \end{pmatrix} \right\rangle \\ &= \langle Xh, PXh\rangle - \langle Xh, tXh\rangle - t\langle h, X^*Xh\rangle + t^2\langle h, Ph\rangle \\ &\leq \|P\|\|Xh\|^2 - 2t\|Xh\|^2 + t^2\|P\|. \end{aligned}$$

Since the above quadratic expression is positive for every $h \in \mathcal{H}$, it follows then that its discriminant must be non-positive which implies that $4\|Xh\|^4 - 4\|P^2\|\|Xh\|^2 \leq 0$. Therefore, $\|Xh\|^2 \leq \|P\|^2$. Taking the supremum over all such h, we have $\|X\|^2 \leq \|P\|^2$. This proves (1).

For (2), the forward implication follows from (1). For the backward implication, suppose that $\|X\| \leq 1$. Let $h, k \in \mathcal{H}$. We see that

$$\begin{aligned} \left\langle \begin{pmatrix} h \\ k \end{pmatrix}, \begin{pmatrix} I & X \\ X^* & I \end{pmatrix} \begin{pmatrix} h \\ k \end{pmatrix} \right\rangle &= \langle h, h + Xk\rangle + \langle k, X^*h + k\rangle \\ &= \|h\|^2 + \langle h, Xk\rangle + \langle k, X^*h\rangle + \|k\|^2 \\ &\geq \|h\|^2 - 2\|h\|\|X\|\|k\| + \|k\|^2 \\ &\geq \|h\|^2 - 2\|h\|\|k\| + \|k\|^2 \\ &= (\|h\| - \|k\|)^2 \geq 0. \end{aligned}$$

This completes the proof. □

Proposition 9.46. *Let $\Phi : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{K})$ be a completely positive map. Then $\|\Phi\| = \|\Phi(I)\|$.*

Proof. Let $X \in \mathcal{B}(\mathcal{H})$ with $\|X\| \leq 1$. Then $\begin{pmatrix} I & X \\ X^* & I \end{pmatrix} \geq 0$, so that $\begin{pmatrix} \Phi(I) & \Phi(X) \\ \Phi(X)^* & \Phi(I) \end{pmatrix} \geq 0$ so that $\|\Phi(X)\| \leq \|\Phi(I)\|$. (Note that any positive map preserves adjoint.) Since this is true for any X with norm at most 1 we infer that $\|\Phi\| \leq \|\Phi(I)\|$. Since $\|I\| = 1$, we have $\|\Phi(I)\| \leq \|\Phi\|$ which implies that $\|\Phi\| = \|\Phi(I)\|$. □

Definition 9.47. Given a set of vectors $h_r \in \mathcal{H}$, the matrix $Gr = (\langle h_r, h_s \rangle)$ is called the *Grammian*.

Proposition 9.48. $(\langle h_i, h_j \rangle)_{i,j=1}^N \geq 0$.

Proof. For any $\lambda_1, ..., \lambda_N \in \mathbb{C}$, we compute

$$\begin{aligned}\left\langle \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_N \end{pmatrix}, (\langle h_i, h_j \rangle) \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_N \end{pmatrix} \right\rangle &= \sum_{i,j=1}^N \overline{\lambda}_i \langle h_i, h_j \rangle \lambda_j \\ &= \left\langle \sum_{i=1}^N \lambda_i h_i, \sum_{j=1}^N \lambda_j h_j \right\rangle \\ &= \left\| \sum_{i=1}^N \lambda_i h_i \right\|^2 \geq 0.\end{aligned}$$

□

Proposition **??** and Proposition **??** can be used to prove the following theorem

Theorem 9.49 (Paulsen-Todorov)**.** *Let G be a graph on N vertices. Then*

$$\chi_{vect}(G) \geq \frac{n}{\theta'(G)}.$$

Proof. Let $\chi_{\text{vect}}(G) = c$. This means we have vectors $\{h_{v,i}\}_{i=1}^c$ for $1 \leq v \leq n$ such that $p(i,j|v,w) = \langle h_{v,i}, h_{w,j} \rangle$ gives a perfect strategy. Moreover, $h_{v,i} \perp h_{v,j}$ for $i \neq j$, while there is some $\varphi \in \mathcal{H}$ with $\sum_{i=1}^c h_{v,i} = \varphi$ for all v. If $(v,w) \in E(G)$, then by the rules of the colouring game, $p(i,i|v,w) = \langle h_{v,i}, h_{w,i} \rangle = 0$ so that $h_{v,i} \perp h_{w,i}$.

Now, let $Q_{ij} = (\langle h_{v,i}, h_{w,j} \rangle)_{v,w=1}^n \in M_n$. Then let $Q = (Q_{ij}) \in M_{nc}$. We can see that Q is a Grammian of the set $\{h_{v,i}\}$ (for some ordering of the set). Hence, $Q \geq 0$. By Choi's theorem, the map $\Phi : M_c \to M_n$ given by $\Phi(E_{ij}) = Q_{ij}$ is completely positive. Evaluating at the identity gives

$$\Phi(I_c) = \sum_{i=1}^c \Phi(E_{ii}) = \sum_{i=1}^c (\langle h_{v,i}, h_{w,i} \rangle).$$

The (v,v)-entry of $\Phi(I_c)$ is $\sum_{i=1}^c \|h_{v,i}\|^2 = \|\varphi\|^2 = 1$ since $h_{v,i} \perp h_{v,j}$ for $i \neq j$ and $\sum_{i=1}^c h_{v,i} = \varphi$. If $(v,w) \in E(G)$, then $h_{v,i} \perp h_{w,i}$ so that the (v,w) entry of $\Phi(I_c)$ is 0. If $v \neq w$ and $(v,w) \notin E(G)$, the (v,w) entry is a sum of non-negative inner products and is non-negative. It follows that $\Phi(I_c) = I_n + K \geq 0$, where $k_{vw} = 0$ if

$(v,w) \in E(G)$, $k_{vv} = 0$ and $k_{vw} \geq 0$. Hence, $\lambda_{\max}(I_n + K) \leq \theta'(G)$. But $\lambda_{\max}(I_n+K) = \|I_n+K\| = \|\Phi(I_c)\|$, so we obtain $\|\Phi(I_c)\| \leq \theta'(G)$.

Let J_k be the $k \times k$ matrix of all 1's, so that $J_k = J_k^*$ and $J_k^2 = kJ_k$. Hence, $\lambda \in \sigma(J_k)$ implies $\lambda^2 = k\lambda$ so that $\lambda \in \{0, k\}$. Since $e_1 + \cdots + e_k$ is an eigenvector for J_k with eigenvalue k, we have $\|J_k\| = \lambda_{\max}(J_k) = k$. Now,

$$\Phi(J_c) = \sum_{i,j=1}^{c} \Phi(E_{ij}) = \sum_{i,j=1}^{c} Q_{ij}.$$

The (v,w)-entry of $\Phi(J_c)$ is

$$\sum_{i,j=1}^{c} \langle h_{v,i}, h_{w,j} \rangle = \left\langle \sum_{i=1}^{c} h_{v,i}, \sum_{j=1}^{c} h_{w,j} \right\rangle = \langle \varphi, \varphi \rangle = 1.$$

Hence, $\Phi(J_c) = J_n$, so that $n = \|J_n\| = \|\Phi(J_c)\| \leq \|\Phi\|\|J_c\| = c\theta'(G)$. It follows that $c \geq \frac{n}{\theta'(G)}$. Since $c = \chi_{\text{vect}}(G)$, we have

$$\chi_{\text{vect}}(G) \geq \frac{n}{\theta'(G)}.$$

□

As an immediate corollary, we have

Corollary 9.50.

$$\frac{n}{\theta'(G)} \leq \chi_t(G) \text{ for all } t \in \{loc, q, qa, qs, qc, vect\}.$$

Definition 9.51 (Induced Subgraph)**.** Given a graph $G = (V, E)$ and a subset $V' \subseteq V$, we define $G_{V'} = (V', E')$, where

$$E' = \{(v,w) \in V' \times V' : (v,w) \in E(G)\}.$$

Example 9.52. Induced subgraph example (insert)

Proposition 9.53. *Let $V' \subseteq V$. Then for $t \in \{loc, q, qs, qa, qc, vect\}$, we have $\chi_t(G_{V'}) \leq \chi_t(G_V)$.*

Proof. The proof is trivial. (Given a solution for $c = \chi_t(G_V)$, just restrict it to V').

□

This leads us to the following two corollaries:

Corollary 9.54 (CMNSW)**.** $\omega(G) \leq \chi_{vect}(G)$.

Proof. Suppose that $\omega(G) = m$, and restrict to those m vertices V' so that $G_{V'} = K_m$, the complete graph on m vertices. We know that $\chi_{\text{vect}}(K_m) \leq \chi_{\text{vect}}(G)$. We claim that $\chi_{\text{vect}}(K_m) = m$ (which gives the result). We know that $\chi_{\text{vect}}(K_m) \leq \chi(K_m) = m$. We now compute $\theta'(K_m)$. If $H \in M_m$ is such that $I + H \geq 0$, $h_{ij} = 0$ for all $(i,j) \in E(K_m)$ and $h_{ii} = 0$, $h_{ij} \geq 0$, then $H = 0$. Hence, $\theta'(K_m) = 1$. By the corollary, $m = \frac{m}{\theta'(K_m)} \leq \chi_{\text{vect}}(G)$ so that $\chi_{\text{vect}}(K_m) = m$. □

Corollary 9.55. *$\chi_t(K_m) = m$ for all $t \in \{loc, q, qa, qs, qc, vect\}$.*

We end this section by collecting some results from a paper due to Cubitt, Mancinska, Roberson, Severini, Stahlke and Winter. First, it is known that

$$\theta(G) = \min\{\lambda : \exists Z \geq 0,\ z_{ii} = \lambda - 1,\ z_{ij} = -1\ \forall i \not\sim j\}.$$

One may define the quantity

$$\theta^+(G) = \min\{\lambda : \exists Z \geq 0,\ z_{ii} = \lambda - 1,\ z_{ij} = -1 \text{ for } i \not\sim j, \text{ and } z_{ij} \geq -1\}.$$

It is known that $\theta^+(G) \geq \theta(G)$. Now, recall that Lovasz's theorem states that $\theta(\overline{G}) \leq \chi(G)$.

Theorem 9.56 (CMRSSW)**.** (1) *$\chi_{vect}(G) = \lceil \theta^+(\overline{G}) \rceil$.*
(2) *$\chi_{vect}(C_5 * K_3) = 7$ while $\chi_q(C_5 * K_3) = 8$.*

In particular, the previous result implies that $C_q(15,7) \subsetneq C_{\text{vect}}(15,7)$.

Remark 9.57. Tsirelson actually thought that $C_q(n,k) = C_{\text{vect}}(n,k)$ for all $n, k \in \mathbb{N}$. Indeed, his idea for proving that $C_q(n,k) = C_{qc}(n,k)$ was to prove that $C_q(n,k) = C_{\text{vect}}(n,k)$ (which is false).

Remark 9.58. The above work shows that $\theta(\overline{G}) \leq \chi_{\text{vect}}(G)$. So while Lovasz thought that his function gave a good lower bound on the chromatic number, it actually gives a lower bound for $\chi_{\text{vect}}(G)$, which could be much smaller.

Remark 9.59. For G, H graphs, we define $G * H$ as follows: we define $V(G * H) = V(G) \times V(H)$, and for $(v_1, w_1), (v_2, w_2) \in V(G) \times V(H)$, we say that $(v_1, w_1) \sim (v_2, w_2)$ if and only if $v_1 \sim_G v_2$ or $w_1 \sim_H w_2$.

10. Synchronous Correlations and Traces

Recall that $C_t^s(n,k)$ is defined as the set of all $p \in C_t(n,k)$ such that $p(i,j|a,a) = 0$ for $i \neq j$. We saw that if we had $p \in C_{\text{vect}}^s(n,k)$ given by $p(i,j|a,b) = \langle h_{a,i} | k_{b,j} \rangle$, then $h_{a,i} = k_{a,i}$ for all a, i.

Definition 10.1 (Concrete C^*-algebra)**.** A (concrete) *C^*-algebra* is a set $\mathcal{A} \subseteq \mathcal{B}(\mathcal{H})$ for some Hilbert space $\mathcal{H}$ satisfying the following:

- If $X, Y \in \mathcal{A}$ and $\lambda \in \mathbb{C}$ then $X+Y \in \mathcal{A}$, $\lambda X \in \mathcal{A}$ and $XY \in \mathcal{A}$.
- If $X \in \mathcal{A}$ then $X^* \in \mathcal{A}$.
- If $(X_n)_{n=1}^{\infty} \subseteq \mathcal{A}$ and $X \in \mathcal{B}(\mathcal{H})$ are such that $\lim_{n\to\infty} \|X_n - X\| = 0$, then $X \in \mathcal{A}$.

We call a C^*-algebra *unital* if $I \in \mathcal{A}$. We set $\mathcal{A}^+ = \{X \in \mathcal{A} : X \geq 0\}$.

Definition 10.2 (Positive Linear Functional)**.** Given a concrete unital C^*-algebra, a map $f : \mathcal{A} \to \mathbb{C}$ is called a *positive linear functional* if it is a linear functional and if $f(X) \geq 0$ whenever $X \in \mathcal{A}^+$. A positive linear functional f is a *state* if $f(I) = 1$.

Example 10.3. Let $\mathcal{A} = M_n$ and $\varphi \in \mathbb{C}^n$. Then $f(X) = \langle \varphi | X\varphi \rangle$ is a positive linear functional. Notice that f is a state whenever $\|\varphi\| = 1$ (i.e. when φ is a state vector).

Example 10.4. If $P \geq 0$ in M_n, then the map $f : M_n \to \mathbb{C}$ given by $f(X) = \mathrm{Tr}(PX)$ is a positive linear functional, which is a state when $\mathrm{Tr}(P) = 1$. (This example is abstracting the idea of mixed states)

Definition 10.5 (Tracial State)**.** Let $\mathcal{A}$ be a unital C^*-algebra. We say that $f : \mathcal{A} \to \mathbb{C}$ is a *tracial state* if it is a state such that $f(XY) = f(YX)$ for all $X, Y \in \mathcal{A}$.

Proposition 10.6. *Let $f : M_n \to \mathbb{C}$ be a linear map such that $f(XY) = f(YX)$ for all $X, Y \in M_n$. Let $a = f(E_{11})$. Then $f(X) = a\,Tr(X)$.*

Proof. Since $E_{ii}E_{ij} = E_{ij}$, we always have

$$f(E_{ij}) = f(E_{ii}E_{ij}) = f(E_{ij}E_{ii}),$$

and this is 0 whenever $i \neq j$. Similarly, since $E_{11} = E_{1i}E_{i1}$, we have

$$a = f(E_{11}) = f(E_{1i}E_{i1}) = f(E_{i1}E_{1i}) = f(E_{ii}).$$

Hence whenever $X \in M_n$, say $X = \sum_{i,j} x_{ij}E_{ij}$, we have $f(X) = \sum_i a x_{ii} = a\mathrm{Tr}(X)$. □

Proposition 10.7. *There exists a unique tracial state on M_n given by*

$$Tr_n(X) = \frac{1}{n}\,Tr(X).$$

Proof. If $X \geq 0$, then $\mathrm{Tr}(X) \geq 0$, while $\frac{1}{n}\mathrm{Tr}(I) = 1$, so Tr_n is a tracial state. Conversely, if $f : M_n \to \mathbb{C}$ is a tracial state, then by the proposition above, $f(X) = a\mathrm{Tr}(X)$ for all $X \in M_n$. Since $f(I) = 1$, we have $a = \frac{1}{n}$. □

Example 10.8. Let $\mathcal{A} = \left\{ \begin{pmatrix} X_1 & 0 \\ 0 & X_2 \end{pmatrix} : X_1 \in M_{n_1},\, X_2 \in M_{n_2} \right\} \subseteq M_{n_1+n_2}$, which is a unital C^*-algebra. Suppose that $f : \mathcal{A} \to \mathbb{C}$ is a tracial state.

Then $g_1 : M_{n_1} \to \mathbb{C}$ with $g_1(X_1) = f\left(\begin{pmatrix} X_1 & 0 \\ 0 & 0 \end{pmatrix}\right)$ satisfies $g_1(X_1Y_1) = g(Y_1X_1)$ for all $X_1, Y_1 \in M_{n_1}$. Hence, $g_1(X_1) = a\mathrm{Tr}(X_1)$ for some $a \in \mathbb{C}$. Similarly, $g_2 : M_{n_2} \to \mathbb{C}$ given by $g_2(X_2) = f\left(\begin{pmatrix} 0 & 0 \\ 0 & X_2 \end{pmatrix}\right)$ must satisfy $g(X_2) = b\mathrm{Tr}(X_2)$ for some $b \in \mathbb{C}$. But $1 = f\left(\begin{pmatrix} I_{n_1} & 0 \\ 0 & I_{n_2} \end{pmatrix}\right) = a\mathrm{Tr}(I_{n_1}) + b\mathrm{Tr}(I_{n_2}) = n_1a + n_2b$. Let $p_1 = n_1a$ and $p_2 = n_2b$. Then $0 \leq g_1(I_{n_1}) = p_1$, and similarly, $p_2 \geq 0$. We also must have $p_1 + p_2 = 1$. Hence for any element of $\mathcal{A}$ we have

$$f\left(\begin{pmatrix} X_1 & 0 \\ 0 & X_2 \end{pmatrix}\right) = a\mathrm{Tr}(X_1) + b\mathrm{Tr}(X_2) = p_1\mathrm{Tr}_{n_1}(X_1) + p_2\mathrm{Tr}_{n_2}(X_2).$$

Conversely, any such map gives a tracial state on $\mathcal{A}$, so we have a one-parameter family of tracial states on $\mathcal{A}$.

Problem 10.9. *Let*

$$\mathcal{A} = \left\{ \begin{pmatrix} X_1 & & \\ & \ddots & \\ & & X_k \end{pmatrix} : X_i \in M_{n_i} \right\}.$$

Prove that $f : \mathcal{A} \to \mathbb{C}$ is a tracial state if and only if there are $p_i \geq 0$ such that $p_1 + \cdots + p_k = 1$ and

$$f\left(\begin{pmatrix} X_1 & & \\ & \ddots & \\ & & X_k \end{pmatrix}\right) = p_1\,Tr_{n_1}(X_1) + \cdots + p_k\,Tr_{n_k}(X_k).$$

Proof. **Homework problem 19**; due 29th March, Tuesday. □

Definition 10.10 (Unital $*$-Representation)**.** Let $\mathcal{A}$ be a unital C^*-algebra and $\mathcal{H}$ be a Hilbert space. A *unital $*$-representation of $\mathcal{A}$ on $\mathcal{H}$* is a map $\pi : \mathcal{A} \to \mathcal{B}(\mathcal{H})$ such that

(1) $\pi(1) = I$ (unital)
(2) π is linear
(3) $\pi(XY) = \pi(X)\pi(Y)$.
(4) $\pi(X^*) = \pi(X)^*$.

With all of these properties together, π is a $*$-homomorphism.

What follows next is a celebrated theorem from the theory of C^*-algebras. This is essential for moving further.

Theorem 10.11 (Gelfand-Naimark-Segal Construction)**.** *Let $\mathcal{A}$ be a unital C^*-algebra and $s : \mathcal{A} \to \mathbb{C}$ be a state. Then there is a Hilbert*

space $\mathcal{H}_s$, a vector $\varphi \in \mathcal{H}_s$ of norm 1 and a unital $*$-representation $\pi_s : \mathcal{A} \to \mathcal{B}(\mathcal{H}_s)$ such that

(1) $s(X) = \langle \varphi | \pi_s(X) \varphi \rangle$.
(2) $\{\pi(X)\varphi : X \in \mathcal{A}\}$ *is dense in* $\mathcal{H}_s$.

This is called the GNS representation of s.

We are now ready to prove the following theorem.

Theorem 10.12 (PSSTW)**.** *Let $p(i,j|v,w) \in C^s_{qc}(n,k)$ be represented by $p(i,j|v,w) = \langle \varphi | E_{vi} F_{wj} \varphi \rangle$ where E_{vi}, F_{wj} are projections with $\sum_i E_{vi} = I$ and $\sum_j F_{wj} = I$ for all i,j, with $E_{vi}F_{wj} = F_{wj}E_{vi}$ for all i,j,v,w. Let $\mathcal{A} \subseteq \mathcal{B}(\mathcal{H})$ be the C^*-algebra generated by the set $\{E_{vi} : 1 \leq v \leq n,\ 1 \leq i \leq k\}$. Let $f : \mathcal{A} \to \mathbb{C}$ be defined by $f(X) = \langle \varphi | X \varphi \rangle$. Then f is a tracial state on $\mathcal{A}$. Conversely, if $\mathcal{A} \subseteq \mathcal{B}(\mathcal{H})$ is a C^*-algebra generated by projections $\{\tilde{E}_{vi} : 1 \leq v \leq n,\ 1 \leq i \leq k\}$ where $\sum_{i=1}^k \tilde{E}_{vi} = I$ for all v, and if $\tau : \mathcal{A} \to \mathbb{C}$ is a tracial state, then $p(i,j|v,w) := \tau(\tilde{E}_{vi}\tilde{E}_{wj}) \in C^s_{qc}(n,k)$.*

Proof. Since $C^s_{qc}(n,k) \subseteq C^s_{\text{vect}}(n,k)$, letting $h_{vi} = E_{vi}\varphi$ and $k_{wj} = F_{wj}\varphi$, we have

$$p(i,j|v,w) = \langle \varphi, E_{vi}F_{wj}\varphi \rangle = \langle E_{vi}\varphi, F_{wj}\varphi \rangle = \langle h_{vi}, k_{wj} \rangle.$$

Moreover, $\sum_i h_{vi} = (\sum_i E_{vi})\,\varphi = \varphi$. Since, by assumption, we have $E_{vi}E_{vj} = 0$ for $i \neq j$, we have $h_{vi} \perp h_{vj}$ for $i \neq j$. Similarly, $k_{wi} \perp k_{wj}$ for all $i \neq j$. Using Cauchy-Schwarz, we can deduce that, since $1 = \sum_i \langle h_{vi}, k_{vi} \rangle$, we must have $h_{vi} = k_{vi}$ for all v,i. In particular, $E_{vi}\varphi = F_{vi}\varphi$, so that $p(i,j|v,w) = \langle \varphi, E_{vi}F_{wj}\varphi \rangle = \langle \varphi, E_{vi}E_{wj}\varphi \rangle$.
Now we prove that $f : \mathcal{A} \to \mathbb{C}$ with $f(X) = \langle \varphi, X\varphi \rangle$ is a tracial state. First, note that it is a state since $\|\varphi\| = 1$. To see that it is tracial, let $W = E_{v_1,i_1} \cdots E_{v_m,i_m}$. Then

$$\begin{aligned} f(WE_{vi}) &= \langle \varphi, WE_{vi}\varphi \rangle \\ &= \langle \varphi, WF_{vi}\varphi \rangle \\ &= \langle \varphi, F_{vi}W\varphi \rangle \\ &= \langle F_{vi}\varphi, W\varphi \rangle \\ &= \langle E_{vi}\varphi, W\varphi \rangle \\ &= \langle \varphi, E_{vi}W\varphi \rangle = f(E_{vi}W). \end{aligned}$$

Now if $X = \sum \lambda_\ell W_\ell$ where the W_ℓ's are words with letters coming from Alice's projections, then

$$f(XE_{vi}) = \sum_\ell \lambda_\ell f(W_\ell E_{vi}) = \sum_\ell \lambda_\ell f(E_{vi}W_\ell) = f(E_{vi}X).$$

But such X form a dense set in $\mathcal{A}$, so that $f(XE_{vi}) = f(E_{vi}X)$ for all $X \in \mathcal{A}$. Similarly, we see that

$$f(E_{wj}(E_{vi}X)) = f((E_{vi}X)E_{wj}) = f(XE_{wj}E_{vi}),$$

for all $X \in \mathcal{A}$. One can see (using induction) that whenever W is a word with letters from Alice's projections and $X \in \mathcal{A}$, then $f(XW) = f(WX)$. Taking linear combinations and using completeness of $\mathcal{A}$, we have $f(XY) = f(YX)$ for all $X, Y \in \mathcal{A}$. So f is a tracial state.

Conversely, suppose that $\mathcal{A} \subseteq \mathcal{B}(\mathcal{H})$ is a C^*-algebra generated by projections $\{E_{vi} : 1 \leq v \leq n,\ 1 \leq i \leq k\}$ where $\sum_i E_{vi} = 1$ for all v, and suppose that $\tau : \mathcal{A} \to \mathbb{C}$ is a tracial state. We still need to construct the operators for Bob, and these may not be on the original Hilbert space. We use the GNS representation of τ, which gives a Hilbert space $\mathcal{H}_\tau$, a unit vector $\varphi \in \mathcal{H}_\tau$ and a $*$-representation $\pi_\tau : \mathcal{A} \to \mathcal{B}(\mathcal{H}_\tau)$. Now set $\pi_\tau(E_{vi}) = \tilde{E}_{vi}$. Then

$$\sum_i \tilde{E}_{vi} = \sum_i \pi_\tau(E_{vi}) = \pi_\tau(1) = I.$$

Using the fact that π_τ is a $*$-homomorphism, we obtain $\tilde{E}_{vi}^2 = \pi_\tau(E_{vi}^2) = \pi_\tau(E_{vi}) = \tilde{E}_{vi}$. Similarly, $\tilde{E}_{vi}^* = \tilde{E}_{vi}$ so $\{\tilde{E}_{vi}\}$ are projections on $\mathcal{H}_\tau$.

From these projections above, we define $F_{w,j}$ on the dense subset $\{\pi_\tau(X)\varphi : X \in \mathcal{A}\} \subseteq \mathcal{H}_\tau$ by $F_{w,j}(\pi_\tau(X)\varphi) = \pi_\tau(XE_{w,j})\varphi$. To see that this is well-defined, suppose that $\pi_\tau(X)\varphi = 0$; we need to show that $\pi_\tau(XE_{w,j})\varphi = 0$. Note that

$$0 = \langle \pi_\tau(X)\varphi | \pi_\tau(X)\varphi \rangle = \langle \varphi | \pi_\tau(X^*X)\varphi \rangle = \tau(X^*X).$$

We also have

$$\begin{aligned} 0 &\leq \|\pi_\tau(XE_{wj})\varphi\|^2 \\ &= \langle \pi_\tau(XE_{w,j})\varphi | \pi_\tau(XE_{w,j})\varphi \rangle \\ &= \langle \varphi | \pi_\tau(E_{wj}^* X^* X E_{wj})\varphi \rangle \\ &= \tau((E_{w,j}^* X^*)(XE_{w,j})) \\ &= \tau(XE_{w,j}E_{w,j}^* X^*) = \tau(XE_{w,j}X^*). \end{aligned}$$

Since $E_{w,j}$ is a projection, $E_{w,j} \leq I$ so that $XE_{w,j}X^* \leq XX^*$. Hence, $\|\pi_\tau(XE_{wj})\varphi\|^2 \leq \tau(XX^*) = \tau(X^*X) = 0$. Therefore, $\pi_\tau(XE_{w,j})\varphi = 0$ so that $F_{w,j}$ is well-defined.

The same calculation as above shows that $\|F_{w,j}\pi_\tau(X)\varphi\| \leq \|\pi_\tau(X)\varphi\|$, so each $F_{w,j}$ is a contraction on the dense subset $\{\pi_\tau(X)\varphi : X \in \mathcal{A}\}$. Hence $F_{w,j}$ extends to all of $\mathcal{H}_\tau$ and is still a contraction. We also have

$$F_{w,j}^2\pi_\tau(X)\varphi = F_{w,j}(\pi_\tau(XE_{w,j})\varphi) = \pi_\tau(XE_{w,j}^2)\varphi = F_{w,j}\pi_\tau(X)\varphi.$$

Hence by continuity, $F_{w,j}^2 = F_{w,j}$. Now since $F_{w,j}$ is an idempotent and a contraction, it must be an orthogonal projection. To see that $\sum_j F_{w,j} = I$, we again compute on the dense subset $\{\pi_\tau(X)\varphi : X \in \mathcal{A}\}$:

$$\left(\sum_i F_{w,i}\right)\pi_\tau(X)\varphi = \sum_i \pi_\tau(XE_{w,i})\varphi = \pi_\tau(X)\varphi.$$

Hence, $\sum_i F_{w,i} = I$ for all w.

We claim that $F_{w,j}\tilde{E}_{v,i} = \tilde{E}_{v,i}F_{w,j}$. With $\tilde{E}_{vi} = \pi_\tau(E)_{vi})$, we have

$$\begin{aligned}(F_{w,j}\tilde{E}_{v,i})(\pi_\tau(X)\varphi) &= F_{w,j}\pi_\tau(E_{v,i}X)\varphi \\ &= \pi_\tau(E_{v,i}XE_{w,j})\varphi.\end{aligned}$$

Similarly,

$$\begin{aligned}(\tilde{E}_{v,i}F_{w,j})\pi_\tau(X)\varphi &= \tilde{E}_{v,i}\pi_\tau(XE_{w,j})\varphi \\ &= \pi_\tau(E_{v,i}XE_{w,j})\varphi.\end{aligned}$$

It follows that $F_{w,j}\tilde{E}_{v,i} = \tilde{E}_{v,i}F_{w,j}$. Therefore, if we define $p(i,j|v,w) = \langle\varphi|\tilde{E}_{v,i}F_{w,j}\varphi\rangle$, then $p \in C_{qc}(n,k)$. Moreover,

$$F_{w,j}\varphi = F_{w,j}\pi_\tau(I)\varphi = \pi_\tau(E_{w,j})\varphi = \tilde{E}_{w,j}\varphi.$$

Hence,

$$p(i,j|v,w) = \langle\varphi|\tilde{E}_{v,i}\tilde{E}_{w,j}\varphi\rangle = \langle\varphi|\pi_\tau(E_{v,i}E_{w,j})\varphi\rangle = \tau(E_{v,i}E_{w,j}).$$

To see that this probability density is synchronous, note that if $i \neq j$, then $E_{v,i}E_{v,j} = 0$ so that $p(i,j|v,v) = \tau(E_{v,i}E_{v,j}) = 0$. Thus, $p \in C_{qc}^s(n,k)$. □

Theorem 10.13. *Let $\mathcal{G} = (I, \mathcal{O}, \lambda)$ be a synchronous game. Then $\mathcal{G}$ has a perfect qc-strategy if and only if there is a unital C^*-algebra generated by projections $\{E_{vi} : v \in I,\ i \in \mathcal{O}\}$, and a tracial state τ on this C^*-algebra such that*

1. *$\sum_i E_{vi} = I$ for all v.*
2. *If $\lambda(v,w,i,j) = 0$ then $E_{v,i}E_{w,j} = 0$.*

Proof. For this proof we assume that $|I| = n$ and $|\mathcal{O}| = k$. Suppose we have a unital C^*-algebra and tracial state with the properties above. Then if τ is our tracial state, by the previous theorem we may set $p(i,j|v,w) = \tau(E_{v,i}E_{w,j}) \in C_{qc}^s(n,k)$. If $\lambda(v,w,i,j) = 0$ then $p(i,j|v,w) = \tau(E_{v,i}E_{w,j}) = \tau(0) = 0$ by assumption. Hence, $\mathcal{G}$ has a perfect *qc*-strategy.

Conversely, suppose that $p(i,j|v,w) \in C_{qc}^s(n,k)$ is a perfect *qc*-strategy for $\mathcal{G}$. We saw that if we took the representation $p(i,j|v,w) =$

$\langle\psi|E_{v,i}F_{w,j}\psi\rangle$ of $p(i,j|v,w)$ and let $\mathcal{A}$ be the C^*-algebra generated by $\{E_{v,i}\}_{v,i}$, then $p(i,j|v,w) = \tau(E_{v,i}E_{w,j})$ gives rise to a trace on $\mathcal{A}$. If $\lambda(v,w,i,j) = 0$ then $\tau(E_{v,i}E_{w,j}) = 0$, and we want to show that $E_{v,i}E_{w,j} = 0$. Take any $X \in \mathcal{A}$; then $X^*X \leq \|X\|^2 I$. Hence, $E^*_{w,j}E^*_{v,i}(X^*X)E_{v,i}E_{w,j} \leq \|X\|^2 E^*_{w,j}E^*_{v,i}E_{v,i}E_{w,j}$. Taking traces we obtain

$$0 \leq \tau(E^*_{w,j}E^*_{v,i}X^*XE_{v,i}E_{w,j}) \leq \|X\|^2\tau(E^*_{w,j}E_{v,i}E_{w,j}) = \|X\|^2\tau(E_{v,i}E_{w,j}) = 0.$$

Therefore, $\tau(E^*_{w,j}E^*_{v,i}X^*XE_{v,i}E_{w,j}) = 0$. Switching some variables around shows that

$$\tau(XE_{v,i}E_{w,j}E^*_{w,j}E^*_{v,i}X^*) = 0.$$

For convenience, let $A = XE_{v,i}E_{w,j}$, so that $AA^* = XE_{v,i}E_{w,j}E^*_{w,j}E^*_{v,i}X^*$. Then by the GNS construction, on some Hilbert space (and for some unit vector φ), we have

$$\begin{aligned} 0 = \tau(AA^*) &= \langle\varphi|\pi_\tau(AA^*)\varphi\rangle \\ &= \|\pi_\tau(A^*)\varphi\| \\ &= \|\pi_\tau(E^*_{w,j}E^*_{v,i}X^*)\varphi\| \\ &= \|\tilde{E}_{w,j}\tilde{E}_{v,i}\pi_\tau(X^*)\varphi\| \end{aligned}$$

But using GNS, $\{\pi_\tau(X^*) : X \in \mathcal{A}\}$ is dense in the new Hilbert space, so that $\tilde{E}_{w,j}\tilde{E}_{v,i} = 0$. Hence, $\tilde{E}_{v,i}\tilde{E}_{w,j} = 0$ as desired. □

Note that in the above proof, $\tilde{E}_{v,i}$'s are not Alice's original projections.

Corollary 10.14. *Let G be a graph. Then $\chi_{qc}(G) \leq k$ if and only if there is a unital C^*-algebra $\mathcal{A}$ with trace and projections $\{E_{v,i}\}_{1\leq i\leq k,\, 1\leq v\leq n}$ in $\mathcal{A}$ with $\sum_{i=1}^k E_{v,i} = I$ for all v such that $v \sim w$ implies that $E_{v,i}E_{w,i} = 0$ for all $1 \leq i \leq k$.*

It is worthwhile noticing that $E_{v,i}$'s, in above corollary, can't be any projections in any C^*-algebra. They have to be projections a C^*-algebra which has got a trace.

We next discuss the characterization of a synchronous game with a perfect q-strategy. We need the following theorem about finite-dimensional C^*-algebra before proceeding into this discussion.

Theorem 10.15. *Let $\mathcal{A}$ be a finite-dimensional C^*-algebra. Then there are $n_1, ..., n_L \in \mathbb{N}$ and a unital $*$-isomorphism $\pi : \mathcal{A} \to M_{n_1} \oplus \cdots \oplus M_{n_L}$. That is to say, $\mathcal{A} \simeq M_{n_1} \oplus \cdots \oplus M_{n_L}$.*

Theorem 10.16. *Let $\mathcal{G} = (I, \mathcal{O}, \lambda)$ be a synchronous game. Then $\mathcal{G}$ has a perfect q-strategy if and only if there is m and projections $\{E_{v,i}\} \subseteq M_m$ such that $\sum_{i=1}^{k} E_{v,i} = I$ for all $v \in I$ and whenever $\lambda(v, w, i, j) = 0$ then $E_{v,i}E_{w,j} = 0$.*

Proof. The backward direction is trivial. For the forward direction, suppose that $\mathcal{G}$ has a perfect q-strategy $p(i, j|v, w) \in C_q^s(n, k)$. We may write $p(i, j|v, w) = \langle \varphi | E_{v,i} \otimes F_{w,j} \varphi \rangle$, where $E_{v,i}$ and $F_{w,j}$ are acting on a finite-dimensional Hilbert space. As before, we look at the algebra $\mathcal{A}$ generated by $\{E_{v,i}\}$; then $\mathcal{A}$ is finite-dimensional. Hence, $\mathcal{A} \simeq M_{n_1} \oplus \cdots \oplus M_{n_L}$. Now on $\mathcal{A}$, $\tau(X) = \langle \varphi | (X \otimes I)\varphi \rangle$ is a trace. But recall that any trace on $M_{n_1} \oplus \cdots \oplus M_{n_L}$ corresponds to $t_1, ..., t_L \geq 0$ with $t_1 + \cdots + t_L = 1$ such that if $X = X_1 \oplus \cdots \oplus X_L \in \mathcal{A}$ (where $X_i \in M_{n_i}$), then $\tau(X) = t_1 \mathrm{Tr}_{n_1}(X_1) + \cdots + t_L \mathrm{Tr}_{n_L}(X_L)$. For simplicity, assume that $t_1 \neq 0$. Write $E_{v,i} = E_{v,i}^1 \oplus \cdots \oplus E_{v,i}^L$ where each block $E_{v,i}^{\ell}$ is in M_{n_ℓ}. Then each block $E_{v,i}^{\ell}$ is still a projection. Now, if $\lambda(v, w, i, j) = 0$ then $\tau(E_{v,i}E_{w,j}) = 0 = \sum_{\ell=1}^{L} t_\ell \mathrm{Tr}_{n_\ell}(E_{v,i}^{(\ell)} E_{w,j}^{(\ell)})$ so that $\mathrm{Tr}_{n_1}(E_{v,i}^{(1)} E_{w,j}^{(1)}) = 0$. Both these matrices are positive, so $E_{v,i}^{(1)} E_{w,j}^{(1)} = 0$. Hence, we pick $m = n_1$ and the set $\{E_{v,i}^{(1)}\}$ gives us the desired result. □

Corollary 10.17. *Let G be a graph. Then $\chi_q(G) \leq k$ if and only if there exists an m, projections $\{E_{v,i} : v \in V, 1 \leq i \leq k\} \subseteq M_m$ such that $\sum_i E_{v,i} = I_m$ and whenever $v \sim w$, we have $E_{v,i}E_{w,i} = 0$ for all i.*

11. Tsirelson's Correlation Matrices

Tsirelson's idea was to work with self-adjoint unitaries (popularly known as reflections) on the set set $\{+1, -1\}$ under the multiplication operation instead of working with projections on the set $\{0, 1\}$ under the addition operation. Recall that the spectrum of projections consists of 0 and 1, and the spectrum of self-adjoint unitaries consists of 1 and -1.

Definition 11.1 (Quantum Correlation Matrix)**.** An $n \times n$ matrix $(c_{s,t})_{s,t=1}^{n}$ of real numbers is called a *quantum correlation matrix* if there exists $A_s = A_s^* \in M_p$ with $-I \leq A_s \leq I$ and there exists $B_t = B_t^* \in M_q$ with $-I \leq B_t \leq I$, along with the existence of $\psi \in \mathbb{C}^p \otimes \mathbb{C}^q$ such that $\|\psi\| = 1$ and $c_{s,t} = \langle \psi | (A_s \otimes B_t)\psi \rangle$. Let $\mathrm{Cor}_q(n) \subseteq M_n(\mathbb{R})$ be the set of all such matrices. Similarly, let $\mathrm{Cor}_{qc}(n) \subseteq M_n(\mathbb{R})$ be the set of all matrices of the form $(c_{s,t}) \in M_n$, where $c_{s,t} = \langle \psi | (A_s \otimes B_t)\psi \rangle$ for some $A_s = A_s^*$, $B_t = B_t^* \in \mathcal{B}(\mathcal{H})$ with $-I \leq A_s \leq I$ and $-I \leq B_t \leq I$ and $\psi \in \mathcal{H}$ with $\|\psi\| = 1$, such that $A_s B_t = B_t A_s$ for all s, t.

Proposition 11.2 (Disambiguation)**.** *A matrix $(c_{s,t})_{s,t=1}^{n}$ is in $Cor_q(n)$ if and only if there are matrices $A_s = A_s^* \in M_p$ and $B_t = B_t^* \in M_q$ with $A_s^2 = I$ and $B_t^2 = I$, along with $\psi \in \mathbb{C}^p \otimes \mathbb{C}^q$ with $\|\psi\| = 1$, such that $c_{s,t} = \langle \psi | A_s \otimes B_t \psi \rangle$.*

Remark 11.3. The above proposition tells us that no matter whether we choose projections or self-adjoint unitaries, the sets we get is the same.

Proof. We use the Halmos trick. For $H = H^*$ and $-I \leq H \leq I$, we may set

$$U = \begin{pmatrix} H & \sqrt{I - H^2} \\ \sqrt{I - H^2} & -H \end{pmatrix}$$

so that $U = U^*$ and $U^2 = I$. We now set

$$U_s = \begin{pmatrix} A_s & \sqrt{I - A_s^2} \\ \sqrt{I - A_s^2} & -A_s \end{pmatrix} \text{ and } V_t = \begin{pmatrix} B_t & \sqrt{I - B_t^2} \\ \sqrt{I - B_t^2} & -B_t \end{pmatrix}.$$

Then $U_s^2 = I$ and $V_t^2 = I$. Let $\tilde{\psi} = \begin{pmatrix} \psi \\ 0 \end{pmatrix}$, so that $\langle \psi | A_s \otimes B_t \psi \rangle = \langle \tilde{\psi} | (U_s \otimes V_t) \tilde{\psi} \rangle$. The other inclusion of sets is obvious, so we are done. □

We will see that in the context of working with reflections, we actually have $\mathrm{Cor}_q(n) = \mathrm{Cor}_{qc}(n)$. First, we need a bit of background on Clifford unitaries.

Definition 11.4 (Clifford Unitaries)**.** A family of $n \times n$ self-adjoint unitary matrices that anti-commute is described as *Clifford unitaries*.

So, a set of matrices $\{X_1, ..., X_m\}$ is Clifford unitaries if we have $X_i = X_i^*$, $X_i^2 = I$ for every i and $X_i X_j = -X_j X_i$ for all $i \neq j$. For the sake of completeness, here is a construction of such a set of unitaries. Let

$$X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

so that

$$XZ = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \text{ and } ZX = -XZ.$$

This set consists of $m = 2$ self-adjoint unitaries that anti-commute and hence is a set of Clifford unitaries. Suppose $m > 2$ and we want to construct a set of Clifford unitaries that has cardinality m. Then we let,

$$C_1 = Z \otimes \underbrace{I_2 \otimes \cdots \otimes I_2}_{m-1 \text{ times}},$$

$$C_2 = X \otimes Z \otimes I_2 \otimes \cdots \otimes I_2.$$

Similarly, for any i we let

$$C_i = \underbrace{X \otimes \cdots \otimes X}_{i-1 \text{ times}} \otimes Z \otimes \underbrace{I_2 \otimes \cdots \otimes I_2}_{m-i \text{ times}}.$$

Then each $C_i = C_i^*$ and $C_i^2 = I$. If $i < j$, then

$$C_i C_j = X^2 \otimes \cdots \otimes X^2 \otimes ZX \otimes X \otimes \cdots \otimes X \otimes I...$$

and

$$C_j C_i = X^2 \otimes \cdots \otimes X^2 \otimes XZ \otimes X \otimes ... \otimes X \otimes I...$$

so that $C_i C_j = -C_j C_i$. (Note that all the C_i's are real matrices and hence $C_i = C_i^t$ for every i.)

Lemma 11.5. *If $A, B \in M_d$ and $\psi = \frac{1}{\sqrt{d}}(e_1 \otimes e_1 + \cdots + e_d \otimes e_d)$, then $\langle \psi | (A \otimes B)\psi \rangle = tr(AB^t)$. (Note: "tr" denotes normalized trace.)*

Proof. Let $A = (a_{ij})$ and $B = (b_{ij})$. Then

$$\begin{aligned}
\langle \psi | (A \otimes B)\psi \rangle &= \frac{1}{d} \sum_{i,j=1}^{d} \langle e_i \otimes e_i | (A \otimes B) e_j \otimes e_j \rangle \\
&= \frac{1}{d} \sum_{i,j=1}^{d} \langle e_i | A e_j \rangle \langle e_i | B e_j \rangle \\
&= \frac{1}{d} \sum_{i,j=1}^{d} a_{ij} b_{ij} = \frac{1}{d} \mathrm{Tr}(AB^t).
\end{aligned}$$

Hence, $\langle \psi | (A \otimes B)\psi \rangle = \mathrm{tr}(AB^t)$. □

Theorem 11.6 (Tsirelson, 1987)**.** *Let $C = (c_{s,t})_{s,t=1}^n$ be a real $n \times n$ matrix. The following are equivalent.*

1. *$C \in Cor_q(n)$.*
2. *There is m and vectors $x_s, y_t \in \mathbb{R}^n$ with $\|x_s\| \leq 1$ and $\|y_t\| \leq 1$ such that $C = (\langle x_s, y_t \rangle)$. Moreover, m can always be taken to be $m = 4n$.*

Consequently, $Cor_q(n) = Cor_{qc}(n) = \{(\langle x_s, y_t \rangle) : x_s, y_t \in \mathbb{R}^{4n}, \|x_s\| \leq 1, \|y_t\| \leq 1\}$.

Proof. Let us show that (i) implies (ii). Write $c_{s,t} = \langle \psi | (A_s \otimes B_t)\psi \rangle = \langle (A_s \otimes I)\psi | (I \otimes B_t)\psi \rangle$. Let $x'_s = (A_s \otimes I)\psi$ and $y'_t = (I \otimes B_t)\psi$. Then $\|x'_s\| \leq 1$ and $\|y'_t\| \leq 1$, while $c_{s,t} = \langle x'_s, y'_t \rangle$. Now, $\mathrm{span}\{x'_s, y'_t\}_{s,t=1}^n$ is m-dimensional where $m \leq 2n$. Identify this subspace of $\mathbb{C}^{2n}$ with $\mathbb{C}^m$, and write $x'_s = (x'_s(1), ..., x'_s(m))$ and $y'_t = (y'_t(1), ..., y'_t(m))$. Then write $x'_s(k) = a_s(k) + i b_s(k)$ where $a_s, b_s \in \mathbb{R}$. Similarly, write $y'_t(k) = \alpha_t(k) +$

$i\beta_t(k)$. Then define $x_s = (a_s(1), -b_s(1), ..., a_s(m), -b_s(m)) \in \mathbb{R}^{2m}$ and define $y_t = (\alpha_t(1), \beta_t(1), ..., \alpha_t(m), \beta_t(m)) \in \mathbb{R}^{2m}$. Then $\|x_s\| = \|x'_s\|$ and $\|y_t\| = \|y'_t\|$. Finally, since $\langle x'_s, y'_t\rangle \in \mathbb{R}$ we still have $\langle x'_s, y'_t\rangle = \langle x_s, y_t\rangle$. Note that $2m \leq 4n$, but by adding 0's we can always make these vectors into $(4n)$-tuples, as desired.

Now we show that (ii) implies (i). Let $x_s, y_t \in \mathbb{R}^n$ for $s, t = 1, ..., n$ be such that $\|x_s\| \leq 1$, $\|y_t\| \leq 1$ and $c_{s,t} = \langle x_s, y_t\rangle$. Now we have to build self-adjoint operators and unit vectors. Write $x_s = (x_s(1), ..., x_s(m))$ and $y_t = (y_t(1), ..., y_t(m))$. Take m Clifford unitaries $C_1, ..., C_m$ on $\mathbb{C}^d$, and set $A_s = \sum_{i=1}^m x_s(i)C_i$. Let $B_t = \sum_{i=1}^m y_t(i)C_i$, so that $A_s = A_s^*$ and $B_t = B_t^*$. Note that

$$A_s^2 = \sum_{i,j=1}^m x_s(i)x_s(j)C_iC_j = \sum_{i=1}^m x_s(i)^2C_i^2 + \sum_{i<j}(x_s(i)x_s(j) - x_s(i)x_s(j))C_iC_j = \left(\sum_{i=1}^m x_s(i)^2\right) \leq I.$$

Hence, $\sigma(A_s) \subseteq [-1, 1]$ so that $-I \leq A_s \leq I$. Similarly, $-I \leq B_t \leq I$. Denote by D^T the transpose of a matrix D. Then

$$A_sB_t^T = \sum_{i,j=1}^m x_s(i)y_t(j)C_iC_j^T = \left[\sum_{i=1}^m x_s(i)y_t(i)C_iC_i^t\right] + \sum_{i\neq j} x_s(i)y_t(j)C_iC_j.$$

Thus, $A_sB_t^T = \sum_{i=1}^m x_s(i)y_t(i)I + \text{stuff}$. Taking the normalized trace gives $\sum_{i=1}^m x_s(i)y_t(i) + \text{stuff}$. Now note that if $i \neq j$, then $\text{Tr}(C_iC_j) = \text{Tr}(C_jC_i)$ and $\text{Tr}(C_iC_j) = \text{Tr}(-C_jC_i) = -\text{Tr}(C_jC_i)$ so that $\text{Tr}(C_iC_j) = 0$. Hence, $\text{Tr}(A_sB_t^T) = \sum_{i=1}^m x_s(i)y_t(i) = \langle x_s, y_t\rangle$. By the lemma, if $\psi = \frac{1}{\sqrt{d}}(e_1 \otimes e_1 + \cdots + e_d \otimes e_d)$ then $\langle x_s, y_t\rangle = \langle\psi|(A_s \otimes B_t)\psi\rangle$, and this gives the desired result. □

Now, let us go back to the other situation and see how things differ. Let us consider the sets $C_q(n, 2)$ and $C_{qc}(n, 2)$. In this case Alice has PVM's $\{E_{s,0}, E_{s,1}\}_{1\leq s\leq m}$ with $E_{s0} + E_{s1} = I$, and Bob has PVM's $\{F_{t,0}, F_{t,1}\}_{1\leq t\leq n}$ with $F_{t0} + F_{t1} = I$. If we had $A_s = A_s^*$ with $A_s^2 = I$ and $B_t^* = B_t$ with $B_t^2 = I$, then $\sigma(A_s), \sigma(B_t) \subseteq \{\pm 1\}$. In this case we let $E_{s0} = \frac{I+A_s}{2}$, $E_{s1} = \frac{I-A_s}{2}$, $F_{s0} = \frac{I+B_t}{2}$, $F_{s1} = \frac{I-B_t}{2}$, so that

$$E_{si} = \frac{I + (-1)^iA_s}{2},\ F_{tj} = \frac{I + (-1)^jB_t}{2}.$$

We obtain $p(i, j|s, t) \in C_q(n, 2)$ given by

$$p(i, j|s, t) = \langle\psi|(E_{si} \otimes F_{tj})\psi\rangle = \frac{1}{4}\langle\psi|(I + (-1)^iA_s \otimes I + (-1)^jI \otimes B_t + (-1)^{i+j}A_s \otimes B_t)\psi\rangle.$$

If we use the Clifford construction from earlier, then

$$p(i,j|s,t) = \frac{1}{4}\mathrm{Tr}(I + (-1)^i A_s + (-1)^j B_t^T + (-1)^{i+j} A_s B_t^T)$$
$$= \frac{1}{4}(1 + 0 + 0 + (-1)^{i+j}\langle x_s, y_t\rangle) = \frac{1}{4} + \frac{1}{4}(-1)^{i+j}\langle x_s, y_t\rangle,$$

since A_s and B_t are linear combination of Clifford unitaries and we know that the Clifford unitaries constructed earlier are of trace zero thereby rendering the trace of A_s and trace of B_t zero.

The above discussion can be summarized into the following theorem.

Theorem 11.7. *Let x_s, y_t be vectors in $\mathbb{R}^m$ for $1 \leq s,t \leq n$ with $\|x_s\| \leq 1$ and $\|y_t\| \leq 1$, such that $\langle x_s, y_t\rangle \in \mathbb{R}$ for all s,t. Then there exists $p(i,j|s,t) \in C_q(n,2)$ such that $p(i,j|s,t) = \frac{1}{4}[1+(-1)^{i+j}\langle x_s, y_t\rangle]$.*

Note that if p is as in the theorem, then $p(0,0|s,t) = p(1,1|s,t)$ and $p(0,1|s,t) = p(1,0|s,t)$ for all s,t.

Theorem 11.8. *Let $p(i,j|s,t) \in C_{vect}(n,2)$ be such that $p(0,0|s,t) = p(1,1|s,t)$ and $p(0,1|s,t) = p(1,0|s,t)$ for all s,t. Then*

1. *There exist vectors x_s, y_t for $1 \leq s,t \leq n$ such that $\|x_s\|, \|y_t\| \leq 1$ and $p(i,j|s,t) = \frac{1}{4}[1 + (-1)^{i+j}\langle x_s, y_t\rangle]$.*
2. *$p(i,j|s,t) \in C_q(n,2)$.*

Proof. The second claim follows from the first by the previous theorem. To show (1), note that since $p \in C_{\mathrm{vect}}(n,2)$, there exist vectors $v_{s0}, v_{s1}, w_{t0}, w_{t1}$ such that $v_{s0} \perp v_{s1}$ and $w_{t0} \perp w_{t1}$ for all s,t, and there exists a vector ψ of norm 1 such that $v_{s0} + v_{s1} = w_{t0} + w_{t1} = \psi$ for all s,t, while $p(i,j|s,t) = \langle v_{si}, w_{tj}\rangle \geq 0$. (The idea behind this is: if we have two orthogonal projections applied to a state vector, then the image will consist of two vectors which will add up to the state vector.) Note that

$$1 = \sum_{i,j\in\{0,1\}} p(i,j|s,t) = 2(p(0,0|s,t) + p(1,0|s,t))$$
$$= 2[\langle v_{s0}, w_{t0}\rangle + \langle v_{s1}, w_{t0}\rangle]$$
$$= 2[\langle \psi, w_{t0}\rangle]$$
$$= 2\langle w_{t0}, w_{t0}\rangle.$$

Hence, $\|w_{t0}\|^2 = \frac{1}{2}$. Similarly, $\|w_{t1}\|^2 = \|v_{s0}\|^2 = \|v_{s1}\|^2 = \frac{1}{2}$. Now set $x_s = v_{s0} - v_{s1}$ and set $y_t = w_{t0} - w_{t1}$. Then $\|x_s\|^2 = \|v_{s0}\|^2 + \|v_{s1}\|^2 = 1$, and similarly, $\|y_t\|^2 = 1$. To see that p is of the desired form, we can

compute

$$\begin{aligned}\frac{1}{4}[1+\langle x_s, y_t\rangle] &= \frac{1}{4}[1+\langle v_{s0}-v_{s1}, w_{t0}-w_{t1}\rangle] \\ &= \frac{1}{4}[1+\langle v_{s0}, w_{t0}\rangle + \langle v_{s1}, w_{t1}\rangle - \langle v_{s1}, w_{t0}\rangle - \langle v_{s0}, w_{t1}\rangle] \\ &= \frac{1}{4}[1+2p(0,0|s,t) - (1-2p(0,0|s,t))] \\ &= p(0,0|s,t) = p(1,1|s,t).\end{aligned}$$

Therefore,

$$p(0,0|s,t) = p(1,1|s,t) = \frac{1}{4}[1+(-1)^{i+j}\langle x_s, y_t\rangle],$$

when $(i,j) = (0,0)$ or $(i,j) = (1,1)$.

Now suppose that $(i,j) = (1,0)$ or $(i,j) = (0,1)$. Then similarly, one can check that

$$\begin{aligned}\frac{1}{4}[1-\langle x_s, y_t\rangle] &= \frac{1}{4}[1+2p(1,0|s,t) - (1-2p(1,0|s,t))] \\ &= p(1,0|s,t) = p(0,1|s,t)\end{aligned}$$

Therefore, (1) holds. □

Theorem 11.9. *Let $\mathcal{G}$ be an n-input, 2-output game, with rule function $\lambda(i,j|s,t) \in \{0,1\}$. Assume that the rules satisfy $\lambda(i,j|s,t) = \lambda(i+1, j+1|s,t)$. The following are equivalent.*

(1) *$\mathcal{G}$ has a perfect q-strategy.*
(2) *$\mathcal{G}$ has a perfect vect-strategy.*
(3) *There are vectors x_s, y_t with $\|x_s\|, \|y_t\| \leq 1$ such that $\langle x_s, y_t\rangle \in \mathbb{R}$ and such that whenever $\lambda(i,j|s,t) = 0$, then $\langle x_s, y_t\rangle = (-1)^{i+j+1}$.*

Proof. Clearly (1) implies (2). Assume that (2) holds, and let $p(i,j|s,t) \in C_{\text{vect}}(n,2)$ be perfect; i.e. whenever $\lambda(i,j|s,t) = 0$ then $p(i,j|s,t) = 0$. Let

$$\tilde{p}(i,j|s,t) = \frac{1}{2}[p(i,j|s,t) + p(i+1,j+1|s,t)].$$

Since $p(i,j|s,t), p(i+1,j+1|s,t) \in C_{\text{vect}}(n,2)$, and $C_{\text{vect}}(n,2)$ is convex, we have $\tilde{p}(i,j|s,t) \in C_{\text{vect}}(n,2)$. If $\lambda(i,j|s,t) = 0$ then $\lambda(i+1,j+1|s,t) = 0$. Therefore, $\tilde{p}(i,j|s,t) = 0$ so that $\tilde{p}$ is a perfect vect-strategy. Moreover, $\tilde{p}(0,0|s,t) = \tilde{p}(1,1|s,t)$ and $\tilde{p}(1,0|s,t) = \tilde{p}(0,1|s,t)$. Thus, there are x_s, y_t with $\|x_s\|, \|y_t\| \leq 1$ such that $\tilde{p}(i,j|s,t) = \frac{1}{4}[1+(-1)^{i+j}\langle x_s, y_t\rangle]$. If $\lambda(i,j|s,t) = 0$, then $\frac{1}{4}[1+(-1)^{i+j}\langle x_s, y_t\rangle] = 0$. Thus, $\langle x_s, y_t\rangle = (-1)^{i+j+1}$. This gives (3).

If (3) holds, then given such vectors x_s, y_t, we may set $p(i,j|s,t) = \frac{1}{4}[1+(-1)^{i+j}\langle x_s, y_t\rangle] \in C_q(n,2)$. One can check that this gives a perfect q-strategy by the same calculation as above. Hence, (1) holds. □

Remark 11.10. Since $\|x_s\|, \|y_t\| \leq 1$ as above and $\langle x_s, y_t\rangle = (-1)^{i+j+1}$, we must have $\|x_s\| = \|y_t\| = 1$ and hence $y_t = (-1)^{i+j+1}x_s$ by the Cauchy-Schwarz inequality.

References

[1] D. Avis, J. Hagesawa, Y. Kikuchi, and Y. Sasaki, A quanutm protocol to win the graph coloring game on all hadamard graphs, IEICE Trans. Fundam. elec- tron. Commun. Comput. Sci., E89-A(5):1378-1381, 2006. arxiv:quant-ph/0509047v4, doi:10.1093/ietfec/e89-a.5.1378.

[2] Jop Briet, Harry Buhrman, Monique Laurent, Teresa Piovesan, Giannicola Scarpa, *Entanglement-assisted zero-error source-channel coding*, arXiv:1308.4283

[3] P.J. Cameron, A. Montanaro, M.W. Newman, S. Severini and A. Winter, On the quantum chromatic number of a graph, The electronic journal of combinatorics 14 (2007), R81, arxiv:quant-ph/0608016.

[4] T. Cubitt, L. Man?cinska, D. Roberson, S. Severini, D. Stahlke, and A. Winter, Bounds on entanglement assisted source-channel coding via the Lova?sz ? number and its variants, IEEE Trans. Inf. Theory 60 (2013), arXiv:1310.7120v1.

[5] R.G. Douglas, *On majorization, factorization, and range inclusion of operators on Hilbert space*, Proc. Amer. Math. Soc. 17 1966, 413?415.

[6] R. Duan, S. Severini, and A. Winter, *Zero-Error Communication via Quantum Channels, Noncommutative Graphs, and a Quantum Lovász Number*, IEEE Transactions on Information Theory, vol. 59, no. 2, pp. 11641174, Feb. 2013.

[7] Dykema, Kenneth J.; Paulsen, Vern *Synchronous correlation matrices and Connes' embedding conjecture*, J. Math. Phys. 57 (2016), no. 1, 015214, 12 pp.

[8] P. Frankl, *Orthogonal vectors in the n-dimensional cube and codes with missing distances*, Combinatorica 6 (1986), no. 3, 279?285.

[9] P. Frankl and V. Rodl, *Frobidden intersections*, Transactions of the AMS, Vol. 300, No. 1, 1987, pp. 259-286.

[10] E. de Klerk, D. V. Pasechnik, *A note on the stability number of an orthogonality graph*, arXiv:0505038v3

[11] L. Lovasz, *On the Shannon capacity of a graph*, IEEE Transactions on Information Theory, vol. 25, no. 1, pp. 17, Jan. 1979.

[12] N. Ozawa, *About the Connes' embedding problem: algebraic approaches*, Japan. J. Math. 8 (2013), no. 1, 147-183.

[13] V. I. Paulsen, *Completely Bounded Maps and Operator Algebras*, Cambridge University Press, 2002.

[14] V.I. Paulsen, S. Severini, D. Stahlke, I.G. Todorov, A. Winter, *Estimating quantum chromatic numbers*, J. Func. Anal.,

[15] V. I. Paulsen and I. G. Todorov, *Quantum chromatic numbers via operator systems*, Quarterly J. Math. 66 (2015), no. 2, 677-692.

[16] David E Roberson, *Variations on a Theme: Graph Homomorphisms*, PhD thesis, University of Waterloo, 2013.

[17] David E. Roberson and Laura Mancinska, *Graph homomorphisms for quantum players*, preprint, arXiv:1212.1724, 2012.
[18] G. Scarpa and S. Severini, *Kochen-Specker sets and the rank-1 quantum chromatic number*, IEEE Trans. Inf. Theory, 58 (2012), no. 4, 2524-2529.
[19] V. B. Scholz and R. F. Werner, *Tsirelson's Problem*, preprint, arXiv:0812.4305v1, 2008.
[20] D. Spielman, *Spectral Graph Theory*, online lecture notes, http://www.cs.yale.edu/homes/spielman/561/ .
[21] M. Szegedy, *A note on the number of Lovasz and the generalized Delsarte bound*, in Proc. 35th Annual Symposium on Foundations of Computer Science, 1994, 1994, pp. 36?39.
[22] B. S. Tsirelson, *Quantum generalizations of Bell's inequality*, Lett. Math. Phys., 4 (1980), no. 4, 93-100.
[23] B. S. Tsirelson, *Some results and problems on quantum Bell-type inequalities*, Hadronic J. Suppl., 8 (1993), no. 4, 329-345.

E-mail address: satish.pandey@uwaterloo.ca, sj2harri@uwaterloo.ca