1. **Non-uniqueness of the decomposition of mixed states.** (4 Points: 2+2)

Consider two macroscopically different preparation schemes of a large number of polarised photons:

Preparation A. For each photon we toss a fair coin. Depending on whether we get head or tail, we prepare the photon to have either vertical or horizontal *linear* polarisation.

Preparation B. For each photon we toss a fair coin. Depending on whether we get head or tail, we prepare the photon to have either left-handed or right-handed *circular* polarisation.

Note: You can simply think of the polarization of the light as a binary variable and of the polarization axis as a local basis. I.e. the vertical and horizontal linear polarizations may be identified with the $|0\rangle$ and $|1\rangle$ eigen states of the Z basis. Likewise you may interpret the left- and right handed circular polarizations as the $|+\rangle$ and $|-\rangle$ eigen states of the X basis.

Now we are given a large number of photons which all were prepared by the same scheme.

a) Argue that having only access to the photons we can not distinguish which of the preparation schemes was used.

PREP. A $\Rightarrow$ $\rho_1 = \underbrace{\text{Prob}(\text{HEAD})}_{=\frac{1}{2}} |0\rangle\langle 0| + \underbrace{\text{Prob}(\text{TAIL})}_{=\frac{1}{2}} |1\rangle\langle 1|$

$$= \frac{1}{2}\left(|0\rangle\langle 0| + |1\rangle\langle 1|\right) = \frac{\mathbb{1}}{2}$$

PREP. B $\Rightarrow$ $\rho_2 = \frac{1}{2}|+\rangle\langle +| + \frac{1}{2}|-\rangle\langle -| = \frac{1}{2}\left(|+\rangle\langle +| + |-\rangle\langle -|\right) = \frac{\mathbb{1}}{2}$

$\Rightarrow$ $\rho_1 = \rho_2$ $\Rightarrow$ indistinguishable

b) Argue that if it were possible to distinguish such types of preparations by measuring the photon, locality would be violated.
(*Hint*: think about how the state we consider can be prepared by ignoring one degree of freedom of a bipartite system as in the last exercise of Sheet 0.)

• For ABSURDUM:

Assumption (*)

Suppose one can distinguish $\rho_1 = \frac{1}{2}|0\rangle\langle 0| + \frac{1}{2}|1\rangle\langle 1|$ and $\rho_2 = \frac{1}{2}|+\rangle\langle +| + |-\rangle\langle -|$, then locality would be violated (one cannot communicate faster than light).

PROOF:

• Suppose we have a state $|\psi\rangle_{AB} = \frac{1}{\sqrt{2}}\left(|0\rangle_A \otimes |0\rangle_B + |1\rangle_A \otimes |1\rangle_B\right)$ between Alice and Bob that are (very) far apart.

• Bob wants to communicate $\boxed{0}$ or $\boxed{1}$ to Alice, and would like to use the strategy of ASSUMPTION (*).

He wants to do some "operation" on $|\psi\rangle_{AB}$ such that at the end Alice has $\rho_1 = \frac{1}{2}\left(|0\rangle\langle 0| + |1\rangle\langle 1|\right)$ or $\rho_2 = \frac{1}{2}\left(|+\rangle\langle +| + |+\rangle\langle +|\right)$.

Associated with "$\boxed{0}$"

Associated with "$\boxed{1}$"

• This protocol does the job:

1) If Bob wants to communicate "$\boxed{0}$", then he measure his qubit in basis $\{|0\rangle, |1\rangle\}$

2) If Bob wants to communicate "$\boxed{1}$", then he measure his qubit in basis $\{|+\rangle, |-\rangle\}$

In fact:

CASE 1) $\Rightarrow$ $\rho_{AB}^{\text{AFTER MEAS.}} = \text{Prob}(|0\rangle_B)\left(|0\rangle_A \otimes |0\rangle_B\right)\left(\langle 0|_A \otimes \langle 0|_B\right) + \text{Prob}(|1\rangle_B)\left(|1\rangle_A \otimes |1\rangle_B\right)\left(\langle 1|_A \otimes \langle 1|_B\right)$

$|\psi\rangle_{AB} = \frac{1}{\sqrt{2}}\left(|0\rangle_A \otimes |0\rangle_B + |1\rangle_A \otimes |1\rangle_B\right)$

$= \frac{1}{2}\left(|0\rangle_A \otimes |0\rangle_B\right)\left(\langle 0|_A \otimes \langle 0|_B\right) + \frac{1}{2}\left(|1\rangle_A \otimes |1\rangle_B\right)\left(\langle 1|_A \otimes \langle 1|_B\right)$

$\Rightarrow \rho_A^{\text{AFTER MEAS.}} = \text{tr}_B\left(\rho_{AB}^{\text{AFTER MEAS.}}\right) = \frac{1}{2}|0\rangle\langle 0| + \frac{1}{2}|1\rangle\langle 1| \rightsquigarrow$ associated with "$\boxed{0}$"

CASE 2) $\Rightarrow$ $\rho_{AB}^{\text{AFTER MEAS.}} = \text{Prob}(|+\rangle_B)\left(|+\rangle_A \otimes |+\rangle_B\right)\left(\langle +|_A \otimes \langle +|_B\right) + \text{Prob}(|-\rangle_B)\left(|-\rangle_A \otimes |-\rangle_B\right)\left(\langle -|_A \otimes \langle -|_B\right)$

$$= \frac{1}{2}\left(|+\rangle_A \otimes |+\rangle_B\right)\left(\langle +|_A \otimes \langle +|_B\right) + \frac{1}{2}\left(|-\rangle_A \otimes |-\rangle_B\right)\left(\langle -|_A \otimes \langle -|_B\right)$$

$\text{Prob}(|+\rangle_A) = \text{Tr}(|+\rangle_A\langle +|_A \, \rho_{AB}) = \frac{1}{2}$

$\rho_{AB} = |\psi_{AB}\rangle\langle\psi_{AB}|$

$|\psi_{AB}\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$

$\Rightarrow \rho_A^{\text{AFTER MEAS.}} = \text{tr}_B\left(\rho_{AB}^{\text{AFTER MEAS.}}\right) = \frac{1}{2}|+\rangle\langle +| + \frac{1}{2}|-\rangle\langle -|$ $\rightsquigarrow$ associated with "$\boxed{1}$"

$\Longrightarrow$ So Bob would be able to use ASSUMPTION (*) to communicate bit "$\boxed{0}$" or "$\boxed{1}$" faster than light.

2. **Impossible machines – no cloning.** (5 Points)

In this problem we will re-derive the impossibility results that you have seen in the lecture but now directly using the structure of quantum theory.

Show that there does not exist a unitary map on two copies of a Hilbert space $\mathcal{H}$ which acts in the following way:

$$\forall |\psi\rangle \in \mathcal{H} : U|\psi\rangle|0\rangle = e^{i\phi(\psi)}|\psi\rangle|\psi\rangle .$$

(*Hint*: Unitary operators are linear.)

For absurdum:

If $\exists\, U : \forall |\psi\rangle \in \mathcal{H} \quad U|\psi\rangle \otimes |0\rangle = e^{i\phi(\psi)}|\psi\rangle \otimes |\psi\rangle$

Then $\langle\psi|\phi\rangle^2 = \langle\psi|\phi\rangle\langle\psi|\phi\rangle = (\langle\psi|\otimes\langle\psi|)|\phi\rangle\otimes|\phi\rangle =$

$= e^{i(\phi(\psi)-\phi(\phi))}\langle\psi|\otimes\langle 0| U^\dagger U |\phi\rangle \otimes |0\rangle = e^{i(\phi(\psi)-\phi(\phi))}\langle\psi|\phi\rangle$

$\Rightarrow \langle\psi|\phi\rangle\left(\langle\psi|\phi\rangle - e^{i(\phi(\psi)-\phi(\phi))}\right) = 0 \Rightarrow \langle\psi|\phi\rangle = 0 \text{ or } e^{i(\phi(\psi)-\phi(\phi))}$

$\Rightarrow |\langle\psi|\phi\rangle| = 0, 1 \quad \forall |\psi\rangle, |\phi\rangle \in \mathcal{H}$

$\Rightarrow$ ABSURD because there are states $|\psi\rangle, |\phi\rangle$ such that $|\langle\psi|\phi\rangle| \neq 0, 1$.

$|\psi\rangle = |0\rangle$
$|\phi\rangle = \frac{|0\rangle + |1\rangle}{\sqrt{2}}$
$\langle\psi|\phi\rangle = \frac{1}{\sqrt{2}} \neq 0, 1$

3. **The most general quantum measurements.** (4 Points: 2+1+1)

   In a quantum mechanics course, measurements are typically introduced as projective measurements of the eigenvalues of observables. But from a theoretical perspective another measurement description is often helpful. For simplicity—and in the spirit of information theory—we assume that the possible measurement outcomes are from a discrete set $\mathcal{X}$. [1]

A measurement with outcomes $\mathcal{X}$ on a quantum system with Hilbert space $\mathcal{H}$ can be described by a *positive operator valued measure* (POVM) on $\mathcal{X}$. We denote by $\mathrm{Pos}(\mathcal{H}) := \{A \in L(\mathcal{H}) \mid A \geq 0\}$ the set of Hermitian positive semi-definite operators on $\mathcal{H}$. A POVM on a discrete space $\mathcal{X}$ is a map $\mu : \mathcal{X} \to \mathrm{Pos}(\mathcal{H})$ such that $\sum_{x \in \mathcal{X}} \mu(x) = \mathrm{Id}$. If the system is in the quantum state $\rho \in \mathcal{D}(\mathcal{H})$, the probability of observing the outcome $x \in \mathcal{X}$ is given by $\mathrm{Tr}(\mu(x)\rho)$.

a) What is the difference between POVM measurements and the measurement description using observables? (Here we refer to the measurement description using observables as the measurement process where the quantum state gets projected on the projector valued measure (PVM) corresponding to the spectral value that is measured during the measurement process as explained in the lecture).

- O observable $(O^\dagger = O)$ $\Rightarrow$ $O = \sum_{i=1}^{d} \lambda_i |V_i\rangle\langle V_i|$ $\Rightarrow$ $\sum_{i=1}^{d} \underbrace{\boxed{|V_i\rangle\langle V_i|}}_{\mu(i)} = \mathbb{1}$

$\Rightarrow$ $\mu(i)\,\mu(j) = \delta_{ij}\,\mu(i)$

$\Rightarrow$ $\mu(i) = |V_i\rangle\langle V_i|$ for $i = 1, \dots, d$ is a POVM, $\underbrace{\text{such that } \mu(i)\mu(j) = 0 \text{ if } i \neq j. \text{ and } \mu^2(i) = \mu(i).}$

So this is something more than a POVM.

We define it PVM

(projective valued measure)

It is often stated that this is the most general form of a quantum measurement. We want to understand this statement in more detail. So what could be regarded as the most general quantum measurement? One can start as follows: A (general) quantum measurement M with outcomes in $\mathcal{X}$ is a map that associates to each quantum state $\rho \in \mathcal{D}(\mathcal{H})$ a probability measure p_ρ on $\mathcal{X}$, i.e. $M : \rho \mapsto p_\rho$ with $p_\rho : \mathcal{X} \to [0, 1]$ such that $\sum_{x \in \mathcal{X}} p_\rho(x) = 1$.

b) Show that any POVM on $\mathcal{X}$ defines a general quantum measurements as defined above.

DEFINITION OF POVM

Given a discrete space $\mathcal{X}$, we define POVM a map $\mu(\cdot)$ such that

$\mu : \mathcal{X} \to \mathrm{Pos}(\mathcal{H})$ and $\sum_{x \in \mathcal{X}} \mu(x) = \mathbb{1}$.

We can define $\mathrm{Prob}(x) = \mathrm{tr}(\mu(x)\rho)$ given $\rho \in D(\mathcal{H})$.

DEFINITION OF GENERAL MEASUREMENT

Given a discrete space $\mathcal{X}$, we define GENERAL MEASUREMENT a map M such that $M : \rho \Rightarrow P_\rho$ where $P_\rho : \mathcal{X} \to [0,1]$ and $\sum_{x\in\mathcal{X}} P_\rho(x) = 1$

- POVM $\Rightarrow$ GENERAL MEASUREMENT.

PROOF:

I define $P_\rho(x) := tr(\mu(x)\rho)$ and I need to show that $\begin{cases} A)\ P_\rho(x) \in [0,1] \\ B)\ \sum_{x\in\mathcal{X}} P_\rho(x) = 1 \end{cases}$

- $P_\rho(x) = tr(\mu(x)\rho) \underset{\uparrow}{=} \sum_i \lambda_i\, tr(|v_i\rangle\langle v_i|\mu(x)) = \sum_i \underbrace{\lambda_i}_{\substack{\geq \\ 0}} \underbrace{\langle v_i|\mu(x)|v_i\rangle}_{\substack{\geq \\ 0}}$

$\rho = \sum_i \lambda_i |v_i\rangle\langle v_i|$ eigendec.

$\rho \geq 0$ $\qquad$ $\mu(x) \geq 0$

$\Rightarrow P_\rho(x) \geq 0$

- $P_\rho(x) = tr(\mu(x)\rho) = \sum_i \lambda_i \langle v_i|\mu(x)|v_i\rangle \leq \sum_i \lambda_i = 1$

$tr(\rho) = 1$

$\langle\psi|\mu(x)|\psi\rangle \leq 1 \quad \forall |\psi\rangle$

PROOF: $\sum_x \mu(x) = \mathbb{1} \Rightarrow \sum_x \langle\psi|\mu(x)|\psi\rangle = 1 \overset{\mu \geq 0}{\Rightarrow} \langle\psi|\mu(x)|\psi\rangle \leq 1$

- $\sum_x P_\rho(x) = 1$

PROOF: $\sum_x P_\rho(x) = \sum_x tr(\rho\,\mu(x)) = tr(\rho \underbrace{\sum_x \mu(x)}_{\mathbb{1}}) = tr(\rho) = 1$

c) Show that any general quantum measurements as defined above defines a unique POVM on $\mathcal{X}$.
(*Hint*: You may assume a general measurement M to be linear. Then interpret.)

DEFINITION OF POVM

Given a discrete space $\mathcal{X}$, we define POVM a map $\mu(\cdot)$ such that $\mu : \mathcal{X} \Rightarrow Pos(\mathcal{H})$ and $\sum_{x\in\mathcal{X}} \mu(x) = \mathbb{1}$.

We can define $Prob(x) = tr(\mu(x)\rho)$ given $\rho \in D(\mathcal{H})$.

DEFINITION OF GENERAL MEASUREMENT

Given a discrete space $\mathcal{X}$, we define GENERAL MEASUREMENT a map $\mathcal{M}$ such that

$\mathcal{M} : \rho \to P_\rho$ where $P_\rho : \mathcal{X} \to [0,1]$ and $\sum_{x \in \mathcal{X}} P_\rho(x) = 1$

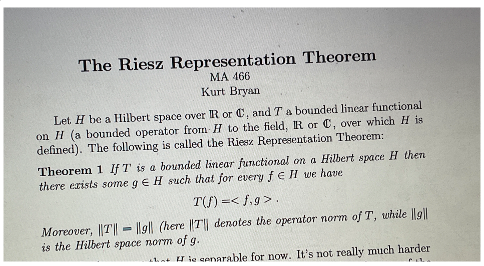
- POVM $\Leftarrow$ GENERAL MEASUREMENT.

PROOF:

Given $P_\rho(x)$, I define $\mu(x)$ such that it's the operator I get from the "RIESZ REPRESENTATION THEOREM."

$$P_\rho(x) \underset{\text{RIESZ THEOREM}}{=} \langle \mu(x) | \rho \rangle_{H.S.} = \operatorname{tr}\left(\mu(x)\rho\right)$$

$\langle A, B \rangle := \operatorname{tr}(A^+ B)$


The Riesz Representation Theorem
MA 466
Kurt Bryan

Let H be a Hilbert space over $\mathbb{R}$ or $\mathbb{C}$, and T a bounded linear functional on H (a bounded operator from H to the field, $\mathbb{R}$ or $\mathbb{C}$, over which H is defined). The following is called the Riesz Representation Theorem:

Theorem 1 *If T is a bounded linear functional on a Hilbert space H then there exists some $g \in H$ such that for every $f \in H$ we have*

$$T(f) = < f, g > .$$

Moreover, $\|T\| = \|g\|$ (here $\|T\|$ denotes the operator norm of T, while $\|g\|$ is the Hilbert space norm of g.

... that H is separable for now. It's not really much harder

$T(f) = P_\rho$

- I have $\sum_x \mu(x) = \mathbb{1}$

$$1 \underset{\text{ASSUMPTION}}{=} \sum_x P_\rho(x) = \operatorname{tr}\left(\left(\sum_x \mu(x)\right)\rho\right) \quad \forall \rho$$

$$\Downarrow$$

$$0 = \operatorname{tr}\left(\left(\sum_x \mu(x) - \mathbb{1}\right)\rho\right) \quad \forall \rho$$

$$\operatorname{tr}(A\rho) = 0 \quad \forall \rho \Rightarrow A = 0$$

$\rho = |\psi\rangle\langle\psi|$

$\operatorname{tr}(A\rho) = \langle\psi|A|\psi\rangle = 0 \quad \forall |\psi\rangle$

$\Rightarrow \langle\psi|A|\psi\rangle = 0 \in \mathbb{R} \Rightarrow A = A^+ \Rightarrow A$ diagonaliz. with 0 eigenvalues $\Rightarrow A = 0$

4. **Encoding classical bits.** (7 Points: 2+2+2+1)

In the last exercise we introduced the description of quantum measurements with the help of POVMs. Now, let $\mathcal{H}$ be a d-dimensional Hilbert space. Our aim is to encode n classical bits into the space of quantum states $\mathcal{D}(\mathcal{H})$. To this end, we choose a set of 2^n states $\{\rho_i\}_{i\in\{0,1\}^n} \subset \mathcal{D}(\mathcal{H})$, each state corresponding to a bit string. To decode the bit string we have to make a measurement described by a POVM $F = \{F_i\}_{i\in\{0,1\}^n}$, where the bit string is the outcome. In this exercise we are going to investigate the following question:

How many classical bits can be encoded and (perfectly) decoded in a d-dimensional quantum system in this way?

Consider a source that outputs the bit string $x \in \{0,1\}^n$ with probability $p(x)$.

a) We say that the decoder is successful if outcome i is returned upon measuring F on ρ_i. Define the expected success probability of the decoder with respect to the distribution p.

$p_{(00000)}, p_{(00001)}, \ldots, p_{(11111)}$

$$P_{succ} := \sum_{x\in\{0,1\}^n} \mathrm{Prob}(x|\rho_x)\cdot\mathrm{Prob}(\rho_x)$$

$$= \sum_{x\in\{0,1\}^n} \mathrm{tr}(\rho_x F_x)\,\mathrm{Prob}(x)$$

b) Prove the technical Lemma that $\rho \leq \mathbb{1}$ for ρ a density matrix.

We need to show that $(\mathbb{1} - \rho) \geq 0$.

If we diagonalize $\rho = \sum_i \lambda_i |v_i\rangle\langle v_i| \Rightarrow \sum_i (1-\lambda_i)|v_i\rangle\langle v_i| = \mathbb{1} - \rho$

$$\Rightarrow \langle\psi|\mathbb{1}-\rho|\psi\rangle = \sum_{i=1}^{d} \underbrace{(1-\lambda_i)}_{\geq 0} \underbrace{|\langle\psi|v_i\rangle|^2}_{\geq 0} \geq 0 \quad \forall |\psi\rangle.$$

- $\rho \geq 0 \Rightarrow \lambda_i \geq 0$
- $\mathrm{tr}(\rho) = 1 \Rightarrow \lambda_i \leq 1$

c) Show that for $p(x) = 2^{-n}$ the expected success probability is bounded by $2^{-n}d$. (*Hint:* Use that $\mathbb{1} \geq \rho_i$ for all i and show that for $A \geq 0$ and $B \geq C$ it holds that $\mathrm{Tr}(AB) \geq \mathrm{Tr}(AC)$ as a starting point.)

LEMMA

$A \geq 0,\ B - C \geq 0 \Rightarrow \mathrm{tr}(A(B-C)) \geq 0$

PROOF: $A \geq 0 \Rightarrow A \underset{\text{EIGEN.}}{=} \sum_i \lambda_i |v_i\rangle\langle v_i|$ with $\lambda_i \geq 0 \Rightarrow \mathrm{tr}(A(B-C)) = \sum_i \underbrace{\lambda_i}_{\geq 0} \underbrace{\langle v_i|B-C|v_i\rangle}_{\geq 0 \Leftarrow (B-C)\geq 0} \geq 0$

$$P_{succ} = \frac{1}{2^n} \sum_{x \in \{0,1\}^n} tr(\rho_x F_x) \underset{\substack{\uparrow \\ \rho_x \leq \mathbb{1}}}{\leq} \frac{1}{2^n} \sum_{x \in \{0,1\}^n} tr(\mathbb{1} F_x) = \frac{1}{2^n} tr\Big(\underbrace{\sum_x F_x}_{= \mathbb{1}}\Big) = \frac{1}{2^n} tr(\mathbb{1})$$

$$= \frac{d}{2^n}$$

$$\Rightarrow P_{succ} \leq \frac{d}{2^n}$$

d) What does this imply regarding our motivating question?

- If $\underset{\substack{\uparrow \\ \text{\# of classical bits}}}{n} > \log_2(d)$ $\left(\text{i.e. } \frac{d}{2^n} < 1\right)$, then P_{succ} cannot be 1.
- So we should require that $n \leq \log_2(d)$ in order to have probability 1 of success.

Freie Universität Berlin
Tutorials on Quantum Information Theory
Winter term 2022/23

Problem Sheet 2
POVMs and encoding classical information

J. Eisert, A. Townsend-Teague, A. Mele, A. Burchards, J. Denzler

1. **Non-uniqueness of the decomposition of mixed states.** (4 Points: 2+2)

   Consider two macroscopically different preparation schemes of a large number of polarised photons:

   Preparation A. For each photon we toss a fair coin. Depending on whether we get head or tail, we prepare the photon to have either vertical or horizontal *linear* polarisation.

   Preparation B. For each photon we toss a fair coin. Depending on whether we get head or tail, we prepare the photon to have either left-handed or right-handed *circular* polarisation.

   Note: You can simply think of the polarization of the light as a binary variable and of the polarization axis as a local basis. I.e. the vertical and horizontal linear polarizations may be identified with the $|0\rangle$ and $|1\rangle$ eigen states of the Z basis. Likewise you may interpret the left- and right handed circular polarizations as the $|+\rangle$ and $|-\rangle$ eigen states of the X basis.

   Now we are given a large number of photons which all were prepared by the same scheme.

   a) Argue that having only access to the photons we can not distinguish which of the preparation schemes was used.
   b) Argue that if it were possible to distinguish such types of preparations by measuring the photon, locality would be violated.
      (*Hint*: think about how the state we consider can be prepared by ignoring one degree of freedom of a bipartite system as in the last exercise of Sheet 0.)

2. **Impossible machines – no cloning.** (5 Points)

   In this problem we will re-derive the impossibility results that you have seen in the lecture but now directly using the structure of quantum theory.

   Show that there does not exist a unitary map on two copies of a Hilbert space $\mathcal{H}$ which acts in the following way:

   $$\forall \,|\psi\rangle \in \mathcal{H} : \, U\,|\psi\rangle\,|0\rangle = \mathrm{e}^{\mathrm{i}\phi(\psi)}\,|\psi\rangle\,|\psi\rangle \, .$$

   (*Hint*: Unitary operators are linear.)

3. **The most general quantum measurements.** (4 Points: 2+1+1)

   In a quantum mechanics course, measurements are typically introduced as projective measurements of the eigenvalues of observables. But from a theoretical perspective another measurement description is often helpful. For simplicity—and in the spirit of information theory—we assume that the possible measurement outcomes are from a discrete set $\mathcal{X}$. [1]

[1] More generally, one can replace $\mathcal{X}$ by the σ-algebra of a measurable Borel space. This is the natural structure from probability theory to describe a set of all possible events in an experiment.

A measurement with outcomes $\mathcal{X}$ on a quantum system with Hilbert space $\mathcal{H}$ can be described by a *positive operator valued measure* (POVM) on $\mathcal{X}$. We denote by $\mathrm{Pos}(\mathcal{H}) := \{A \in L(\mathcal{H}) \mid A \geq 0\}$ the set of Hermitian positive semi-definite operators on $\mathcal{H}$. A POVM on a discrete space $\mathcal{X}$ is a map $\mu : \mathcal{X} \to \mathrm{Pos}(\mathcal{H})$ such that $\sum_{x\in\mathcal{X}} \mu(x) = \mathrm{Id}$. If the system is in the quantum state $\rho \in \mathcal{D}(\mathcal{H})$, the probability of observing the outcome $x \in \mathcal{X}$ is given by $\mathrm{Tr}(\mu(x)\rho)$.

a) What is the difference between POVM measurements and the measurement description using observables? (Here we refer to the measurement description using observables as the measurement process where the quantum state gets projected on the projector valued measure (PVM) corresponding to the spectral value that is measured during the measurement process as explained in the lecture).

It is often stated that this is the most general form of a quantum measurement. We want to understand this statement in more detail. So what could be regarded as the most general quantum measurement? One can start as follows: A (general) quantum measurement M with outcomes in $\mathcal{X}$ is a map that associates to each quantum state $\rho \in \mathcal{D}(\mathcal{H})$ a probability measure p_ρ on $\mathcal{X}$, i.e. $M : \rho \mapsto p_\rho$ with $p_\rho : \mathcal{X} \to [0, 1]$ such that $\sum_{x\in\mathcal{X}} p_\rho(x) = 1$.

b) Show that any POVM on $\mathcal{X}$ defines a general quantum measurements as defined above.

c) Show that any general quantum measurements as defined above defines a unique POVM on $\mathcal{X}$.
(*Hint*: You may assume a general measurement M to be linear. Then interpret.)

4. **Encoding classical bits.** (7 Points: 2+2+2+1)

In the last exercise we introduced the description of quantum measurements with the help of POVMs. Now, let $\mathcal{H}$ be a d-dimensional Hilbert space. Our aim is to encode n classical bits into the space of quantum states $\mathcal{D}(\mathcal{H})$. To this end, we choose a set of 2^n states $\{\rho_i\}_{i\in\{0,1\}^n} \subset \mathcal{D}(\mathcal{H})$, each state corresponding to a bit string. To decode the bit string we have to make a measurement described by a POVM $F = \{F_i\}_{i\in\{0,1\}^n}$, where the bit string is the outcome. In this exercise we are going to investigate the following question:

How many classical bits can be encoded and (perfectly) decoded in a d-dimensional quantum system in this way?

Consider a source that outputs the bit string $x \in \{0, 1\}^n$ with probability $p(x)$.

a) We say that the decoder is successful if outcome i is returned upon measuring F on ρ_i. Define the expected success probability of the decoder with respect to the distribution p.

b) Prove the technical Lemma that $\rho \leq \mathbb{1}$ for ρ a density matrix.

c) Show that for $p(x) = 2^{-n}$ the expected success probability is bounded by $2^{-n}d$.
(*Hint:* Use that $\mathbb{1} \geq \rho_i$ for all i and show that for $A \geq 0$ and $B \geq C$ it holds that $\mathrm{Tr}(AB) \geq \mathrm{Tr}(AC)$ as a starting point.)

d) What does this imply regarding our motivating question?