Sequential Quantum Measurements

Master's Thesis
University of Turku
Theoretical Physics
2018
B.Sc. Tuomas Tammero
Supervisors:
Docent Teiko Heinosaari
Prof. Sabrina Maniscalco

UNIVERSITY OF TURKU
Department of Physics and Astronomy

TAMMERO, TUOMAS: Sequential Quantum Measurements

Master's thesis, 48 pages
Theoretical physics
December 2018

This thesis studies the field of quantum measurements and more precisely sequential measurements. The aim is to provide answers to questions like when is performing a sequential or a repeated measurement meaningful, or even possible. This is important in order to understand what type of information can be acquired when performing measurements on quantum states.

In basic quantum theory observables are identified with self-adjoint operators. We introduce the concept of a positive operator valued measure (POVM) to gain better understanding on observables. The downside to POVMs is that they don't produce a new state after the measurement, only measurement statistics. In order to perform multiple measurements on a quantum state, the concepts of measurement model and instrument are defined in this thesis. Using instruments to perform measurements is essential in order to describe a state after the initial measurement is performed. This way it's possible to make measurements sequentially.

After defining the necessary tools for sequential measurements and a few different properties related to them, some applicable situations for these kind of measurements are shown. This way sequential measurements end up being somewhat meaningful in order to gain information from quantum systems.

Keywords: Measurement, Sequential Measurement, Observable, Measurement Model, Instrument

TURUN YLIOPISTO
Fysiikan ja tähtitieteen laitos

TAMMERO, TUOMAS: Kvanttijonomittaukset

Pro gradu -tutkielma, 48 sivua
Teoreettinen fysiikka
Joulukuu 2018

Tämän tutkielman tarkoituksena on tarkastella kvanttimittauksia ja vielä tarkemmin sanottuna kvanttijonomittauksia. Tavoitteena on vastata muutamiin kysymyksiin, esimerkiksi milloin jonomittauksen tai toistetun mittauksen suorittaminen on mielekästä tai ylipäätään mahdollista. Mittausten mielekkyyden ja mahdollisuuden tärkeys nousee esille, kun halutaan tietää minkälaista informaatiota on mahdollista saada suorittamalla mittauksia kvanttitiloille.

Kvanttiteorian alkeissa suureet yhdistetään itseadjungoituihin operaattoreihin. Tässä tutkielmassa esitellään positiivioperaattorimitan käsite, jotta suureita voisi ymmärtää syvällisemmin. Tosin mittauksen suorittaminen positiivioperaattorimittaa käyttäen ei riitä silloin kun halutaan mallintaa kvanttitilaa mittauksen jälkeen, vaan jäljelle jää ainoastaan statistiikkaa itse mittauksesta. Jotta kvanttitilaan voisi suorittaa useita mittauksia, tarvitaan mittausmallin ja instrumentin käsitteitä. Instrumenttien käyttäminen on ensisijaisen tärkeää, kun tilaan halutaan suorittaa uusia mittauksia ensimmäisen mittauksen jälkeen. Tällä tavoin on mahdollista suorittaa jonomittauksia.

Sen jälkeen kun tarvittavat työkalut jonomittausten suorittamiseen on määritelty, esitellään muutamia näihin liittyviä ominaisuuksia. Käyttämällä määriteltyjä työkaluja ja ominaisuuksia on mahdollista esitellä sovellutuksia jonomittauksille. Tällä tavoin jonomittaukset osoittautuvat hyödyllisiksi, kun kvanttitilasta pyritään saamaan mahdollisimman paljon informaatiota.

Avainsanat: Mittaus, Jonomittaus, Suure, Mittausmalli, Instrumentti

Contents

Introduction **1**

1 Quantum Mechanical Framework **2**

1.1 Hilbert spaces 2

1.2 Hilbert Space Operators 3

1.3 Observables 8

1.4 Dilation Theorems 13

1.5 Measurement Models 15

2 Formulation of Sequential Measurements **18**

2.1 Instruments 18

2.2 Sequential Measurements 21

2.3 Repeatable Measurements 23

3 Properties of Quantum Measurements **24**

3.1 Compatibility 24

3.2 Non-Disturbance 26

3.3 Joint Measurability 27

3.4 Universality 29

3.5 Broadcasting 33

4 Applications of Sequential Measurements **36**

4.1 Saturation of Repeated Measurements 37

4.2 Approximate Repeatability 44

References **48**

Introduction

Measurements in the quantum world are extremely different from what they are in the classical sense. Classically, it's possible to measure almost anything with precision, given the limitations of our measurement devices. As an example a classical system containing only a single particle is fully described by it's location and momentum, both of which can be measured. In quantum theory, states are used to describe the system and a measurement is performed directly on the state. These measurements performed on states often disrupt the state in one way or another. We might not be able to perform another measurement to a state after an initial measurement is made or the measurement itself could alter the state somehow.

The objective of this thesis is to showcase and study a variety of measurement setups, which affect the measured state in different ways. Starting from the rudiments of quantum mechanical framework and the most basic notion of an observable, a positive operator valued measure, we work our way to measurement models and instruments. The aim is to introduce the necessary tools needed to perform multiple measurements in the initial state, which is essentially what sequential measurements mean. A special case of sequential measurement is repeatable measurements, where a measurement of a single observable is performed multiple times. We study how measuring observables sequentially alter the quantum state in each step and when performing these types of measurements is meaningful.

Quantum measurements have a variety of features which are introduced in this thesis. For example the notion of non-disturbance, which shows what kind of measurements can be performed without disturbing the initial state. Another feature closely related to non-disturbance worth mentioning is joint measurability, where we'll study what kind of observables can be measured jointly on the quantum state.

In the final chapter we'll go through some applicable scenarios where a sequential measurement can be performed, including the canonical position observable. The point is to show that after thoroughly showcasing quantum measurements and the related properties, the studied measurement setups are in fact useful when studying real quantum systems.

1 Quantum Mechanical Framework

In this chapter, all the important bits and pieces of quantum theory needed to understand this work are presented with examples. The source of these important definitions is the book[10]. While the book is extremely comprehensive, only the most important parts needed to present quantum measurements and more precisely, sequential measurements, are reviewed here.

1.1 Hilbert spaces

The most fundamental mathematical concept in quantum theory is that of Hilbert space. While the definition itself in not too complex, it's important to recall what is meant by Hilbert space, since they are needed throughout this thesis. We'll start off by defining an inner product in a complex vector space.

Definition 1.1. Let $\mathcal{H}$ be a complex vector space. The function $\langle\cdot|\cdot\rangle$ on $\mathcal{H}\times\mathcal{H}$ is called an inner product if,

i) $\langle\varphi|c\psi+\phi\rangle = c\,\langle\varphi|\psi\rangle + \langle\varphi|\phi\rangle\,, \quad \forall c\in\mathbb{C}, \quad \varphi,\psi,\phi\in\mathcal{H}$

ii) $\overline{\langle\varphi|\psi\rangle} = \langle\psi|\varphi\rangle$

iii) $\langle\varphi|\varphi\rangle > 0, \quad \text{if } \varphi\neq 0$

Inner product defines a norm $\|\varphi\| := \sqrt{\langle\varphi|\varphi\rangle}$ in $\mathcal{H}$, making $\mathcal{H}$ a normed space. Moreover, a metric $d(\cdot,\cdot)$ can be defined using this norm, $d(\varphi,\psi) = \|\varphi-\psi\|$, which leads to the definition of a complete metric space[10].

Definition 1.2. Let $\{\varphi_i\}$ be a sequence in $\mathcal{H}$. The sequence is called a Cauchy-sequence, if for every $\varepsilon > 0$, there exists $N_\varepsilon \in \mathbb{N}, N_\varepsilon > 0$ such that $d(\varphi_i,\varphi_j) < \varepsilon$ when $i,j > N_\epsilon$. A metric space $\mathcal{H}$ is said to be complete, if every Cauchy-sequence in $\mathcal{H}$ converges.

Example 1.1.

Let $\varphi_i = \frac{1}{2^i}$ and $\varepsilon > 0$. Then choose $N_\varepsilon > 0$ so that $\frac{1}{2^{N_\varepsilon}} > \varepsilon$, $\frac{2}{2^{N_\varepsilon}} = \varepsilon$ and $i, j > N_\varepsilon$ which leads to

$$\begin{aligned} d(\varphi_i, \varphi_j) &= \left\| \frac{1}{2^i} - \frac{1}{2^j} \right\| \leq \frac{1}{2^i} + \frac{1}{2^j} \\ &\leq \frac{1}{2^{N_\varepsilon}} + \frac{1}{2^{N_\varepsilon}} = \varepsilon, \end{aligned}$$

which proves that $\{\varphi_i\}$ is Cauchy.

To show an example of a sequence, which is not Cauchy, let $\varphi_i = \ln i$. From this we get

$$\begin{aligned} d(\varphi_{i+1}, \varphi_i) &= \|\ln(i+1) - \ln i\| = \ln(1 + \frac{1}{i}) \\ &\to \ln 1 = 0, \end{aligned}$$

from which it's clear that the sequence is not Cauchy.

Combining the already defined inner product and completeness leads to the following definition.

Definition 1.3. A inner product space, which is complete, is called a Hilbert space.

From now on $\mathcal{H}$ always denotes a complex and separable Hilbert space, meaning a Hilbert space with a countable orthonormal basis. Though there are a number of interesting Hilbert spaces to consider, we'll mostly focus on the two-dimensional complex space, $\mathcal{H} = \mathbb{C}^2$.

1.2 Hilbert Space Operators

There are multiple important sets of operators with a variety of different attributes, all necessary to define for their own purposes. But first, we need the very basic definition of a bounded operator.

Definition 1.4. Let $T : \mathcal{H} \to \mathcal{H}$ be a linear mapping. T is a bounded operator if there is a $t \geq 0$, for which

$$\|T\psi\| \leq t \|\psi\|, \quad \forall \psi \in \mathcal{H}.$$

The set of bounded operators is denoted with $\mathcal{L}(\mathcal{H})$. Moreover, each bounded operator T has an adjoint operator T^* defined by

$$\langle \varphi | T^* \psi \rangle = \langle T\varphi | \psi \rangle \,.$$

Using the adjoint operator, the absolute value of operator can also be defined as

$$|T| = (T^*T)^{1/2}.$$

The following sets of bounded linear operators are used throughout this thesis and form an integral part of quantum theory.

- The set of selfadjoint operators, $\mathcal{L}_s(\mathcal{H})$. A linear operator L if selfadjoint, if $L = L^*$.
- The set of projections, $\mathcal{P}(\mathcal{H})$. A selfadjoint operator P is a projection, if $P^2 = P$.
- The set of unitary operators, $\mathcal{U}(\mathcal{H})$. A bounded operator U is unitary if $U^*U = UU^* = \mathbb{1}$.
- The set of trace class operators, $\mathcal{T}(\mathcal{H})$. A bounded operator T is trace class if $\mathrm{tr}[|T|] < \infty$.
- The set of states, $\mathcal{S}(\mathcal{H})$. A trace class operator ρ is a state, if it's positive and of trace one. The state ρ is pure, if it's also a one-dimensional projection.
- The set of effects, $\mathcal{E}(\mathcal{H})$. A selfadjoint operator E is an effect, if $O \leq E \leq \mathbb{1}$.

Example 1.2. We'll show some examples of these kind of operators in the Hilbert space $\mathbb{C}^2$.

- An example of a selfadjoint operator in $\mathbb{C}^2$ would be any real valued 2x2 matrix.
- An example of a projection in $\mathbb{C}^2$ is the matrix $P = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$, since clearly $P^2 = P$.

- An example of a unitary operator in $\mathbb{C}^2$ is a diagonal matrix
$$U = \begin{pmatrix} ie^{-ix} & 0 \\ 0 & ie^{-iy} \end{pmatrix}, \quad \text{where } x, y \in \mathbb{R}.$$
- An example of a trace class operator in $\mathbb{C}^2$ is any real or complex valued 2x2 matrix.

When working with sequences of operators and their convergence, the convergence is referred as weak or strong. This means that the convergence happens with respect to either the weak or strong operator topology.

Definition 1.5. Let $\{T_i\}$ be a sequence of bounded operators, $T_i \in \mathcal{L}(\mathcal{H})$ for every index i.

a) The sequence converges strongly to an operator $T \in \mathcal{L}(\mathcal{H})$ if $\lim T_i\varphi = T\varphi$, $\forall \varphi \in \mathcal{H}$.

b) The sequence converges weakly to an operator $T \in \mathcal{L}(\mathcal{H})$ if $\lim \langle \varphi | T_i \psi \rangle = \langle \varphi | T\psi \rangle$, $\forall \varphi, \psi \in \mathcal{H}$.

Throughout this thesis convergence is always referred to be either strong or weak and by the above definition, the convergence can be thought to happen with respect to either the norm or inner product.

Next up we will prove two closely related theorems, the Hilbert projection theorem and Riezs's lemma, both of which will later be needed when proving the Kolmogorov extension theorem for POVMs.

Theorem 1.1 (The Hilbert Projection Theorem)**.** Let $\mathcal{H}$ be a Hilbert space with a closed subspace $\mathcal{N}$. Every $\psi \in \mathcal{H}$ can bee written as $\psi = \phi + \zeta$ where $\phi \in \mathcal{N}$ and $\zeta \in \mathcal{N}^{\perp}$.

Proof First we need to show that for $\psi \in \mathcal{H}$ there exists a unique element $\phi \in \mathcal{N}$, which is closest to ψ. Let $d = \inf_{\varphi \in \mathcal{N}} \|\psi - \varphi\|$ and let $\{\varphi_n\}$ be a sequence in $\mathcal{N}$ for which

$$\|\psi - \varphi_n\| \to d.$$

It follows that

$$\begin{aligned}
\|\varphi_n - \varphi_m\|^2 &= \|(\varphi_n) - (\varphi_m)\|^2 \\
&= 2\|\varphi_n - \psi\|^2 + 2\|\varphi_m - \psi\|^2 - \|-2\psi + \varphi_n + \varphi_m\|^2 \\
&= 2\|\varphi_n - \psi\|^2 + 2\|\varphi_m - \psi\|^2 - 4\left\|\psi - \frac{1}{2}(\varphi_n + \varphi_m)\right\|^2 \\
&\leq 2\|\varphi_n - \psi\|^2 + 2\|\varphi_m - \psi\|^2 - 4d^2 \\
&\to 2d^2 + 2d^2 - 4d^2 = 0,
\end{aligned}$$

when $n \to \infty$ and $m \to \infty$. We can see that $\{\varphi_n\}$ is a Cauchy sequence and it converges to an element $\phi \in \mathcal{N}$ and $\|\psi - \phi\| = d$. The element ϕ is also unique, since if there was another element ϕ' for which $\|\psi - \phi'\| = d$

$$\begin{aligned}
&\|\psi - \phi\| = \|\psi - \phi'\| \\
\Leftrightarrow &\|\psi - \phi\|^2 = \|\psi - \phi'\|^2 \\
\Leftrightarrow &\langle \psi - \phi | \psi - \phi \rangle - \langle \psi - \phi' | \psi - \phi' \rangle = 0 \\
\Leftrightarrow &\langle \phi' - \phi | \phi' - \phi \rangle = 0 \\
\Leftrightarrow &\phi' = \phi.
\end{aligned}$$

Now with ϕ being the element closest to ψ we can define $\zeta = \psi - \phi \Leftrightarrow \psi = \phi + \zeta$. Let $\varphi \in \mathcal{N}$ and $\alpha \in \mathbb{R}$. For $d = \|\psi - \phi\|$ we have

$$\begin{aligned}
d^2 \leq \|\psi - (\phi + \alpha\varphi)\|^2 &= \|\zeta - \alpha\varphi\|^2 \\
&= d^2 - 2\alpha \cdot \mathrm{Re}(\langle \zeta | \varphi \rangle) + \alpha^2 \|\varphi\|^2 \\
\Rightarrow &- 2\alpha \cdot \mathrm{Re}(\langle \zeta | \varphi \rangle) + \alpha^2 \|\varphi\|^2 \geq 0,
\end{aligned}$$

from which we get $\mathrm{Re}(\langle \zeta | \varphi \rangle) = 0$. Similarly, by using $i\alpha \in \mathbb{C}$ instead of the real scalar we get $\mathrm{Im}(\langle \zeta | \varphi \rangle) = 0$ and it follows that $\zeta \in \mathcal{N}^\perp$. Furthermore ζ is unique, since if there were another ζ' for which $\psi = \phi + \zeta = \phi + \zeta'$, we get

$$\begin{aligned}
&\|\zeta - \alpha\varphi\|^2 = \|\zeta' - \alpha\varphi\|^2 \\
\Leftrightarrow &\langle \zeta - \alpha\varphi | \zeta - \alpha\varphi \rangle = \langle \zeta' - \alpha\varphi | \zeta' - \alpha\varphi \rangle \\
\Leftrightarrow &\langle \zeta' - \zeta | \zeta' - \zeta \rangle = 0 \\
\Leftrightarrow &\zeta' = \zeta.
\end{aligned}$$

Theorem 1.2 (Riesz's lemma)**.** For every functional $T : \mathcal{H} \to \mathbb{C}$ there exists a unique $\psi_T \in \mathcal{H}$ such that

$$T\varphi = \langle \psi_T | \varphi \rangle \quad \forall \varphi \in \mathcal{H},$$
$$\|\psi_T\|_{\mathcal{H}} = \|T\|_{\mathcal{H}^*} .$$

Proof Let $\mathcal{N}$ be the set of vectors $\varphi \in \mathcal{H}$ which $T\varphi = 0$. The continuity of T guarantees that the subspace $\mathcal{N}$ of $\mathcal{H}$ is closed. If $\mathcal{N} = \mathcal{H}$ then it would follow that $T\varphi = \langle 0, \varphi \rangle = 0, \forall \varphi \in \mathcal{H}$, which completes the proof for $\mathcal{N} = \mathcal{H}$. Assume that $\mathcal{N} \neq \mathcal{H}$, then by the projection theorem[14] there exists a nonzero vector $\varphi_0 \in \mathcal{N}^\perp$ and we define

$$\psi_T = T^* \varphi_0 \|\varphi_0\|^{-2} \varphi_0,$$

which we will show to have the correct properties. Firstly, for $\varphi \in \mathcal{N}$ it's clear that $T\varphi = 0 = \langle \psi_T | \varphi \rangle$. If $\varphi = \alpha \varphi_0$ for some scalar α then

$$T\varphi = T\alpha\varphi_0 = \alpha T\varphi_0 = \langle T^* \varphi_0 \|\varphi_0\|^{-2} \varphi_0 | \alpha\varphi_0 \rangle = \langle \psi_| \alpha\varphi_0 \rangle .$$

From the linearity of functions T and $\langle \psi_T | \cdot \rangle$ and the fact that both agree on $\mathcal{N}$ and φ_0, it follows that they must also agree on the space, which is spanned by $\mathcal{N}$ and φ_0. The vector space spanned by $\mathcal{N}$ and φ_0 is $\mathcal{H}$, since every vector $\psi \in \mathcal{H}$ can be written as

$$\psi = \left(\psi - \frac{T\psi}{T\varphi_0} \varphi_0 \right) + \frac{T\psi}{T\varphi_0} \varphi_0$$

and so, $T\varphi = \langle \psi_T, \varphi \rangle , \forall \varphi \in \mathcal{H}$. If there exists another vector $\psi' \in \mathcal{H}$ for which $T\varphi = \langle \psi' | \varphi \rangle$, then it follows that

$$\|\psi' - \psi_T\|^2 = T(\psi' - \psi_T) - T(\psi' - \psi_T) = 0$$
$$\Rightarrow \psi' = \psi_T,$$

which proves the uniqueness of ψ_T. Furthermore, the fact that $\|T\|_{\mathcal{L}(\mathcal{H})} = \|\psi_T\|_{\mathcal{H}}$ holds follows from

$$\|T\| = \sup_{\|\varphi\| \leq 1} |T\varphi| = \sup_{\|\varphi\| \leq 1} |\langle \psi_T, \varphi \rangle| \leq \sup_{\|\varphi\| \leq 1} \|\psi_T\| \|\varphi\| = \|\psi_T\|$$

and

$$\|T\| = \sup_{\|\varphi\|\leq 1} |T\varphi| \geq \left| T\left(\frac{y_T}{\|y_T\|}\right)\right| = \langle \psi_T | \psi_T / \|\psi_T\| \rangle = \|\psi_T\| \,,$$

which completes the proof [14]. □

Before introducing observables, we need to define a key feature for them, which is their spectral decomposition. Spectral decomposition often makes handling the math around observables much easier.

Theorem 1.3. For a trace class operator T, there exists an orthonormal basis $\{\varphi_i\}$ and a sequence of complex numbers $\{\lambda_i\}$ for which

$$T = \sum_i \lambda_i \left|\varphi_i\right\rangle \left\langle\varphi_i\right| .$$

For operators not in the trace class, a spectral decomposition exists, but the decomposition itself might have to be written as an integral instead of a sum.

1.3 Observables

Before any measurements can be performed, observables need to be defined. In introduction level quantum physics, observables are often presented as selfadjoint operators. For the purposes of this thesis, this representation of observables is not sufficient. Here observables are identified as positive operator valued measures (POVM's), which are mappings from a chosen σ-algebra $\mathcal{F}$, to the set of effects in the Hilbert space, $\mathcal{E}(\mathcal{H})$.

Definition 1.6. A σ-algebra $\mathcal{F}$ on a nonempty set Ω is a collection of subsets, which has the following three properties

i) $\emptyset \in \mathcal{F}, \quad \Omega \in \mathcal{F}$

ii) $X \in \mathcal{F} \Rightarrow \Omega \setminus X \in \mathcal{F}$

iii) $X_1, X_2, ... \in \mathcal{F} \Rightarrow \cup_i X_i \in \mathcal{F}$

A measurable space is defined by the pair $(\Omega, \mathcal{F})$. Any set on the σ-algebra is called an event, which makes the σ-algebra a collection of events. By denoting $\mathfrak{B}(\Omega)$ we always denote the Borel σ-algebra of the set Ω, which is defined as the smallest σ-algebra containing all the open sets of Ω[16].

Example 1.3. Let's consider the set $\Omega = \{0, 1, 2\}$. The following collections of subsets are valid σ-algebras for the predefined set

- $\{\emptyset, \Omega\}$
- $\{\emptyset, \Omega, 0, 1, 2, \{0, 1\}, \{0, 2\}, \{1, 2\}\}$

With the σ-algebra defined, it's time to introduce the concept of a positive operator valued measure.

Definition 1.7. A POVM is a mapping $\mathsf{A} : \mathcal{F} \to \mathcal{E}(\mathcal{H})$, for which

i) $\mathsf{A}(\emptyset) = 0$

ii) $\mathsf{A}(\Omega) = \mathbb{1}$

iii) $\mathsf{A}(\cup_i X_i) = \sum_i \mathsf{A}(X_i)$, with respect to the weak operator topology, for any sequence of disjoint sets $\{X_i\} \in \mathcal{F}$

Moreover, the mapping A is a POVM if and only if $X \mapsto \mathrm{tr}[\rho A(X)]$ represents a probability measure for all states $\rho \in \mathcal{S}(\mathcal{H})$.

Example 1.4. Let's consider the Hilbert space $\mathcal{H} = \mathbb{C}^2$. Now, any selfadjoint operator A can be written using the operators $\mathbb{1}, \sigma_x, \sigma_y$ and σ_z by composing a linear combination. This is a way to obtain a general effect on $\mathcal{H}$,

$$A = \frac{1}{2}(\alpha \mathbb{1} + \vec{r} \cdot \vec{\sigma}), \quad \alpha \in \mathbb{R}, \vec{r} \in \mathbb{R}^3.$$

The eigenvalues of A are $\lambda_\pm = \frac{1}{2}(\alpha \pm \|\vec{r}\|)$. A is in fact an effect if the following requirements are held

$$0 \leq \lambda_-, \quad \lambda_+ \leq 1 \Rightarrow \|\vec{r}\| \leq \alpha \leq 2 - \|\vec{r}\| .$$

This is the general definition of an observable. When an observable is said to be discrete, then A is a function, the outcome set Ω is finite and the sigma-algebra of the observable is just the power set of the outcome set Ω. We use the notation $\mathsf{A}(X)$ when working with the general case and distinguish discrete observables using lowercase notation for the outcome, $\mathsf{A}(x)$.

An important subset of POVMs are projection-valued measures, PVMs. PVMs are identified as sharp observables and have a little bit stricter definition than POVMs.

Definition 1.8. An observable A is sharp, if $\mathsf{A}(X)$ is a projection for every $X \in \mathcal{F}$.

Example 1.5. Let $\mathcal{H} = \mathbb{C}^2$. In $\mathcal{H}$ sharp observables have only two possible outcomes, labelled here as ± 1. Respectively, the corresponding operators $\mathsf{A}(-1)$ and $\mathsf{A}(1)$ are either O, $\mathbb{1}$ or some other one dimensional projection. If we exclude O and $\mathbb{1}$, choose a vector $\vec{r} \in \mathbb{R}^3$, we can define

$$\mathsf{A}(-1) = \frac{1}{2}(\mathbb{1} - \vec{r} \cdot \vec{\sigma}), \quad \mathsf{A}(1) = \frac{1}{2}(\mathbb{1} + \vec{r} \cdot \vec{\sigma}), \tag{1}$$

where $\vec{\sigma}$ denotes a vector with the Pauli matrices as elements. The corresponding selfadjoint operator for the observable A is

$$A = \sum_i i\mathsf{A}(i) = \vec{r} \cdot \vec{\sigma}.$$

Observables can be characterized in multiple ways. One of these characterizations is informationally complete observables, which play an important role when discussing measurement features like non-disturbance and universality.

Definition 1.9. A collection of observables $\{\mathsf{A}, \mathsf{B}, \ldots\}$ is informationally complete, if for every pair of states the probability distributions of these observables being equal implies that the states are equal.

Often in literature a single observable A is said to be informationally complete instead of a set of observables. In this case, we require that the probability distribution of this observable being equal to every pair of states implies that the states are equal.

Example 1.6. Let $\mathcal{H} = \mathbb{C}^2$ and let P be a set of $\dim(\mathcal{H}) = 4$ normalized vectors $\{|\psi_1\rangle, |\psi_2\rangle, |\psi_3\rangle, |\psi_4\rangle\}$, for which

$$|\langle\psi_k|\psi_j\rangle| = \frac{1}{3}, \quad k \neq j.$$

The POVM elements corresponding to P are the scaled operators $|\psi_k\rangle\langle\psi_k|/3 = \Pi_k/3$. In order to have an informationally complete P, the operators Π_k must be linearly independent, since they need to span the entire space of operators.[15]

In the qubit case, the four vectors needed to construct these operators are easy find. Let the first of the four vectors be

$$|\psi_1\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

Then we can assume that the first component of each of the other three vectors is $1/\sqrt{3}$ to meet the necessary requirement. The other component similarly needs to be rotated accordingly, so to say, since the first component of the three vectors is same. In fact, the following three vectors can be used to construct the wanted POVM

$$|\psi_2\rangle = \begin{pmatrix} \frac{1}{\sqrt{3}} \\ \sqrt{\frac{2}{3}} \end{pmatrix}, \quad |\psi_2\rangle = \begin{pmatrix} \frac{1}{\sqrt{3}} \\ \sqrt{\frac{2}{3}}e^{2i\pi/3} \end{pmatrix}, \quad |\psi_3\rangle = \begin{pmatrix} \frac{1}{\sqrt{3}} \\ \sqrt{\frac{2}{3}}e^{-2i\pi/3} \end{pmatrix},$$

and the elements of P are $|\psi_1\rangle\langle\psi_1|, |\psi_2\rangle\langle\psi_2|, |\psi_3\rangle\langle\psi_3|$ and $|\psi_4\rangle\langle\psi_4|$.

When it comes to sequential measurements, we need to define a pre-order of observables called post-processing. This gives us a way to compare two observables and the amount of information yielded from each.

Definition 1.10. Let A and B be observables with outcome sets Ω_{A} and Ω_{B}, respectively. Denoting $\mathsf{A} \preceq_{post} \mathsf{B}$ means that there exists a mapping κ, for which

$$\kappa : \Omega_{\mathsf{A}} \times \Omega_{\mathsf{B}} \to [0,1]$$

$$\sum_x \kappa(x,y) = 1, \quad \forall y \in \Omega_{\mathsf{B}},$$

$$\mathsf{A}(x) = \sum_y \kappa(x,y)\mathsf{B}(y), \quad \forall x \in \Omega_{\mathsf{A}}.$$

A mapping κ satisfying these conditions is called a Markov kernel, which form a convex set. Furthermore, denoting $\mathsf{A} \simeq_{post} \mathsf{B}$ requires that both $\mathsf{A} \preceq_{post} \mathsf{B}$ and $\mathsf{B} \preceq_{post} \mathsf{A}$ hold. Denoting $\mathsf{A} \prec_{post} \mathsf{B}$ requires that $\mathsf{A} \preceq_{post} \mathsf{B}$, but not $\mathsf{B} \preceq_{post} \mathsf{A}$.

To show that post-processing really is a pre-order, we need to prove it's reflexivity and transitivity. Reflexivity follows straight from the definition of the mapping κ. For three observables $\mathsf{A}, \mathsf{B}, \mathsf{C}$, let $\mathsf{A} \preceq_{post} \mathsf{B}$ and $\mathsf{B} \preceq_{post} \mathsf{C}$, then

$$\begin{aligned}
\mathsf{A}(x) &= \sum_y \kappa_1(x,y)\mathsf{B}(y), \quad \mathsf{B}(y) = \sum_Z \kappa_2(y,x)\mathsf{C}(z) \\
\Rightarrow \mathsf{A}(x) &= \sum_{y,z} \kappa_1(x,y)\kappa_2(y,z)\mathsf{C}(z)
\end{aligned}$$

Since both $\sum_y \kappa_1(x,y) = 1$ and $\sum_Z \kappa_2(y,z) = 1$ and both κ_1 and κ_2 are mappings to the set $[0,1]$, post-processing is transitive and hence, a pre-order.

Example 1.7. Let A be a qubit observables of the form

$$\mathsf{A}(\pm 1) = \frac{1}{2}(\mathbb{1} \pm \vec{a} \cdot \vec{\sigma}).$$

To obtain another qubit observable B, we need the applied post-processing κ as

$$\begin{aligned}
\kappa(+,+) &= \kappa(-,-) = s \\
\kappa(+,-) &= \kappa(-,+) = 1 - s, \qquad \text{where } \ s \in \mathbb{C}.
\end{aligned}$$

Using the above definition of post processing we get

$$\begin{aligned}
\mathsf{B}(+1) &= \kappa(+,+)\mathsf{A}(+1) + \kappa(+,-)\mathsf{A}(-1) \\
&= \frac{1}{2}(s(\mathbb{1} + \vec{a} \cdot \vec{\sigma}) + (1-s)(\mathbb{1} - \vec{a} \cdot \vec{\sigma})) \\
&= \frac{1}{2}(\mathbb{1} + (2s-1)\vec{a} \cdot \vec{\sigma}), \\
\mathsf{B}(-1) &= \frac{1}{2}(\mathbb{1} - (2s-1)\vec{a} \cdot \vec{\sigma}) \\
\Rightarrow \mathsf{B}(\pm 1) &= \frac{1}{2}(\mathbb{1} \pm (2s-1)\vec{a} \cdot \vec{\sigma}).
\end{aligned}$$

So the chosen post-processing applied on the observable A yields an observable B, which is almost the same as A, but with the added noise factor $(2s-1)$.

We defined post-processing for discrete observables, since this is what we'll use later in this thesis. In the definition of post processing for a general observable, the sum is replaced with an integral and $\kappa(\cdot, Y)$ is a probability measure and $\kappa(X, \cdot)$ is $\mathcal{F}_{\mathsf{A}}$-measurable for every $X \in \mathcal{F}_{\mathsf{A}}$.

1.4 Dilation Theorems

In this section we'll present two important and well-known dilation theorems, which are going to be very useful later in this thesis. The source of the proofs for these theorems is the book by Paulsen[13].

Theorem 1.4. (Stinespring's dilation theorem) Let $\mathcal{E}^* : \mathcal{L}(\mathcal{H}) \to \mathcal{L}(\mathcal{H})$ be a completely positive linear map. There exists another Hilbert space $\mathcal{K}$, a unital *-homomorphism $\alpha : \mathcal{L}(\mathcal{H}) \to \mathcal{L}(\mathcal{K})$ and a bounded operator $V : \mathcal{H} \to \mathcal{K}$ such that

$$\mathcal{E}^*(T) = V^*\alpha(T)V, \quad \forall T \in \mathcal{L}(\mathcal{H}). \tag{2}$$

In addition if $V^*V = \mathbb{1}$ then $\mathcal{E}^*$ is unital.

Proof Let's define a symmetric bilinear function $\langle\cdot|\cdot\rangle$ in the tensor product $\mathcal{L}(\mathcal{H})\otimes\mathcal{H}$ by

$$\langle T \otimes \varphi | L \otimes \psi\rangle = \langle \mathcal{E}^*(L^*T)\varphi|\psi\rangle_{\mathcal{H}}$$

where $\langle|\rangle_{\mathcal{H}}$ is the inner product in $\mathcal{H}$.

The function $\langle|\rangle$ is a positive semidefinite, which follows from the complete positivity of $\mathcal{E}^*$, since

$$\left\langle \sum_{i=1}^{n} T_i \otimes \varphi_i | \sum_{j=1}^{n} T_j \otimes \varphi_j \right\rangle = \left\langle \mathcal{E}_n^*(T_j^*T_i) \begin{pmatrix} \varphi_1 \\ \vdots \\ \varphi_n \end{pmatrix} | \begin{pmatrix} \varphi_1 \\ \vdots \\ \varphi_n \end{pmatrix} \right\rangle_{\mathcal{H}^{(n)}} \geq 0,$$

where $\langle|\rangle_{\mathcal{H}^{(n)}}$ is the inner product in $\mathcal{H}^{(n)}$, the direct sum containing the Hilbert space $\mathcal{H}$ n times. This inner product is defined by

$$\left\langle \begin{pmatrix} \varphi_1 \\ \vdots \\ \varphi_n \end{pmatrix} \middle| \begin{pmatrix} \psi_1 \\ \vdots \\ \psi_n \end{pmatrix} \right\rangle_{\mathcal{H}^{(n)}} = \langle\varphi_1|\psi_1\rangle_{\mathcal{H}} + \cdots \langle\varphi_n|\psi_n\rangle_{\mathcal{H}}\,.$$

Since every positive semidefinite bilinear function satisfies the Cauchy-Schwarz inequality,

$$|\langle\varphi|\psi\rangle|^2 \leq \langle\varphi|\varphi\rangle\langle\psi|\psi\rangle, \tag{3}$$

we can define

$$\begin{aligned}\mathcal{N} &= \{u \in \mathcal{L}(\mathcal{H}) | \langle u|u\rangle = 0\} \\ &= \{u \in \mathcal{L}(\mathcal{H}) \otimes \mathcal{H} | \langle u|v\rangle = 0, \forall v \in \mathcal{L}(\mathcal{H}) \otimes \mathcal{H}\}\end{aligned}$$

to be a subspace of $\mathcal{L}(\mathcal{H}) \otimes \mathcal{H}$. The bilinear function on the newly defined quotient space $\mathcal{L}(\mathcal{H}) \otimes \mathcal{H}/\mathcal{N}$

$$\langle u + \mathcal{N}|v + \mathcal{N}\rangle = \langle u|v\rangle$$

is an inner product. Now let $\mathcal{K}$ denote the Hilbert space, which is the completion of the quotient space $\mathcal{L}(\mathcal{H}) \otimes \mathcal{H}/\mathcal{N}$ equipped with the inner product defined above.

For $T \in \mathcal{L}(\mathcal{H})$, we'll define a linear map $\alpha(T) : \mathcal{L}(\mathcal{H}) \otimes \mathcal{H} \to \mathcal{L}(\mathcal{H}) \otimes \mathcal{H}$ by

$$\alpha(T)\left(\sum_i T_i \otimes \varphi_i\right) = \sum_i (TT_i) \otimes \varphi_i.$$

From the factorization of the operator follows the inequality

$$T_i^* T^* T T_j \leq \|T^*T\| (T_i^* T_j)$$

and furthermore

$$\begin{aligned}&\left\langle \alpha(T)\left(\sum T_i \otimes \varphi_i\right), \alpha(T)\left(\sum T_j \otimes \varphi_j\right)\right\rangle \\ =&\sum_{i,j} \langle \mathcal{E}^*(T_j^* T^* T a_i)\varphi_j|\varphi_j\rangle_{\mathcal{H}} \\ \leq&\|T^*T\| \sum_{i,j} \langle \mathcal{E}^*(T_j^* T_i)\varphi_i|\varphi_j\rangle_{\mathcal{H}} \\ =&\|T\|^2 \left\langle \sum T_i \otimes \varphi_i \sum T_j \otimes \varphi_j\right\rangle.\end{aligned}$$

So $\mathcal{N}$ is left invariant by $\alpha(T)$, but a quotient linear transformation on $\mathcal{L}(\mathcal{H}) \otimes \mathcal{H}$ is induced by $\alpha(T)$. From the inequality we can also see that $\alpha(T)$ is bounded and extends to be a bounded operator on $\mathcal{K}$ and furthermore, α is a unital $*$-homomorphism.

We'll now define a map $V : \mathcal{H} \to \mathcal{K}$ by

$$V(\varphi) = \mathbb{1} \otimes \varphi + \mathcal{N}.$$

We can also verify, that V is in fact bounded, since

$$\|V(\varphi)\|^2 = \langle \mathbb{1} \otimes \varphi | \mathbb{1} \otimes \varphi \rangle = \langle \mathcal{E}^*(\mathbb{1})\varphi | \varphi \rangle \leq \|\mathcal{E}^*(1)\| \cdot \|\varphi\|^2 .$$

Furthermore

$$\langle V^* \alpha(T) V \varphi | \psi \rangle_{\mathcal{H}} = \langle \alpha(T) \mathbb{1} \otimes \varphi, \mathbb{1} \otimes \psi \rangle_{\mathcal{K}} = \langle \mathcal{E}^*(T)\varphi | \psi \rangle_{\mathcal{H}} , \quad \forall x, y \in \mathcal{H},$$

which completes the proof. □

Another important dilation theorem follows from Stinespring's theorem straightforwardly.

Theorem 1.5. (Naimark's dilation theorem) Let A be a POVM associated with the Hilbert space $\mathcal{H}$. There then exists a Hilbert space $\mathcal{K}$, an isometric map $V : \mathcal{H} \to \mathcal{K}$, and a sharp observable $\hat{\mathsf{A}}$, such that

$$\mathsf{A}(X) = V^* \hat{\mathsf{A}}(X) V$$

for every X. The triplet $(\mathcal{K}, \hat{\mathsf{A}}, V)$ is called the Naimark dilation of A.

Proof Let ϕ be a positive and linear map corresponding to the observable A, which makes ϕ completely positive. By applying Stinespring's theorem, we can obtain a $*$-homomorphism α mapping to $\mathcal{L}(\mathcal{H})$ and a bounded and linear $V : \mathcal{H} \to \mathcal{K}$, such that $\mathcal{E}^*(X) = V^* \alpha(X) V$. By letting $\hat{\mathsf{A}}$ be the PVM corresponding to α, the equation in the theorem holds. □

1.5 Measurement Models

Before defining a measurement model for quantum observables, let's recall what is meant by a quantum channel.

Definition 1.11. A quantum channel $\mathcal{V}$ is a completely positive trace preserving linear map $\mathcal{V} : \mathcal{T}(\mathcal{H}) \to \mathcal{T}(\mathcal{K})$, where $\mathcal{T}(\mathcal{H})$ denotes the input state space and $\mathcal{T}(\mathcal{K})$ is the output state space.

The Hilbert spaces $\mathcal{H}$ and $\mathcal{K}$ are not required to be the same space. Moreover, quantum channels are defined as mappings from one set of trace class operators to another, but their usefulness stems from mapping quantum states.

Example 1.8. Let ρ be a state in $\mathcal{H}$ and U some unitary operator. A unitary channel is defined as $\sigma_U(\rho) = U\rho U^*$. To see that this in fact a channel, let our Hilbert space be a composite system, $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$. Then,

$$\begin{aligned}\langle\psi|(\sigma_U \otimes \mathbb{1})\rho\psi\rangle &= \langle\psi|(U \otimes \mathbb{1})\rho(U^* \otimes \mathbb{1})\psi\rangle \\ &= \langle(U^* \otimes \mathbb{1})\psi|\rho(U^* \otimes \mathbb{1})\psi\rangle \\ &\geq 0,\end{aligned}$$

since $(U^* \otimes \mathbb{1})\psi$ just maps the vector ψ to another vector. It's clear that σ_U is completely positive since ρ is positive. The unitary channel is also trace preserving, since

$$\begin{aligned}\mathrm{tr}[\sigma_U(\rho)] &= \mathrm{tr}[U\rho U^*] = \mathrm{tr}[U^*U\rho] \\ &= \mathrm{tr}[\rho] \quad \text{for every } \rho \in \mathcal{S}(\mathcal{H}).\end{aligned}$$

Now that channels and observables in the quantum sense have been defined, it's time to look at what does it mean to perform measurements. To perform a single measurement, the concept of POVM suffices to give the meaningful statistics. But very often, this is not satisfiable at all. Especially in the case of sequential quantum measurements, we are interested in the state after the first measurement and not just the outcome. This is why a measurement model is needed, so that the state possibly altered by the first measurement is preserved and a second measurement can be performed.

Definition 1.12. Let A be a POVM on a system with the Hilbert space $\mathcal{H}$ and let $(\Omega, \mathcal{F})$ be the respective outcome space. To perform a measurement of A we'll need

- A probe system with the Hilbert space $\mathcal{K}$
- $\xi \in \mathcal{K}$ as the initial state of the probe system

- A channel $\mathcal{V} : \mathcal{T}(\mathcal{H} \otimes \mathcal{K}) \to \mathcal{T}(\mathcal{H} \otimes \mathcal{K})$ describing the interaction between the probe and the system itself
- A pointer observable F on the probe system, but with the same outcome space as A

These four items define a measurement model $\mathcal{M} = \langle \mathcal{K}, \xi, \mathcal{V}, \mathsf{F} \rangle$, if they satisfy what's called the probability reproducibility condition

$$\mathrm{tr}[\rho \mathsf{A}(X)] = \mathrm{tr}[(\mathcal{V}(\rho \otimes \xi)(\mathbb{1} \otimes \mathsf{F}(X))], \quad \forall X \in \mathcal{F} \text{ and } \rho \in \mathcal{S}(\mathcal{H}). \tag{4}$$

This condition means, that if a direct measurement of the observable A was performed on the system, this would lead to the same outcome as if the measurement of F was performed on the probe system after the two systems have interacted.

Example 1.9. Let A be a three outcome qubit observable defined as

$$\begin{aligned}
\mathsf{A}(1) &= \frac{1}{3}(\mathbb{1} + \sigma_y), \\
\mathsf{A}(2) &= \frac{1}{3}(\mathbb{1} + \frac{\sqrt{3}}{2}\sigma_x - \frac{1}{2}\sigma_y), \\
\mathsf{A}(3) &= \frac{1}{3}(\mathbb{1} - \frac{\sqrt{3}}{2}\sigma_x - \frac{1}{2}\sigma_y).
\end{aligned}$$

For our measurement model let the Hilbert space of the probe system be the same as $\mathcal{H}$ and we'll choose the pointer observable to be a trivial observable, i.e. $\mathcal{K} = \mathbb{C}^2$ and $\mathsf{F}(X) = \frac{1}{3}\mathbb{1}$.

We need to measure the observable A on a state ρ, so let the initial state be a general pure qubit state. We'll choose the initial state of the probe system to be some other pure qubit state, which in matrix form are

$$\rho = \begin{pmatrix} a & b^* \\ b & 1-a \end{pmatrix}, \quad \xi = \begin{pmatrix} c & d^* \\ d & 1-c \end{pmatrix}.$$

To simplify the calculations a bit, let b and d be real, so $b = b^*$ and $d = d^*$.

A suitable channel for our measurement model is the swap gate, σ_{SWAP}, which operates on a composite system as

$$\sigma_{SWAP}(\rho \otimes \xi) = \xi \otimes \rho.$$

So the measurement model for our observable is $\mathcal{M} = \langle \mathcal{H}, \xi, \sigma_{SWAP}, \mathsf{F} \rangle$. To conclude that this in fact is a proper measurement model for $\mathsf{A}(X)$, the probability reproducibility condition from equation (4) needs to be verified.

From the left hand side of equation (4) we get

$$\begin{aligned}
\mathrm{tr}[\rho \mathsf{A}(1)] =& \frac{1}{3}\,\mathrm{tr}\left[\begin{pmatrix} a + ib^* & b - ia \\ i(1-a) + b & -a - ib + 1 \end{pmatrix}\right] = \frac{1}{3}, \\
\mathrm{tr}[\rho \mathsf{A}(2)] =& \mathrm{tr}[\rho \mathsf{A}(3)] = \frac{1}{3}.
\end{aligned}$$

Since we chose the pointer observable as a trivial observable, we only need calculate the following

$$\begin{aligned}
&\mathrm{tr}[\sigma_{SWAP}(\rho \otimes \xi)(\mathbb{1} \otimes \mathsf{F}(X))] = \frac{1}{3}\,\mathrm{tr}[(\xi \otimes \rho)(\mathbb{1} \otimes \mathbb{1})] = \frac{1}{3}\,\mathrm{tr}[\xi \otimes \rho] \\
=& \frac{1}{3}\,\mathrm{tr}\left[\begin{pmatrix} c\begin{pmatrix} a & b^* \\ b & 1-a \end{pmatrix} & \cdots \\ \cdots & (1-c)\begin{pmatrix} a & b^* \\ b & 1-a \end{pmatrix} \end{pmatrix}\right] \\
=& \frac{1}{3}.
\end{aligned}$$

The chosen measurement model satisfies the probability reproducibility condition and is in fact a measurement of the observable A.

2 Formulation of Sequential Measurements

Before discussing sequential and repeatable measurements in detail, we need to define instruments. Instruments are closely related to the previously defined measurement models and are extremely useful tools in the kind of measurements studied in this thesis.

2.1 Instruments

Observables are identified with positive operator-valued measures (POVMs), which represent the outcome of a measurement, but they don't tell how a measurement

alters the original quantum state. If a measurement is made using only a POVM, there is no possibility of making additional measurements to the already measured state. Measurement and state after measurement can be described using an instrument, which allows a measurement to the possibly changed quantum state. A general definition for an instrument can be given using the definition of a measurement model or measure. After performing the measurement of the observable A by using the measurement $\mathcal{M} = \langle \mathcal{K}, \xi, \mathcal{V}, \mathsf{F} \rangle$, a measurement of another observable B can be done on the system. The result of this is the joint probability distribution for the values of A and B. After the measurement of A an B give the results x and y respectively, the probability for a measured state ρ and state of the probe system ξ is acquired from the formula

$$\mathrm{tr}[\mathcal{V}(\rho \otimes \xi)(\mathsf{B}(y) \otimes \mathsf{F}(x))]. \tag{5}$$

A measurement model for B is not required similarly as for the measurement of A, since the measurement of B is considered as a direct measure to the state. The probability formula can also be written as

$$\mathrm{tr}[\mathcal{V}(\rho \otimes \xi)(\mathsf{B}(y) \otimes \mathsf{F}(x))] = \mathrm{tr}[\mathsf{B}(y)\, \mathrm{tr}_{\mathcal{K}}[\mathcal{V}(\rho \otimes \xi)(\mathbb{1} \otimes \mathsf{F}(x))]],$$

where the lower index on trace means partial trace over the Hilbert space $\mathcal{K}$. The partial trace in the previous equation is defined as

$$\mathcal{I}_X^{\mathcal{M}}(\rho) := \mathrm{tr}_{\mathcal{K}}[\mathcal{V}(\rho \otimes \xi)(\mathbb{1} \otimes \mathsf{F}(x))]. \tag{6}$$

Definition 2.1. A mapping $\mathcal{I}_X^{\mathcal{M}} : (\mathcal{F}, \Omega) \to \mathcal{T}(\mathcal{H}, \mathcal{K})$ is the instrument of a measurement $\mathcal{M}$, if the following three conditions hold

i) $\mathcal{I}_X^{\mathcal{M}}$ is linear, completely positive and trace preserving

ii) $\mathrm{tr}[\mathcal{I}_\Omega^{\mathcal{M}}(\rho)] = 1$ and $\mathrm{tr}[\mathcal{I}_\emptyset^{\mathcal{M}}(\rho)] = 0$

iii) $\mathrm{tr}[\mathcal{I}_{\cup_i X_i}^{\mathcal{M}}(\rho)] = \sum_i \mathrm{tr}[\mathcal{I}_{X_i}^{\mathcal{M}}(\rho)]$ for a state ρ and a collection of disjoint sets $\{X_i\}$

By $\mathcal{T}(\mathcal{H}, \mathcal{K})$ we mean maps from $\mathcal{T}(\mathcal{H})$ to $\mathcal{T}(\mathcal{K})$. This can be understood as instruments mapping states to states in a possibly different Hilbert space, since $\mathcal{H} = \mathcal{K}$ is not required, although often that is the case.

The instrument $\mathcal{I}_X^{\mathcal{M}}$ uniquely defines the measured observable by probabilities

$$p_\rho(\mathsf{A} \in X) = \mathrm{tr}[\mathcal{I}_X^{\mathcal{M}}(\rho)],$$

where $p_\rho(\mathsf{A} \in X)$ means the probability that the measurement of the observable A on the state ρ gives the result from the set X. So each measurement defines a unique instrument and it can be said that $\mathcal{I}_X^{\mathcal{M}}$ is induced by the measurement $\mathcal{M}$. The following theorem defines this more accurately.

Theorem 2.1 (Ozawa's theorem)**.** For each instrument $\mathcal{I}$ there exists a measurement $\mathcal{M}$ for which $\mathcal{I} = \mathcal{I}^{\mathcal{M}}$. Furthermore, it's possible to choose $\mathcal{M}$ so that ξ is a pure state, $\mathcal{V}$ is a unitary channel and F is a sharp observable[12].

Ozawa's theorem defines the instrument for observable $\mathsf{A}(X)$ in the Schrödinger picture. The dual of this is the instrument in the Heisenberg picture, $\mathcal{I}_X^{\mathcal{M}*}$, which defines the observable as

$$\mathsf{A}(X) = \mathcal{I}_X^{\mathcal{M}*}(\mathbb{1}) \quad \forall X \in \mathcal{F}.$$

Using probabilities this can be also be written in the form

$$\mathrm{tr}[\rho \mathsf{A}(X)] = \mathrm{tr}[\mathcal{I}_X^{\mathcal{M}}(\rho)] \quad \forall X \in \mathcal{F}, \rho \in \mathcal{S}(\mathcal{H}).$$

So each measurement model defines a unique instrument and each instrument defines a unique observable. On the other hand every observable defines an equivalence class of instruments and similarly each instrument defines an equivalence class of measurement models.

As a special case of an instrument we have an instrument describing a discrete observable, which we'll need later on.

Definition 2.2. An instrument describing a discrete observable A is a collection of completely positive linear transformations $\mathcal{I}_x$ to $\mathcal{T}(\mathcal{H})$, for which $\mathcal{I}_x^*(\mathbb{1}) = \mathsf{A}(x)$ for all $x \in \Omega_{\mathsf{A}}$.

A similar definition for an instrument describing a continuous observable can be obtained by using a collection from the related sigma-algebra $X \in \mathcal{F}_{\mathsf{A}}$ instead of the element $x \in \Omega_{\mathsf{A}}$ in the definition above.

The dual of the instrument $\mathcal{I}$ in the previous definition is defined by the equation $\mathrm{tr}[L\mathcal{I}_x(T)] = \mathrm{tr}[\mathcal{I}_x^*(L)T]$, where $L \in \mathcal{L}(\mathcal{H})$ and $T \in \mathcal{T}(\mathcal{H})$. I_x^* is the instrument of the observable A in Heisenberg picture and $\mathcal{I}_x$ the instrument in the Schrödinger picture. The instrument $\mathcal{I}_x$ performs a measure of the observable A to a state ρ giving result x and what's left is the new state $\mathcal{I}_x(\rho)$. The probability for this is $\mathrm{tr}[\mathcal{I}_x(\rho)] = \mathrm{tr}[\rho\mathsf{A}(x)]$.

An example of an instrument for a discrete variable is the Lüders instrument, which is defined using the square root of the observable in the following way.

Definition 2.3. The Lüders instrument for a discrete observable is

$$\mathcal{I}_x^L(\rho) = \mathsf{A}(x)^{1/2}\rho\mathsf{A}(x)^{1/2} \tag{7}$$

Clearly $\mathcal{I}^L$ is an instrument, since $\mathrm{tr}[\mathcal{I}_x^L(\rho)] = \mathrm{tr}[\mathsf{A}(x)^{1/2}\rho\mathsf{A}(x)^{1/2}] = \mathrm{tr}[\rho\mathsf{A}(x)]$. Secondly, when summing over all the measurement results x we get $\sum_x tr[\mathcal{I}_x^L(\rho)] = tr[\rho]$, so $\mathcal{I}_x^L$ is an instrument of the observable A.

2.2 Sequential Measurements

Sequential measurement means a situation, where a measurement is performed on a state using an instrument, after which the new, possibly altered state, goes through a measurement of some other observable. Let these observables be A and B. Now it's useful to define a measurement for both of these observables. The measurement of A is $\mathcal{M}_a = \langle \mathcal{K}_a, \xi_a, \mathcal{V}_a, \mathsf{F}_a \rangle$ and respectively for B the measurement is $\mathcal{M}_b = \langle \mathcal{K}_b, \xi_b, \mathcal{V}_b, \mathsf{F}_b \rangle$. Now the sequential measurement means measuring the observable A and B after that.

Let I_{ab} be a mapping $\mathcal{T}(\mathcal{H} \otimes \mathcal{K}_a \otimes \mathcal{K}_b) \to \mathcal{T}(\mathcal{H} \otimes \mathcal{K}_b \otimes \mathcal{K}_a)$, which switches the position of Hilbert spaces $\mathcal{K}_a$ and $\mathcal{K}_b$ in the tensor product. In addition we'll define $\widetilde{\mathcal{V}}_a = \mathcal{V}_a \otimes \mathbb{1}_a$ and $\widetilde{\mathcal{V}}_b = I_{ab}^{-1} \circ \mathcal{V}_b \otimes \mathbb{1}_a \circ I_{ab}$. For the partial trace the upper index $*$ means mapping from the Hilbert space $\mathcal{H} \otimes \mathcal{K}_a \otimes \mathcal{K}_b$ and partial trace without the index is a mapping from the Hilbert space, which is the tensor product of two spaces, i.e. $\mathrm{tr}_{ab}^* : \mathcal{T}(\mathcal{H} \otimes \mathcal{K}_a \otimes \mathcal{K}_b) \to \mathcal{T}(\mathcal{H})$ and $\mathrm{tr}_a : \mathcal{T}(\mathcal{H} \otimes \mathcal{K}_a) \to \mathcal{T}(\mathcal{H})$. Using

these notations we get

$$\begin{aligned}
\mathcal{I}_{X\times Y}^{\mathcal{M}_{ab}}(\rho) &= \mathrm{tr}_{ab}^*[\widetilde{\mathcal{V}}_b \circ \widetilde{\mathcal{V}}_a(\rho \otimes \xi_a \otimes \xi_b)\mathbb{1} \otimes \mathsf{F}_a(X) \otimes \mathsf{F}_b(Y)] \\
&= \mathrm{tr}_b[\mathrm{tr}_a^*[\widetilde{\mathcal{V}}_b(\mathcal{V}_a(\rho \otimes \xi_a) \otimes \xi_b)\mathbb{1} \otimes \mathsf{F}_a(X) \otimes \mathsf{F}_b(Y)]] \\
&= \mathrm{tr}_b[\mathcal{V}_b(\mathrm{tr}_a[\mathcal{V}_a(\rho \otimes \xi_a)\mathbb{1} \otimes \mathsf{F}_a(X)] \otimes \xi_b)\mathbb{1} \otimes \mathsf{F}_b(Y)] \\
&= \mathrm{tr}_b[\mathcal{V}_b(\mathcal{I}_X^{\mathcal{M}_a}(\rho) \otimes \xi_b)\mathbb{1} \otimes \mathsf{F}_b(Y)] \\
&= \mathcal{I}_Y^{\mathcal{M}_b}(\mathcal{I}_X^{\mathcal{M}_a}(\rho)) \\
&= \mathcal{I}_Y^{\mathcal{M}_b} \circ \mathcal{I}_X^{\mathcal{M}_a}(\rho).
\end{aligned}$$

The instrument $\mathcal{I}_{X\times Y}^{\mathcal{M}_{ab}}(\rho)$ of the joint measurement $\mathcal{M}_{ab} = \langle \mathcal{K}_a \otimes \mathcal{K}_b, \xi_a \otimes \xi_b, \widetilde{\mathcal{V}}_b \circ \widetilde{\mathcal{V}}_a, \mathsf{F}_a \otimes \mathsf{F}_b \rangle$ is the composition of the instruments defined by the measurements $\mathcal{M}_a$ and $\mathcal{M}_b$. Generally $\mathcal{I}_{X\times Y}^{\mathcal{M}_{ab}}(\rho) \neq \mathcal{I}_{X\times Y}^{\mathcal{M}_{ba}}(\rho)$, so the obtained measurement results depend on the order in which the measurements of the observables A and B are performed. If the instruments of the joint measurement are equal and don't depend on the order of the measurements, then A and B are compatible observables.

Next we will consider the probabilities of sequential measurements. For all states $\rho \in \mathcal{S}(\mathcal{H})$ a mapping

$$X \times Y \mapsto [0,1], \quad Z \mapsto \mathrm{tr}[\mathcal{I}_Z^{\mathcal{M}_{ab}}(\rho)]$$

is a probability measure for the joint probability of the observables A and B. The probability of performing a measurement on ρ obtaining the result $X \times Y$ is

$$\begin{aligned}
\mathrm{tr}[\mathcal{I}_{X\times Y}^{\mathcal{M}_{ab}}(\rho)] &= \mathrm{tr}[\mathcal{I}_Y^{\mathcal{M}_b} \circ \mathcal{I}_X^{\mathcal{M}_a}(\rho)] \\
&= \mathrm{tr}[\mathcal{I}_X^{\mathcal{M}_a}(\rho)\mathsf{B}(Y)].
\end{aligned}$$

This probability can be interpreted as measuring the observable B for the unnormalized state $\mathcal{I}_X^{\mathcal{M}_a}(\rho)$ giving the result Y. This unnormalized state however depends on the measurement of A giving the result X. In this case it's natural to describe the situation using conditional probabilities. Then it can be said that the measurement of the observable B gives the result Y if the measurement of the A gives the result X.

2.3 Repeatable Measurements

After talking about sequential measurements it's natural to move on to repeatable measurements. Let $\mathcal{M}_a = \mathcal{M}$ be the measurement of A like in the case of sequential measurements. Now the measurement of the same observable is repeated on the state instead of measuring different observables. If the measurement of A is performed twice, the later measurement might give more information. The measurement $\mathcal{M}$ is repeatable if repeating the measurement doesn't lead to a new result from the probabilistic point of view, i.e.

$$\begin{aligned}\mathrm{tr}[\mathcal{I}^{\mathcal{M}_{aa}}_{X\times Y}(\rho)] &= \mathrm{tr}[\mathcal{I}^{\mathcal{M}_a}_{Y} \circ \mathcal{I}^{\mathcal{M}_a}_{X}(\rho)] \\ &= \mathrm{tr}[\mathcal{I}^{\mathcal{M}_a}_{X\cap Y}(\rho)], \quad \forall X, Y \in \mathcal{F}, \rho \in \mathcal{S}(\mathcal{H}),\end{aligned}$$

which can be written in an equivalent form

$$\mathrm{tr}[\mathcal{I}^{\mathcal{M}_a}_{X} \circ \mathcal{I}^{\mathcal{M}_a}_{X}(\rho)] = \mathrm{tr}[\mathcal{I}^{\mathcal{M}_a}_{X}(\rho)]. \tag{8}$$

This condition describes the so called weak repeatability of a measurement, which distinguishes it from the strong repeatability of a measurement

$$\mathcal{I}^{\mathcal{M}_{aa}}_{X\times Y}(\rho) = \mathcal{I}^{\mathcal{M}_a}_{X\cap Y}(\rho) \quad \forall X, Y \in \mathcal{F}, \rho \in \mathcal{S}(\mathcal{H}).$$

Repeatable measurements of the observable A can be performed, if the measurement $\mathcal{M}$ is repeatable.

Example 2.1. The Lüders channel of a sharp qubit observable is discussed with more detail in example (3.1), but it can also be used to show a repeatable measurement. Let A be a sharp qubit observable defined in equation (1). The corresponding instrument is

$$\mathcal{I}_{\pm}(\rho) = \sqrt{\mathsf{A}(\pm 1)}\rho\sqrt{\mathsf{A}(\pm 1)},$$

which clearly is weakly repeatable, since

$$\begin{aligned}\mathrm{tr}[\mathcal{I}_{\pm} \circ \mathcal{I}_{\pm}(\rho)] &= \mathrm{tr}[\mathcal{I}_{\pm}(\sqrt{\mathsf{A}(\pm 1)}\rho\sqrt{\mathsf{A}(\pm 1)})] \\ &= \mathrm{tr}[\sqrt{\mathsf{A}(\pm 1)}\sqrt{\mathsf{A}(\pm 1)}\rho\sqrt{\mathsf{A}(\pm 1)}\sqrt{\mathsf{A}(\pm 1)}] \\ &= \mathrm{tr}[\sqrt{\mathsf{A}(\pm 1)}\rho\sqrt{\mathsf{A}(\pm 1)}] \\ &= \mathrm{tr}[\mathcal{I}_{\pm}(\rho)],\end{aligned}$$

since the sharp observable A is a projection.

A fundamental result when discussing repeatable measurements is that only discrete observables admit repeatable measurements, as stated in the following theorem [1].

Theorem 2.2. Let A be an observable defined on a measurable space $(\Omega, \mathcal{F})$. If A admits a repeatable measurement, then the measurement is discrete.

We omit the proof of this theorem to the source[1]. An obvious but important consequence of this theorem is the fact that no continuous observable can admit repeatability. Later we provide a looser definition of repeatability to be used with continuous observables.

3 Properties of Quantum Measurements

When considering quantum measurements, there are multiple interesting possibilities, when measuring multiple observables. Next, we'll discuss a few different measurement scenarios and introduce a few important concepts. Two important situations regarding sequential measurements are non-disturbing measurements and jointly measurable observables. Non-disturbance means that measuring B after A gives the same measurement outcome for B as just measuring this single observable. Joint measurement on two observables relates very closely on sequential measurements, since a pair of jointly measurable observables can be modelled as a sequential measurement. These combined with compatibility of a channel and an observable lead up to the universality property of an observable.

3.1 Compatibility

The first important property for a quantum observable is it's compatibility with a quantum channel. This means an observable and a channel that are both part of the same measurement, which is explained in the following definition.

Definition 3.1. For an observable A on $\mathcal{H}$, a channel $\mathcal{V} : \mathcal{L}(\mathcal{K}) \to \mathcal{L}(\mathcal{H})$ is called an A-channel if there exists an instrument $\mathcal{I}$, for which

$$\mathcal{I}_x(\mathbb{1}_{\mathcal{K}}) = \mathsf{A}(x), \quad \sum_x \mathcal{I}_x(\rho) = \mathcal{V}(\rho) \ . \tag{9}$$

This means that the channel $\mathcal{V}$ and observable A are compatible if they are parts of the same instrument $\mathcal{I}$.

Example 3.1. Consider the sharp qubit observable introduced in the first chapter, $\mathsf{A}(\pm 1) = \frac{1}{2}(\mathbb{1} \pm \vec{r} \cdot \vec{\sigma})$. Let's choose the Lüders instrument of the observable A, $\mathcal{I}_{\pm 1}(\rho) = \sqrt{\mathsf{A}(\pm 1)}\rho\sqrt{\mathsf{A}(\pm 1)}$. $\mathcal{I}$ is clearly an A-compatible instrument, since

$$\mathrm{tr}[\mathcal{I}_{\pm 1}(\rho)] = \mathrm{tr}[A(\pm 1)\rho], \quad \sum_x \mathrm{tr}[\mathcal{I}_x(\rho)] = \mathrm{tr}[\rho].$$

By definition, the A-compatible channel is

$$\mathcal{V}(\rho) = \mathcal{I}_{-1}(\rho) + \mathcal{I}_{+1}(\rho) = \sqrt{\mathsf{A}(-1)}\rho\sqrt{\mathsf{A}(-1)} + \sqrt{\mathsf{A}(+1)}\rho\sqrt{\mathsf{A}(+1)}.$$

A quick calculation shows that

$$\begin{aligned}
\mathsf{A}(+1)^2 &= \frac{1}{4}\begin{pmatrix} 1+r_z & r_x - ir_y \\ r_x + ir_y & 1 - r_z \end{pmatrix}\begin{pmatrix} 1+r_z & r_x - ir_y \\ r_x + ir_y & 1 - r_z \end{pmatrix} \\
&= \frac{1}{4}\begin{pmatrix} (1+r_z)^2 + (r_x - ir_y)(r_x + ir_y) & (r_x - ir_y)(1-r_z) + (r_x - ir_y)(1+r_z) \\ (r_x + ir_y)(1-r_z) + (r_x + ir_y)(1+r_z) & (1-r_z)^2 + (r_x - ir_y)(r_x + ir_y) \end{pmatrix} \\
&= \frac{1}{4}\begin{pmatrix} 1 + 2r_z + r_z^2 + r_x^2 + r_y^2 & 2r_x - 2ir_y \\ 2r_x + 2ir_y & 1 - 2r_z + r_z^2 + r_x^2 + r_y^2 \end{pmatrix} \\
&= \frac{1}{2}\begin{pmatrix} 1+r_z & r_x - ir_y \\ r_x + ir_y & 1 - r_z \end{pmatrix} = \mathsf{A}(+1) \Leftrightarrow \sqrt{\mathsf{A}(+1)} = \mathsf{A}(+1).
\end{aligned}$$

A similar calculation shows that $\sqrt{\mathsf{A}(-1)} = \mathsf{A}(-1)$ holds as well, so we can simplify

the formula for the A-channel a bit

$$\begin{aligned}
\mathcal{V}(\rho) =& \sqrt{\mathsf{A}(-1)}\rho\sqrt{\mathsf{A}(-1)} + \sqrt{\mathsf{A}(+1)}\rho\sqrt{\mathsf{A}(+1)} \\
=& \mathsf{A}(-1)\rho\mathsf{A}(-1) + \mathsf{A}(+1)\rho\mathsf{A}(+1) \\
=& \frac{1}{4}\left[(\mathbb{1} - \vec{r}\cdot\vec{\sigma})\rho(\mathbb{1} - \vec{r}\cdot\vec{\sigma}) + (\mathbb{1} + \vec{r}\cdot\vec{\sigma})\rho(\mathbb{1} + \vec{r}\cdot\vec{\sigma})\right] \\
=& \frac{1}{4}[\rho - \rho\vec{r}\cdot\vec{\sigma} - \vec{r}\cdot\vec{\sigma}\rho + \vec{r}\cdot\vec{\sigma}\rho\vec{r}\cdot\vec{\sigma} + \rho + \rho\vec{r}\cdot\vec{\sigma} + \vec{r}\cdot\vec{\sigma}\rho + \vec{r}\cdot\vec{\sigma}\rho\vec{r}\cdot\vec{\sigma}] \\
=& \frac{1}{2}[\rho + \vec{r}\cdot\vec{\sigma}\rho\vec{r}\cdot\vec{\sigma}].
\end{aligned}$$

3.2 Non-Disturbance

In the world of quantum measurements, a distinctive feature is that performing a measurement for two different observables usually disturb one another. The aim is to show that there exists observables that can be measured without any disturbance between the measurement.

Definition 3.2. An observable A can be measured without disturbing the observable B if there is an instrument $\mathcal{I}$ implementing A and

$$\mathrm{tr}[\mathcal{I}_\Omega(\rho)\mathsf{B}(y)] = \mathrm{tr}[\rho\mathsf{B}(y)], \quad \forall\rho \in \mathcal{S}(\mathcal{H}), y \in \Omega_{\mathsf{B}} \tag{10}$$

This means that the measurement results of the observable B are equal for all input states ρ and output states $\mathcal{I}_\Omega(\rho)$ pairwise.

Example 3.2. An example of a situation where the non-disturbance condition holds is when A and B commute and $\mathcal{V}$ is chosen to be the Lüders channel of A,

$$[\mathsf{A}, \mathsf{B}] = 0, \quad \mathcal{I}^L(\rho) = \sqrt{\mathsf{A}(x)}\rho\sqrt{\mathsf{A}(x)}.$$

Now we can see via a quick calculation that

$$\begin{aligned}
\mathrm{tr}[\mathcal{I}^L(\rho)\mathsf{B}(y)] &= \sum_x \mathrm{tr}[\sqrt{\mathsf{A}(x)}\rho\sqrt{\mathsf{A}(x)}\mathsf{B}(y)] \\
&= \sum_x \mathrm{tr}[\mathsf{A}(x)\rho\mathsf{B}(y)] \\
&= \mathrm{tr}[\rho\mathsf{B}(y)].
\end{aligned}$$

There is in fact a theorem related to the previous example which states that the two observables commute if and only if the Lüders channel does not disturb the measurement. We omit the proof of this theorem to [2].

Non-disturbance may hold also in different cases than the one described above, but the general case is that the first measurement disturbs the measurement performed on the output state of the channel. In the case of two jointly measurable observables which do disturb one another, the joint measurement cannot be implemented as a sequential measurement, since there is no A-instrument which satisfies the non-disturbance condition[9].

Another example of pairs of observables that do not disturb one another we'll consider an informationally complete observable B. In this case each state ρ is uniquely determined by the probabilities $\mathrm{tr}[\rho\mathsf{B}(y)]$. The non-disturbance condition (10) now yields $\mathcal{I}_\Omega(\rho) = \rho$. But for non-trivial observables A the state is altered in some manner and even the informationally complete observable B is disturbed.

Non-disturbance does not imply a symmetric relation between the two observables A and B. The pair is said to be non-disturbing if A can be measured without disturbing B or vice versa. In the case that the measurement of either observable does not disturb the other, the pair of observables is said to be mutually non-disturbing, which is a stronger relation than just non-disturbance[6].

The concept of non-disturbance brings us close to the next important feature, jointly measurable observables.

3.3 Joint Measurability

Two observables A and B are said to be jointly measurable if there exists a third observable M with the outcome space $\Omega_\mathsf{A} \times \Omega_\mathsf{B}$. Also, the observables A and B are obtained as the marginals of M, i.e.

$$\mathsf{A}(x) = \sum_y \mathsf{M}(x,y), \quad \mathsf{B}(y) = \sum_x \mathsf{M}(x,y) \quad \forall x \in \Omega_\mathsf{A}, y \in \Omega_\mathsf{B}.$$

A special case of jointly measurable observables is a pair of observables $\mathsf{A}(x)$ and $\mathsf{B}(y)$ that commute, that is

$$[\mathsf{A}(x), \mathsf{B}(y)] = 0,$$

though commutativity is by no means required for two observables to be jointly measurable.

In the case of non-disturbing measurements, the observables A and B are also jointly measurable. This can be shown with ease by choosing the common observable as $\mathsf{M}(x, y) = \mathcal{I}_x^*(\mathsf{B}(y))$.

Example 3.3. For a general effect in $\mathcal{H} = \mathbb{C}^2$, we'll mark

$$\mathsf{A}(\alpha, \vec{r}) = \frac{1}{2}(\alpha \mathbb{1} + \vec{r} \cdot \vec{\sigma}), \quad \alpha \in \mathbb{R}, \vec{r} \in \mathbb{R}^3,$$

and recall the requirement for $\mathsf{A}(\alpha, \vec{r})$ to be an effect,

$$\|\vec{r}\| \leq \alpha \leq 2 - \|\vec{r}\| .$$

Next we'll define a qubit observable E, for which

$$\mathsf{E}^{\alpha,\vec{r}}(1) = \mathsf{A}(\alpha, \vec{r}), \quad \mathsf{E}^{\alpha,\vec{r}}(0) = \mathbb{1} - \mathsf{A}(\alpha, \vec{r}).$$

Now let $\mathsf{E}^{\alpha,\vec{a}}$ and $\mathsf{E}^{\beta,\vec{b}}$ be two different observables. There are many ways to check which conditions need to hold for these two observables to be jointly measurable[8]. The following formulation for unbiased variables is probably the simplest.

If $\alpha = \beta = 1$, and $\|\vec{a}\| \leq 1$, $\left\|\vec{b}\right\| \leq 1$, then $\mathsf{E}^{1,\vec{a}}$ and $\mathsf{E}^{1,\vec{b}}$ are jointly measurable if and only if

$$\left\|\vec{a} + \vec{b}\right\| + \left\|\vec{a} - \vec{b}\right\| \leq 2.$$

If this requirement is fulfilled, the joint observable M is a four outcome variable defined by

$$\mathsf{E}^{1,\vec{a}}(1) = \mathsf{M}(1,1) + \mathsf{M}(1,0), \quad \mathsf{E}^{1,\vec{b}}(1) = \mathsf{M}(1,1) + \mathsf{M}(0,1)$$
$$\mathsf{E}^{1,\vec{a}}(0) = \mathsf{M}(0,0) + \mathsf{M}(0,1), \quad \mathsf{E}^{1,\vec{b}}(0) = \mathsf{M}(0,0) + \mathsf{M}(1,0).$$

In fact, it follows that the joint observable is determined by a single effect, $\mathsf{M}(1,1)$. The other effects can be derived from this as well as the original observables E. Thus, the joint observable is of the form $\mathsf{M}(1,1) = \frac{1}{2}(\gamma\mathbb{1} + \vec{g}\cdot\vec{\sigma})$ for some parameters $\gamma \in \mathbb{R}, \vec{g} \in \mathbb{R}^3$. Keeping in mind that every one of $\mathsf{M}(0,0), \mathsf{M}(0,1), \mathsf{M}(1,0)$ also needs to be an effect leads to some constraints for these parameters γ and $\vec{g}$. After a bit of calculation, we can write these using the variables defined in the effects $\mathsf{A}(\alpha,\vec{a})$, $\mathsf{A}(\beta,\vec{b})$ and $\mathsf{M}(1,1)$

$$\begin{aligned}
\mathsf{M}(0,0) &= \frac{1}{2}\left((2-\alpha-\beta+\gamma)\mathbb{1} + (-\vec{a}-\vec{b}+\vec{g})\cdot\vec{\sigma}\right)\\
\mathsf{M}(0,1) &= \frac{1}{2}\left((\beta-\gamma)\mathbb{1} + (\vec{b}-\vec{g})\cdot\vec{\sigma}\right)\\
\mathsf{M}(1,0) &= \frac{1}{2}\left((\alpha-\gamma)\mathbb{1} + (\vec{a}-\vec{g})\cdot\vec{\sigma}\right)
\end{aligned}$$

and the corresponding eigenvalues for $\mathsf{M}(0,0)$, $\mathsf{M}(0,1)$ and $\mathsf{M}(1,1)$ are, respectfully

$$\begin{aligned}
\lambda_\pm &= \frac{1}{2}\left(2-\alpha-\beta+\gamma \pm \left\|-\vec{a}-\vec{b}+\vec{g}\right\|\right)\\
\lambda_\pm &= \frac{1}{2}\left(\beta-\gamma \pm \left\|\vec{b}-\vec{g}\right\|\right)\\
\lambda_\pm &= \frac{1}{2}(\alpha-\gamma \pm \|\vec{a}-\vec{g}\|).
\end{aligned}$$

Since $\mathsf{M}(0,0)$, $\mathsf{M}(0,1)$ and $\mathsf{M}(1,1)$ are effects, the eigenvalues $\lambda_\pm$ need to satisfy $0 \leq \lambda_-$ and $\lambda_+ \leq 1$. Through this we acquire the following inequations

$$\begin{aligned}
\left\|-\vec{a}-\vec{b}+\vec{g}\right\| \leq 2-\alpha-\beta+\gamma &\leq 2-\left\|-\vec{a}-\vec{b}+\vec{g}\right\|\\
\left\|\vec{b}-\vec{g}\right\| \leq \beta-\gamma &\leq 2-\left\|\vec{b}-\vec{g}\right\|\\
\|\vec{a}-\vec{g}\| \leq \alpha-\gamma &\leq 2-\|\vec{a}-\vec{g}\|,
\end{aligned}$$

so the choice of the initial parameters α, β, $\vec{a}$ and $\vec{b}$ restrict the parameters γ and $\vec{g}$ used to define the effect $\mathsf{M}(1,1)$.

3.4 Universality

It has been shown that any pair of observables, which are jointly measurable, can be modelled as sequential measurements even if the later observable isn't decided but after the first measurement is done. This means that it is possible to measure

an observable in such a way that measurements on the output state are not limited any more than what joint measurability limits them. This feature of the first measurement is called universality and it does not hold for all measurements [7].

By focusing on two measurements it's enough to describe the first one as a pair $(\mathsf{A}, \mathcal{V})$, which consist of an observable and a channel, respectively. The second measurement can be specified just as an observable B, since the output state won't be needed. The measurement distributions for an initial state ρ then are $\mathrm{tr}[\rho\mathsf{A}(x)]$ and $\mathrm{tr}[\mathcal{V}(\rho)\mathsf{B}(y)]$ and the outcome of the second measurement is useful to write in the Heisenberg picture, $\mathrm{tr}[\rho\mathcal{V}^*(\mathsf{B}(y))]$.

Since there usually is some disturbance while performing a sequential measurement, measuring observables A and B means measuring A and B', which is an altered version of B. The reason for this modification is to reduce the disturbance caused by the first measurement. In this case the channel $\mathcal{V}$ and the observable B' should give

$$\mathrm{tr}[\mathcal{V}(\rho)\mathsf{B}(y)] = \mathrm{tr}[\rho\mathsf{B}'(y)], \tag{11}$$

for all states ρ and measurement outcomes Y. In the Heisenberg picture this can be written as

$$\mathcal{V}^*(\mathsf{B}'(y)) = \mathsf{B}(y), \quad \forall y \in \Omega_{\mathsf{B}}. \tag{12}$$

The interpretation of these equations is as follows: After measuring A a measurement of B' is performed, after which B is obtained via the channel in the Heisenberg picture. An alternate approach is to tailor the channel and observable B' in a way, which gives the measurement of B. With these jointly measurable observables, it's useful to define the universality property.

Definition 3.3. A channel $\mathcal{V}$ related to the observable A has the universal property if for each observable B jointly measurable with A there exists another observable B' such that

$$\mathrm{tr}[\mathcal{V}(\rho)\mathsf{B}'(y)] = \mathrm{tr}[\rho\mathsf{B}(y)] \tag{13}$$

for every input state ρ and outcome y.

What this means is that the only limitation for the measurements performed later is joint measurability. The exciting fact regarding this feature is that these kinds of measurements actually exists[7].

Theorem 3.1. For every observable A an A-channel $\mathcal{V}_{\mathsf{A}}$ having the universality feature exists.

Proof Let $(\mathcal{K}, \hat{\mathsf{A}}, V)$ be a minimal Naimark dilation of the observable A. The dilation being minimal means that the set $\{\sum_x c_x \hat{\mathsf{A}}(x) V\psi, \quad c_x \in \mathbb{C}, \psi \in \mathcal{H}\}$ is dense in the Hilbert space $\mathcal{K}$. Let $\mathcal{V}_{\mathsf{A}}$ be a channel with input space $\mathcal{H}$ and output space $\mathcal{K}$ defined as

$$\mathcal{V}_{\mathsf{A}}(\rho) = \sum_x \hat{\mathsf{A}}(x) V \rho V^* \hat{\mathsf{A}}(x),$$

which is an A-channel since

$$\operatorname{tr}\left[\hat{\mathsf{A}}(x) V \rho V^* \hat{\mathsf{A}}(x)\right] = \operatorname{tr}\left[\rho V^* \hat{\mathsf{A}}(x) V\right] = \operatorname{tr}\left[\rho \mathsf{A}(x)\right].$$

Next we'll show that the channel $\mathcal{V}_{\mathsf{A}}$ in fact has the universal property.

Since every observable B, which is jointly measurable with A, can be written as $\mathsf{B}(y) = (\mathcal{V}^{\mathsf{B}})^*(|y\rangle\langle y|)$, $\{|y\rangle\}_{y\in\Omega_{\mathsf{B}}}$ being an orthonormal basis and $\mathcal{V}^{\mathsf{B}}$ a channel defined by the equation

$$\mathcal{V}^{\mathsf{B}}(\rho) = \sum_y \operatorname{tr}[\rho \mathsf{B}(y)] \, |y\rangle\langle y| \, .$$

We need to show that there exists a new channel Γ^{B} related to the already defined channels,

$$\mathcal{V}^{\mathsf{B}} = \Gamma^{\mathsf{B}} \circ \mathcal{V}_{\mathsf{A}}. \tag{14}$$

Using the channel Γ^{B} in the Heisenberg picture, we'll define a new observable

$$\mathsf{B}'(y) = (\Gamma^{\mathsf{B}})^*(|y\rangle\langle y|).$$

Moreover, the observable B' satisfies the equation (13) since

$$\begin{aligned}
\operatorname{tr}[\mathcal{V}_{\mathsf{A}}(\rho)\mathsf{B}'(y)] &= \operatorname{tr}[\Gamma^{\mathsf{B}}(\mathcal{V}_{\mathsf{A}}(\rho)) \, |y\rangle\langle y| \\
&= \operatorname{tr}[\mathcal{V}^{\mathsf{B}}(\rho) \, |y\rangle\langle y|] \\
&= \operatorname{tr}[\rho \mathsf{B}(y)].
\end{aligned}$$

The only thing needed to complete the proof is to show that for each observable B, which is jointly measurable with the observable A, the channel Γ^{B} defined in equation (14) exists.

Since A and B are jointly measurable, they have a joint observable M with a Naimark dilation $(\mathcal{K}', \hat{\mathsf{M}}, V')$. We'll define a sharp observable $\hat{\mathsf{A}}'$ as the margin of $\hat{\mathsf{M}}$, $\hat{\mathsf{A}}'(X) = \sum_y \hat{\mathsf{M}}(x,y)$. From $\mathsf{A}(x) = \sum_Y \mathsf{M}(x,y)$ it follows that $(\mathcal{K}', \hat{\mathsf{A}}', V')$ is a Naimark dilation of the observable A. We chose the original Naimark dilation of A to be minimal, so there exists an isometric operator J such that

$$J : \mathcal{K} \to \mathcal{K}', \quad \hat{A}'(x)J = J\hat{A}(x), \quad V' = JV.$$

From $\hat{\mathsf{M}}(x,y)\hat{\mathsf{M}}(x',y') = \delta_{xx'}\hat{\mathsf{M}}(x,y)$ it follows that

$$\hat{\mathsf{M}}(x,y)J\hat{\mathsf{A}}(x') = \delta_{xx'}\hat{\mathsf{M}}(x,y)J. \tag{15}$$

Using this auxiliary result we can rewrite the channel $\mathcal{V}^{\mathsf{B}}$ as

$$\mathcal{V}^{\mathsf{B}}(\rho) = \sum_{x,y} \operatorname{tr}\left[\hat{\mathsf{A}}(x)V\rho V^*\hat{\mathsf{A}}(x)J^*\hat{\mathsf{M}}(x,y)J\right] |y\rangle\langle y|$$

and define the channel Γ^{B} as

$$\Gamma^{\mathsf{B}}(\rho) = \sum_{x,y} \operatorname{tr}\left[\rho J^*\hat{\mathsf{M}}(x,y)J\right] |y\rangle\langle y|.$$

Now by the result can be verified with a straightforward calculation

$$\begin{aligned}
&\Gamma^{\mathsf{B}} \circ \mathcal{V}_{\mathsf{A}}(\rho) \\
=&\Gamma^{\mathsf{B}}\left(\sum_{x'} \hat{A}(x')V\rho V^*\hat{A}(x')\right) \\
=&\sum_{x,y} \operatorname{tr}\left[\left(\sum_{x'} \hat{A}(x')V\rho V^*\hat{A}(x')\right) J^*\mathsf{M}(x,y)J\right] |y\rangle\langle x| \\
=&\sum_{x,y} \operatorname{tr}\left[\hat{\mathsf{A}}(x)V\rho V^*\hat{A}(x)J^*\hat{M}(x,y)J\right] |y\rangle\langle y| \\
=&\mathcal{V}^{\mathsf{B}}(\rho),
\end{aligned}$$

where equation (15) was used again. □

Therefore the joint measurability turns into a useful attribute for two observables, since these channels can be found and sequential measurements can be performed on the joint observables.

Example 3.4 (Using Naimark Dilation)**.** Let us consider the general qubit observable from example 1.4 in $\mathcal{H} = \mathbb{C}^2$. Using the eigenvalues $\lambda_\pm = \frac{1}{2}(\alpha \pm \|\vec{r}\|)$ and orthonormal basis $\varphi_0 = \left(\begin{smallmatrix}1\\0\end{smallmatrix}\right), \varphi_1 = \left(\begin{smallmatrix}0\\1\end{smallmatrix}\right)$ the spectral decomposition of the general qubit effect can be presented as

$$\mathsf{A} = \sum_i |\beta_i\rangle \langle\beta_i| ,$$

where $\beta_i = \varphi_i/\lambda_i$, $\lambda_0 := \lambda_-$, $\lambda_1 := \lambda_+$. We'll choose the dilated vector space to be $\mathcal{K} = \mathbb{C}^2 \oplus \mathbb{C}^2$ with an orthogonal basis $\{\psi_j\}$. Let $V : \mathcal{H} \to \mathcal{K}$ be an isometry, for which $\mathsf{A} = V^*\hat{\mathsf{A}}V$ and $V^*\psi_i = \beta_i$. An example of an isometry like this would be

$$V^*(\beta_i, \beta_j) = a\beta_i, \quad (\beta_i, \beta_j) \in \mathcal{K}, a \in \mathbb{C}.$$

With this, an instrument can be defined

$$\mathcal{I}(\rho) = \hat{\mathsf{A}}V\rho V^*\hat{\mathsf{A}}$$

and a channel $\mathcal{V}_\mathsf{A}(\rho)$, summing over every X as

$$\mathcal{V}_\mathsf{A}(\rho) = \sum_X \hat{\mathsf{A}}(X)V\rho V^*\hat{\mathsf{A}}(X) = \sum_X \mathcal{I}_X(\rho)$$

we get

$$\mathrm{tr}[\mathcal{I}_X(\rho)] = \mathrm{tr}[\hat{\mathsf{A}}V\rho V^*\hat{\mathsf{A}}] = \mathrm{tr}[\hat{\mathsf{A}}V\rho V^*] = \mathrm{tr}[\rho\mathsf{A}].$$

So there exists a channel $\mathcal{V}_\mathsf{A}$ compatible with A.

3.5 Broadcasting

As mentioned before, quantum channels are linear transformations are completely positive trace preserving linear maps. A channel is said to be a broadcasting channel

if the channel takes a single system as input and outputs two similar systems[6], that is

$$\mathcal{V} : \mathcal{S}(\mathcal{H}) \to \mathcal{S}(\mathcal{H}_A \otimes \mathcal{H}_B), \quad \mathcal{H}_A = \mathcal{H}_B = \mathcal{H}.$$

This type of channel broadcasts a state, if the partial traces of the output state are equal to the input state,

$$\mathrm{tr}_A[\mathcal{V}(\rho)] = \mathrm{tr}_B[\mathcal{V}(\rho)] = \rho.$$

Similarly, a subset of states is broadcastable if there exists a channel, which broadcasts each state in that subset. The previous equation is equivalent to the following[6]

$$\mathrm{tr}[\rho \mathsf{A}(x)] = \mathrm{tr}[\mathcal{V}(\rho)\mathsf{A}(x) \otimes \mathbb{1}] = \mathrm{tr}[\mathcal{V}(\rho)\mathbb{1} \otimes \mathsf{A}(x)] \tag{16}$$

for all POVMs A in $\mathcal{H}$ and all outcomes X. To change this up a bit, we can choose the broadcasting procedure to apply an all states, but not all POVMs. By limiting the broadcasting to only some observables, the following definition is acquired.

Definition 3.4. An observable A is broadcasted by a channel $\mathcal{V}$ if the condition in equation (16) holds for all states $\rho \in \mathcal{S}(\mathcal{H})$. A subset of observables is broadcastable if there exists a channel, which broadcasts all observables in that subset.

While the channel $\mathcal{V}$ in the definition is in the Schrödinger picture, an equivalent condition can be stated in the Heisenberg picture of the channel, $\mathcal{V}^*$

$$\mathsf{A}(x) = \mathcal{V}^*(\mathsf{A}(x) \otimes \mathbb{1}) = \mathcal{V}^*(\mathbb{1} \otimes \mathsf{A}(x)).$$

Example 3.5. Let A be a sharp observable in $\mathcal{H} = \mathbb{C}^2$ of the form $\mathsf{A}(x) = \mathrm{diag}(\alpha_0(x), \alpha_1(x)) = \mathrm{diag}(\alpha_0, \alpha_1)$ where $\alpha_0 + \alpha_1 = 1$ and let $\{\varphi_0, \varphi_1\}$ be an orthonormal basis of $\mathcal{H}$. We'll define a channel $\mathcal{V} : \mathcal{H} \to \mathcal{H} \otimes \mathcal{H}$ as

$$\mathcal{V}(\rho) = \sum_{i=0}^{1} \langle \varphi_i | \rho \varphi_i \rangle \, |\varphi_i \otimes \varphi_i\rangle \langle \varphi_i \otimes \varphi_i| \, .$$

This channel broadcasts the observable $\mathsf{A}(x)$, since for a general state in the qubit space $\rho = \left(\begin{smallmatrix} a & b \\ b^* & 1-a \end{smallmatrix}\right)$, where $a, b \in \mathbb{C}$ we get

$$\begin{aligned}
\operatorname{tr}[\rho\mathsf{A}(X)] &= \operatorname{tr}\left[\left(\begin{smallmatrix} a & b \\ b^* & 1-a \end{smallmatrix}\right)\left(\begin{smallmatrix} \alpha_0 & 0 \\ 0 & \alpha_1 \end{smallmatrix}\right)\right] \\
&= \operatorname{tr}\left[\left(\begin{smallmatrix} \alpha_0 a & \alpha_1 b \\ \alpha_0 b^* & \alpha_1(1-a) \end{smallmatrix}\right)\right] \\
&= \alpha_0 a + \alpha_1(1-a), \\
\operatorname{tr}[\mathcal{V}(\rho)\mathsf{A}(X)\otimes\mathbb{1}] &= \operatorname{tr}\left[\left(\begin{smallmatrix} a & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1-a \end{smallmatrix}\right)\left(\begin{smallmatrix} \alpha_0 & 0 & 0 & 0 \\ 0 & \alpha_0 & 0 & 0 \\ 0 & 0 & \alpha_1 & 0 \\ 0 & 0 & 0 & \alpha_1 \end{smallmatrix}\right)\right] \\
&= \operatorname{tr}\left[\left(\begin{smallmatrix} \alpha_0 a & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \alpha_1(1-a) \end{smallmatrix}\right)\right] \\
&= \alpha_0 a + \alpha_1(1-a), \\
\operatorname{tr}[\mathcal{V}(\rho)\mathbb{1}\otimes\mathsf{A}(X)] &= \operatorname{tr}\left[\left(\begin{smallmatrix} a & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1-a \end{smallmatrix}\right)\left(\begin{smallmatrix} \alpha_0 & 0 & 0 & 0 \\ 0 & \alpha_1 & 0 & 0 \\ 0 & 0 & \alpha_0 & 0 \\ 0 & 0 & 0 & \alpha_1 \end{smallmatrix}\right)\right] \\
&= \alpha_0 a + \alpha_1(1-a) \\
\Rightarrow \operatorname{tr}[\rho\mathsf{A}(X)] &= \operatorname{tr}[\mathcal{V}(\rho)\mathsf{A}(X)\otimes\mathbb{1}] = \operatorname{tr}[\mathcal{V}(\rho)\mathbb{1}\otimes\mathsf{A}(X)],
\end{aligned}$$

hence the channel $\mathcal{V}$ broadcasts the observable A.

Let us consider a pair of broadcastable observable A and B,

$$\begin{aligned}
\operatorname{tr}[\rho\mathsf{A}(x)] &= \operatorname{tr}[\mathcal{V}(\rho)\mathsf{A}(x)\otimes\mathbb{1}] = \operatorname{tr}[\mathcal{V}(\rho)\mathbb{1}\otimes\mathsf{A}(x)], \\
\operatorname{tr}[\rho\mathsf{B}(y)] &= \operatorname{tr}[\mathcal{V}(\rho)\mathsf{B}(y)\otimes\mathbb{1}] = \operatorname{tr}[\mathcal{V}(\rho)\mathbb{1}\otimes\mathsf{B}(y)],
\end{aligned}$$

for all input states $\rho \in \mathcal{S}(\mathcal{H})$ and outcomes x and y. The interpretation of this situation is that firstly, two copies of the initial state ρ is made and both are sent to different recipients. These two recipients can then independently decide whether they want to measure A or B on the received copies of ρ. The broadcastibility conditions make sure that the outcomes for this measurement scheme are equal to separately performed measurements of A and B to ρ.

A weaker condition for A and B is for them to be one-side broadcastable, that is

$$\operatorname{tr}[\rho\mathsf{A}(x)] = \operatorname{tr}[\mathcal{V}(\rho)\mathsf{A}(x)\otimes\mathbb{1}], \quad \operatorname{tr}[\rho\mathsf{B}(y)] = \operatorname{tr}[\mathcal{V}(\rho)\mathbb{1}\otimes\mathsf{B}(y)],$$

for all input states and and outcomes. Naturally, two broadcastable observables are also one-side broadcastable. But the special feature that follows from this is that a

pair of broadcastable observables is also compatible. They have a joint observable M defined as

$$\mathsf{M}(x, y) = \mathcal{V}^*(\mathsf{A}(x) \otimes \mathsf{B}(y)).$$

If $\mathcal{H} = \mathbb{C}^2$, then one-side broadcastibility implies more features on the qubit observables.

Theorem 3.2. Let A and B be qubit observables. The following conditions for A and B are equivalent:

i) A and B are one-side broadcastable

ii) A and B are mutually non-disturbing

iii) A and B are non-disturbing

iv) A and B are mutually commuting

We omit the proof of this theorem to[6]. To briefly discuss the relation of these claims, from the hierarchy of each it follows that i) $\Rightarrow$ ii) $\Rightarrow$ iii) $\Leftrightarrow$ iv) and the to complete the proof it's enough to show that iv) $\Rightarrow$ i)[6].

The properties of quantum measurements defined in this chapter are closely related. An observable's compatibility with a channel is used whenever an instrument is needed to model the measurement. Non-disturbance leads to joint measurability, which can later be modelled as a sequential measurement. Universality property then again is a very useful feature of a sequential measurement scheme. As seen in theorem (3.2), one-side broadcastability is an equivalent condition with non-disturbance and mutual non-disturbance.

4 Applications of Sequential Measurements

The purpose of this chapter is to study what actually happens, when sequential measurements are performed on quantum states instead of just defining these measurements. There are multiple different scenarios to study, which give different amounts

of information on the quantum states. There are repeatable measurements, which can be performed a number of times and it's important to study how many times different measurements can be performed and still gain more information on the initial state.

4.1 Saturation of Repeated Measurements

This discussion considering the saturation of repeated measurements follows mostly the study performed in[5]. As was mentioned when defining repeatable measurements, an observable $\mathsf{A}_n^{\mathcal{I}}$ corresponding to the instrument $\mathcal{I}$ repeated n times is defined by the equation

$$\mathsf{A}_n^{\mathcal{I}}(x_1, \ldots, x_n) = \mathcal{I}_{x_1} \circ \cdots \circ \mathcal{I}_{x_n}(\mathbb{1}).$$

By adding one more measurement of A by using instrument $\mathcal{I}$, we get

$$\begin{aligned}\mathsf{A}_{n+1}^{\mathcal{I}}(x_1, \ldots, x_{n+1}) &= \mathcal{I}_{x_1} \circ \cdots \circ \mathcal{I}_{x_{n+1}}(\mathbb{1}) \\ &= \mathcal{I}_{x_1}(\mathsf{A}_n^{\mathcal{I}}(x_2, \ldots, x_{n+1})).\end{aligned}$$

Combining this with the fact that $\sum_x \mathcal{I}_x(\mathbb{1}) = \mathbb{1}$ we are left with the equation

$$\sum_{x_{n+1}} \mathsf{A}_{n+1}^{\mathcal{I}}(x_1, \ldots, x_{n+1}) = \mathsf{A}_n^{\mathcal{I}}(x_1, \ldots, x_n).$$

By the definition of post-processing, this equation leads to $\mathsf{A}_n \preceq_{post} \mathsf{A}_{n+1}$. By extending this to all indexes n, we get the sequence

$$\mathsf{A}_1^{\mathcal{I}} \preceq_{post} \mathsf{A}_2^{\mathcal{I}} \preceq_{post} \mathsf{A}_3^{\mathcal{I}} \preceq_{post} \cdots ,$$

which means that by repeating the same measurement over and over again might increase the amount of information learned about the initial state. The amount of times a meaningful measurement like this can be performed is defined as the saturation step of the instrument.

Definition 4.1. Let $\{\mathsf{A}_n^{\mathcal{I}}\}$ be a sequence of observables related to the instrument $\mathcal{I}$, deifined by the equation

$$\mathsf{A}_n^{\mathcal{I}}(x_1, \ldots, x_n) = \mathcal{I}_{x_1} \circ \cdots \circ \mathcal{I}_{x_n}(\mathbb{1}).$$

The saturation step $\mathfrak{s}(\mathcal{I})$ for the instrument is defined as the smallest positive integer n, for which $\mathsf{A}_n^{\mathcal{I}} \simeq_{post} \mathsf{A}_{n+1}^{\mathcal{I}}$. In the case that $\mathsf{A}_n^{\mathcal{I}} \prec_{post} \mathsf{A}_{n+1}^{\mathcal{I}}, \forall n \in \mathbb{N}$, we can denote $\mathfrak{s}(\mathcal{I}) = \infty$.

Now there are multiple different scenarios to observe. We'll consider cases where the instrument saturates after the first step, instrument that saturate after some finite number of steps $n > 1$ and instruments that never saturate.

As was stated in section 2.3, an instrument is said to be repeatable, if the second repetition of the instrument gives the same outcome as the first step. So it's clear that a repeatable instrument wouldn't give any additional information when the measurement is done again, and a repeatable measurement saturates after the initial step.

Getting different measurement outcomes x_1 and x_2 from a repeatable measurement on any state can't happen, which means that for all states ρ and if $x_1 \neq x_2$

$$\begin{aligned} &\mathrm{tr}[\mathcal{I}_{x_1} \circ \mathcal{I}_{x_2}(\mathbb{1})\rho] = 0, \\ \Rightarrow &\mathcal{I}_{x_1} \circ \mathcal{I}_{x_2}(\mathbb{1}) = \delta_{x_1 x_2} \mathcal{I}_{x_1}(\mathbb{1}). \end{aligned}$$

This can be written using the corresponding observables

$$\begin{aligned} &\mathsf{A}_2^{\mathcal{I}}(x_1, x_2) = \delta_{x_1 x_2} \mathsf{A}^{\mathcal{I}}(x_1) = \sum_x \delta_{x x_1} \delta_{x_1 x_2} \mathsf{A}^{\mathcal{I}}(x) \\ \Rightarrow &\mathsf{A}_2^{\mathcal{I}} \simeq_{post} \mathsf{A}^{\mathcal{I}}. \end{aligned}$$

So for an repeatable instrument $\mathcal{I}$ the saturation step is $\mathfrak{s}(\mathbb{1}) = 1$ just as was expected, since a repeatable measurement shouldn't yield any additional information after the initial measurement.

Repeating a repeatable measurement isn't a very interesting subject based on the definition itself, but rather we are interested in gaining more information by repeating a measurement. To consider instruments, which saturate at a finite step $\mathfrak{s}(\mathcal{I}) > 1$, let $\mathcal{H}$ be a Hilbert space with some finite dimension $\dim(\mathcal{H}) = d \geq 2$ and

let $\{\varphi_i\}$ be an orthonormal basis of $\mathcal{H}$. Let $\mathcal{I}$ be an instrument defined by

$$\Omega = \{0, 1\},$$
$$\mathcal{I}_x(T) := L_x^* T L_x,$$
$$L_0 := |\varphi_d\rangle\langle\varphi_d|, \quad L_1 \sum_{i=1}^{d-1} |\varphi_{i+1}\rangle\langle\varphi_i|.$$

For this instrument we have the following proposition.

Proposition 4.1. Let $\mathsf{A}_n^{\mathcal{I}}$ be the observable corresponding to the instrument $\mathcal{I}$. Then the following applies.

a) For $1 \leq n \leq d-1$, then $\mathsf{A}_n^{\mathcal{I}} \simeq_{post} \mathsf{P}_n$ where the sharp observable P_n is defined as $\mathsf{P}_n(j) \begin{cases} \sum_{l=1}^{d-n} |\varphi_l\rangle\langle\varphi_l|, & \text{when } j = 1 \\ |\varphi_{d-n+j-1}\rangle\langle\varphi_{d-n+j-1}|, & \text{when } 2 \leq j \leq n+1. \end{cases}$

b) For $d-1 \leq n$, then $\mathsf{A}_n^{\mathcal{I}} \simeq_{post} \mathsf{P}_{d-1}$, where P_{d-1} is defined by

$$\mathsf{P}_{d-1} := |\varphi_j\rangle\langle\varphi_j|, \quad 1 \leq j \leq n.$$

Proof

a) Assume that $1 \leq n \leq d-1$ holds. The operators of the form $L_{x_n} L_{x_{n-1}} \cdots L_{x_1}$ are

$$L_1^n = \sum_{l=1}^{d-n} |\varphi_{l+n}\rangle\langle\varphi_l|,$$
$$L_0^j L_1^{n-j} = |\varphi_{d-n+j}\rangle\langle\varphi_{d-n+j}|, \quad 1 \leq j \leq n.$$

From $L_1 L_0$ it follows that all operators of this form are zero. The non-zero observables corresponding to multiple measurements $\mathsf{A}(x_1, \ldots, x_n)$ are

$$\mathsf{A}_n(1, \ldots, 1) = \sum_{l=1}^{d-n} |\varphi_l\rangle\langle\varphi_l|,$$
$$\mathsf{A}_n(1, \ldots, 1, 0, \ldots, 0) = |\varphi_{d-n+j}\rangle\langle\varphi_{d-n+j}|, \quad 1 \leq j \leq n,$$

where $\mathsf{A}_n(1, \ldots, 1, 0, \ldots, 0)$ means outcome 1 is given $n-j$ times and outcome 0 is given j times. From this it follows that $\mathsf{A}_d \simeq_{post} \mathsf{P}_d$.

b) Assume that $d-1 \leq n$. Then we need show that $\mathsf{A}_{d-1} \simeq_{post} \mathsf{A}_d$. The from a) we know that $\mathsf{A}_{d-1} \simeq_{post} \mathsf{P}_{d-1}$ and hence the observable A_d is

$$\mathcal{I}_x(\mathsf{P}_{d-1}(j)), \quad x \in \Omega, 1 \leq j \leq d.$$

The non-zero elements of the observable are of the form

$$\mathcal{I}_0(\mathsf{P}_{d-1}(d)) = \mathsf{P}_{d-1}(d), \qquad \mathcal{I}_1(\mathsf{P}_{d-1}(d)) = \mathsf{P}_{d-1}(j-1), \quad 2 \leq j \leq d.$$

It follows that $\mathsf{A}_d \simeq_{post} \mathsf{P}_{d-1} \simeq_{post} \mathsf{A}_{d-1}$, and thus the proof holds. □

By the definition of the observables P_i we can see that

$$\mathsf{P}_1 \prec_{post} \mathsf{P}_2 \prec_{post} \cdots \prec_{post} \mathsf{P}_{d-2} \prec_{post} \mathsf{P}_{d-1}$$

holds. Thus using the previous proposition we can conclude that

$$\begin{aligned} &\mathsf{A}_1^{\mathcal{I}} \prec_{post} \mathsf{A}_2^{\mathcal{I}} \prec_{post} \cdots \prec_{post} \mathsf{A}_{d-2}^{\mathcal{I}} \prec_{post} \mathsf{A}_{d-1}^{\mathcal{I}} \simeq_{post} \mathsf{A}_d^{\mathcal{I}} \simeq_{post} \cdots \\ \Rightarrow &\mathfrak{s}(\mathcal{I}) = d-1. \end{aligned}$$

This clearly means that the constructed measurement yields more information after each step until the limit $d-1$ is reached. Yet, this doesn't need to be the limit and there still exists measurements which provide more information no matter how many times a measurement is made.

The final thing with respect to repeating a measurement is to showcase an instrument, which doesn't saturate at all. Let A and B be POVM's and let the related outcome spaces $(\Omega_{\mathsf{A}}, \mathcal{F}_{\mathsf{A}})$ and $(\Omega_{\mathsf{B}}, \mathcal{F}_{\mathsf{B}})$ be of the form of the standard Borel σ-algebra, that is $\mathcal{F}_{\mathsf{A}} = \mathfrak{B}(\Omega_{\mathsf{A}})$ and $\mathcal{F}_{\mathsf{B}} = \mathfrak{B}(\Omega_{\mathsf{A}})$. For $\mathsf{A} \preceq_{post} \mathsf{B}$ to hold, there needs to exist a Markov kernel

$$\kappa : \mathcal{F}_{\mathsf{A}} \times \Omega_{\mathsf{B}} \to [0,1]$$

where $\kappa(\cdot, y)$ is a probability measure on $(\Omega_{\mathsf{A}}, \mathcal{F}_{\mathsf{A}})$ for every $y \in \Omega_{\mathsf{B}}$. Furthermore, $\kappa(X, \cdot)$ needs to be $\mathcal{F}_{\mathsf{B}}$-measurable for every $x \in \mathcal{F}_{\mathsf{A}}$ and

$$\mathsf{A}(X) = \int_{\Omega_{\mathsf{B}}} \kappa(x,y) \mathrm{d}\mathsf{B}(y), \quad \forall x \in \Omega_{\mathsf{A}}.$$

Let $\mathcal{I}$ be an instrument on a finite set Ω and let $\mathsf{A}_\infty^{\mathcal{I}}$ be an observable with the outcome space $(\Omega^\infty, \mathcal{F}^\infty)$. $\mathsf{A}_\infty^{\mathcal{I}}$ is called the infinite composition of the instrument and it's defined by

$$\mathsf{A}_\infty^{\mathcal{I}}(\{x_1\} \times \cdots \times \{x_n\} \times \Omega^\infty) = \mathsf{A}_n^{\mathcal{I}}(x_1, \ldots, x_n), \quad \forall n \geq 1, (x_1, \ldots, x_n) \in \Omega^n,$$

where $\mathsf{A}_n^{\mathcal{I}}$ denotes the n repetitions of $\mathcal{I}$ and the σ-algebra $\mathcal{F}^\infty$ is generated by the so called cylindrical sets $\{x_1\} \times \cdots \times \{x_n\} \times \Omega^\infty$.

The existence and uniqueness of $\mathsf{A}_\infty^{\mathcal{I}}$ is not something that can be taken for granted, but something that needs to be proved. For this we need the following theorem called the quantum Kolmogorov extension theorem[11][17].

Theorem 4.1. Let $(\Omega_i, \mathcal{F}_i)$ be the outcome space with the regular Borel set where $(i = 1, 2, \ldots)$ and let A_n be a POVM with the outcome space $(\Pi_{i=1}^n \Omega_i, \Pi_{i=1}^n \mathcal{F}_i)$, where $(n = 1, 2, \ldots)$. If the set of observables $\{\mathsf{A}_n\}$ satisifies

$$\mathsf{A}_n(x) = \mathsf{A}_{n+1}(x \times \Omega_{n+1}), \quad n \geq 1, \quad x \in \Pi_{i=1}^n \Omega_i, \tag{17}$$

then there exists a unique POVM A_∞ equipped with the infinite outcome space $(\Pi_{i=1}^\infty \Omega_i, \Pi_{i=1}^\infty \mathcal{F}_i) = (\Omega^\infty, \mathcal{F}^\infty)$ such that

$$\mathsf{A}_n(x) = \mathsf{A}_\infty(x \times \Pi_{i=n+1}^\infty \Omega_i), \quad n \geq 1, \quad x \in \Omega^n.$$

Furthermore, each unit vector $\psi \in \mathcal{H}$ corresponds to a unique probability measure μ^ψ on the outcome set $(\Omega^\infty, \mathcal{F}^\infty)$, for which

$$\mu^\psi(x \times \Omega^\infty) = \langle \psi | \mathsf{A}_n(x) \psi \rangle, \quad \forall n \in \mathbb{N}, x \in \Omega^n.$$

Proof Let $\psi \in \mathcal{H}$ be an arbitrary unit vector defining the probability measure

$$\mu_K^\psi(\cdot) = \langle \psi | \mathsf{A}_K(\cdot) \psi \rangle,$$

where

$$\mathsf{A}_K(x) = \mathsf{A}_n(x \times \Pi_{i \in \{1, \ldots, n\}} \Omega_i))$$

and K is a subset of $\{1, \ldots, n\}$.

To show that the probability measure μ^ψ corresponds to some POVM A as $\mu^\psi(\cdot) = \langle \psi | \mathsf{A}(\cdot) \psi \rangle$ we'll define a complex number

$$\tilde{\mu}^{\psi}(x) \begin{cases} \|\psi\|^2 \mu^{\psi/\|\psi\|}(x), & \text{if } \psi \neq 0 \\ 0 & \text{if } \psi = 0 \end{cases}$$

which means that ψ doesn't necessarily need to be a unit vector. We can further define for any pair of vectors $\psi, \phi \in \mathcal{H}$ a complex number $\mu_{\psi,\phi}(x)$ as

$$\mu_{\psi,\phi}(x) := \tilde{\mu}^{\frac{1}{2}\psi+\frac{1}{2}\phi}(x) - \tilde{\mu}^{\frac{1}{2}\psi-\frac{1}{2}\phi}(x) + i\tilde{\mu}^{\frac{1}{2}\psi-\frac{i}{2}\phi}(x) - i\tilde{\mu}^{\frac{1}{2}\psi+\frac{i}{2}\phi}(x).$$

This new complex number is defined via the so called polarization of $\tilde{\mu}$. By using these complex numbers, for a cylindrical set x

$$x = x' \times \Omega^{\infty}$$

we get

$$\tilde{\mu}^{\psi}(x) = \langle \psi | \mathsf{A}_n(x)\psi \rangle \,, \quad \mu_{\psi,\phi}(x) = \langle \psi | \mathsf{A}_n(x)\psi \rangle \,.$$

Being a complex linear combination of finite measures, $\mu_{\psi,\phi}$ is now a complex measure on $(\Omega^{\infty}, \mathcal{F}^{\infty})$. As mentioned before, these cylindrical sets work as a generator of the σ-algebra $\mathcal{F}^{\infty}$ and therefore the two complex measures we defined coincide and the following equations hold.

$$\begin{aligned} &\mu_{\psi+\psi',\phi}(x) = \mu_{\psi,\phi}(x) + \mu_{\psi',\phi}(x), \quad \forall \psi, \psi', \phi \in \mathcal{H}, \\ &\mu_{c\psi,\phi}(x) = c^* \mu_{\psi,\phi}(x), \quad \forall c \in \mathbb{C}^2, \forall \psi, \phi \in \mathcal{H}, \\ &\mu_{\phi,\psi}(x) = \mu_{\psi,\phi}(x)^*, \quad \forall \psi, \phi \in \mathcal{H} \end{aligned}$$

for all cylindrical sets x and $\forall x \in \Omega^{\infty}$. Now, for any fixed set $x \in \Omega^{\infty}$ the mapping $(\psi, \phi) \mapsto \mu_{\psi,\phi}(x)$ is a Hermitian sesquilinear form on the Hilbert space $\mathcal{H}$. This form is bounded and has norm $\|\mu_{\psi,\phi}(x)\| \leq 1$, since

$$\mu_{\psi,\psi}(x) = \tilde{\mu}^{\psi}(x) = \|\psi\|^2 \mu^{\psi/\|\psi\|}(x) \leq \|\psi\|^2 \,.$$

From Riesz's lemma[14] it follows that there exists a bounded operator A_n, for which

$$\mu_{\psi,\phi}(x) = \langle \psi | \mathsf{A}_n(x)\phi \rangle \,.$$

Now that the existence of A_n has been shown, we need to verify it's uniqueness. If there exists two observables $\mathsf{B}_n(x) \neq \mathsf{A}_n(x)$, then there exists a unit vector $\psi \in \mathcal{H}$, for

which $\langle\psi|\mathsf{B}_n(x)\psi\rangle \neq \langle\psi|\mathsf{A}_n(x)\psi\rangle$ which contradicts the fact that $\tilde{\mu}^{\psi}(x) = \mu_{\psi,\psi}(x) = \langle\psi|A_n(x)\psi\rangle$.

What we now need to do is to verify that A_n is a POVM. Since

$$\mu_{\psi,\psi}(x) = \tilde{\mu}(x) \geq 0$$

A_n is positive for each $\psi \in \mathcal{H}$. It is also σ-additive with respect to the weak operator topology, since the measure $\mu_{\psi,\phi}(\cdot)$ is σ-additive. When restricted to the defined cylindrical sets we get $\mathsf{A}_n(\Omega^\infty) = \mathbb{1}$. Hence, A_n is a POVM.

Only thing remaining now is to show that A_K is a marginal of A. For a cylindrical set x we have

$$\langle\psi|\mathsf{A}_n(x)\psi\rangle = \tilde{\mu}^{\psi}(x) = \|\psi\|^2\, \mu^{\psi/\|\psi\|}(x) = \|\psi\|^2\, \mu_K^{\psi/\|\psi\|}(x_K) = \langle\psi|\mathsf{A}_K(x_K)\psi\rangle\,,$$

where $x_K \in \mathcal{F}^K$ and $\psi \in \mathcal{H} \setminus \{0\}$. Since a positive bounded operator L by the numbers defined as $\langle\psi|L\psi\rangle$, we have $\mathsf{A}_n(x) = A_K(x_K)$. $\square$

Using the Kolmogorov Consistency condition (17), an infinite composition of an instrument can finally be constructed. Let $\mathcal{I}$ be a completely positive instrument with the regular Borel outcome space. A_n is the observable corresponding to n repetitions of the instrument $\mathcal{I}$. Now it's straightforward to see that for each $x \in \Omega^n$ we have

$$\begin{aligned}
\mathsf{A}_{n+1}(x \times \Omega) &= (\mathcal{I}^n \circ \mathcal{I})_{x\times\Omega}(\mathbb{1}) \\
&= \mathcal{I}_x^n(\mathcal{I}_\Omega(\mathbb{1})) \\
&= \mathcal{I}_x^n(\mathbb{1}) \\
&= \mathsf{A}_n(x).
\end{aligned}$$

Now there exists an observable A_∞ defined as

$$\mathsf{A}_\infty(\Pi_{i=1}^n x_i \times \Omega^\infty) = \mathcal{I}_{x_1} \circ \cdots \circ \mathcal{I}_{x_2}(\mathbb{1}),$$

which is the infinite composition of the instrument $\mathcal{I}$. From the definition of the infinite composition we can see that

$$\mathsf{A}_n \preceq_{post} \mathsf{A}_\infty, \quad \forall n \geq 1$$

and thus, we have constructed an insaturable observable. This shows that it's possible to perform sequential measurements, where with each step more information on the initial state is gained.

4.2 Approximate Repeatability

Repeatability of an instrument has been discussed a couple of times throughout this thesis, but the set of observables related to a repeatable instrument is very limited. This chapter considering approximate repeatability follows mostly the study performed in [3].

First of all, only discrete observables can be repeatable. Secondly, no additional information on a second measurement of a repeatable instrument is gained, since the output is the same as in the initial measurement. To make things more interesting, we discuss the approximate repeatability of an instrument and the related observable. Let A be an observable and the associated outcome space $(\mathbb{R}, \mathcal{F})$ be the Borel σ-algebra of real numbers, that is $\mathcal{F} = \mathfrak{B}(\mathbb{R})$. Furthermore, we need the following few definitions before introducing the concept of approximate repeatability.

Definition 4.2. An effect $\mathsf{A}(X)$ is said to be actual in a state ρ, if

$$\mathrm{tr}[\rho \mathsf{A}(X)] = 1$$

If an observable is actual in some state ρ, the observable is said to be actualizable.

An example of an actualizable effect would be any non-zero projection. The set of effects being actualizable is very limited, so next we'll need a sort of looser definition of actualizability.

Definition 4.3. Let $\frac{1}{2} \leq c < 1$. An effect $\mathsf{A}(X)$ is c-actual in a state ρ if

$$\mathrm{tr}[\rho \mathsf{A}(X)] > c.$$

Furthermore, $\mathsf{A}(X)$ is said to be c-actualizable. If $\mathsf{A}(X)$ is c-actualizable for every c, then it is said to be almost actualizable.

If $\mathsf{A}(X)$ is actualizable for a fixed value of c, then for some outcome Y for which $X \subseteq Y$ and of course $\mathsf{A}(X) \leq \mathsf{A}(Y)$, then $\mathsf{A}(Y)$ is also c-actualizable. Let $I_{x,r} = (x - r/2, x + r/2)$, $\forall x \in \mathbb{R}, r > 0$. Using $I_{x,r}$ we can define the minimal width r of an interval, where the effect $\mathsf{A}(I_{x,r})$ is c-actualizable.

Definition 4.4. For $\frac{1}{2} \leq c < 1$, the resolution width $\gamma(\mathsf{A}, c)$ is defined as

$$\gamma(\mathsf{A}, c) := \inf\{r > 0 | \mathsf{A}(I_{x,r}) \text{ is } c\text{-actualizable} \, \forall x \in \mathbb{R}\}$$

and c is said to be the confidence level of A. Furthermore, by letting c approach one, we denote

$$\gamma(\mathsf{A}, 1) := \lim_{c \to 1-} \gamma(\mathsf{A}, c)$$

that is, for A equipped with confidence level 1, $\gamma(\mathsf{A}, 1)$ is the resolution width.

To apply repeatable measurements on continuous observables, we need to provide an alternative definition on what repeatability means, using the definition of actualizability. For $\varepsilon > 0$ we denote

$$X_\varepsilon := \bigcup_{x \in X} I_{x,\varepsilon} = \left\{ y \in \mathbb{R} | \, |x - y| < \frac{\varepsilon}{2}, \quad \exists x \in X \subseteq \mathbb{R} \right\},$$

and use this to define a new type of repeatability.

Definition 4.5. For an instrument $\mathcal{I}$, $\varepsilon > 0$ and $\frac{1}{2} \leq c < 1$

1) $\mathcal{I}$ is ε-repeatable if

$$\mathrm{tr}[\mathcal{I}_{X_\varepsilon}(\mathcal{I}_X(\rho))] = \mathrm{tr}[\mathcal{I}_X(\rho)], \quad \forall \rho \in \mathcal{S}(\mathcal{H}), X \in \mathcal{F}$$

2) $\mathcal{I}$ is (ε, c)-repeatable if

$$\begin{aligned} \mathrm{tr}[\mathcal{I}_X(\rho)] &\neq 0, \\ \mathrm{tr}[\mathcal{I}_{X_\varepsilon}(\mathcal{I}_X(\rho))] &> c \cdot \mathrm{tr}[\mathcal{I}_X(\rho)], \quad \forall \rho \in \mathcal{S}(\mathcal{H}), X \in \mathcal{F} \end{aligned}$$

With these less strict definitions of repeatability, repeating the measurements of continuous observables can be made. It has also been proved that if a POVM A

corresponding to an interval is actualizable, there exists an ε-repeatable instrument $\mathcal{I}$ associated with the POVM A for every $\varepsilon > 0$[4].

To provide an example of an ε-repeatable instrument, we consider the Hilbert space of square integrable functions on $\mathbb{R}$, $\mathcal{H} = L^2(\mathbb{R})$, where we define the following observable

$$\mathsf{Q}(X)\psi(x) = \chi_X(x)\psi(x),$$

called the canonical position observable, with the outcome space $(\mathbb{R}, \mathfrak{B}(\mathbb{R}))$[10]. Here χ_X denotes the characteristic function of the set X. Furthermore, for a probability measure μ on $\mathbb{R}$ we define another observable

$$\mathsf{Q}_\mu(X) = \int \mu(X - x)\mathrm{d}\mathsf{Q}(x), \quad X \in \mathfrak{B}(\mathbb{R})$$

called a position observable. The resolution width of this Q_μ is

$$\gamma(\mathsf{Q}_\mu, c) = \inf\{r > 0 \,|\, \text{ess sup}_{x\in\mathbb{R}}\mu(I_{x,r}) > c\},$$

which is a finite number. The denotion ess sup means the essential supremum of a real valued function with respect to the Lebesque measure.

Proposition 4.2. For a position observable Q_μ and $\frac{1}{2} \leq c < 1$ each $\epsilon > \gamma(\mathsf{Q}_\mu, c)$ there exists a Q_μ-compatible and (ϵ, c)-repeatable instrument of the form

$$\mathcal{I}_X(\rho) = \int_X U_x\rho_0 U_x^* \,\mathrm{tr}[\rho\mathsf{Q}_\mu(\mathrm{d}x)], \quad X \in \mathfrak{B}(\mathbb{R}),\ \rho_0, \rho \in \mathcal{S}(\mathcal{H}), \tag{18}$$

where the unitary operator U_x corresponds to a position shift.

Proof From $\varepsilon > \gamma(\mathsf{Q}_\mu, c)$ it follows that

$$\mathrm{tr}[\rho_0\mathsf{Q}_\mu(I_{0,\epsilon})] > c, \quad \exists\rho_0 \in \mathcal{S}(\mathcal{H}).$$

The state ρ_0 generates an instrument $\mathcal{I}$. Now, for any state $\rho \in \mathcal{S}(\mathcal{H})$ and outcome $X \in \mathfrak{B}(\mathbb{R})$

$$\begin{aligned}
\mathrm{tr}[\mathcal{I}_{X_\varepsilon}(\mathcal{I}_X(\rho))] &= \mathrm{tr}[\mathcal{I}_X(\rho)\mathsf{Q}_\mu(X_\varepsilon)] \\
&= \int_X \mathrm{tr}[U_x\rho_0 U_x^*\mathsf{Q}_\mu(X_\varepsilon)]\,\mathrm{tr}[\rho\mathsf{Q}_\mu(\mathrm{d}x)] \\
&= \int_X \mathrm{tr}[\rho_0\mathsf{Q}_\mu(X_\varepsilon - x)]\,\mathrm{tr}[\rho\mathsf{Q}_\mu(\mathrm{d}x)].
\end{aligned}$$

And it follows that

$$I_{0,\varepsilon} \subset X_\varepsilon - x, \quad \forall x \in X$$
$$\Rightarrow \mathrm{tr}[\rho_0 \mathsf{Q}_\mu(X_\varepsilon - x)] > c.$$

And hence, requiring $\mathrm{tr}[\rho \mathsf{Q}_\mu(X)] \neq 0$, we get

$$\mathrm{tr}[\mathcal{I}_{X_\varepsilon}(\mathcal{I}_X(\rho))] > \int_X c \cdot \mathrm{tr}[\rho \mathsf{Q}_\mu(\mathrm{d}x)] = c \cdot \mathrm{tr}[\mathcal{I}_X(\rho)],$$

which is what we wanted to prove. □

So in conclusion, there are actual real measurements for continuous observables, where an altered definition of repeatable measurement can be applied and more information on the initial state can possibly be gained.

References

[1] P. Busch, P. Lahti, J-P. Pellonpää, and K. Ylinen. *Quantum Measurement.* Springer, 2016.

[2] P. Busch and J. Singh. Lüders theorem for unsharp quantum measurements. *Phys. Lett. A*, 249:10–12, 11 1998.

[3] C. Carmeli, T. Heinonen, and A. Toigo. Intrinsic unsharpness and approximate repeatability of quantum measurements. *J. Phys. A*, 40(6), 2007.

[4] E. B. Davies and J. T. Lewis. An operational approach to quantum probability. *Comm. Math. Phys.*, 17(3):239–260, 1970.

[5] E. Haapasalo, T. Heinosaari, and Y. Kuramochi. Saturation of repeated quantum measurements. *J. Phys. A*, 49(33), 2016.

[6] T. Heinosaari. Simultaneous measurement of two quantum observables: Compatibility, broadcasting, and in-between. *Phys. Rev. A*, 93, Apr 2016.

[7] T. Heinosaari and T. Miyadera. Universality of sequential quantum measurements. *Phys. Rev. A.*, 91(2), 2015.

[8] T. Heinosaari, D. Reitzner, and P. Stano. Notes on joint measurability of quantum observables. *Found, Phys.*, 38(12):1133–1147, Dec 2008.

[9] T. Heinosaari and M. Wolf. Non-disturbing quantum measurements. *J. Math. Phys*, 51(9):1–20, 2010.

[10] T. Heinosaari and M. Ziman. *The Mathematical Language of Quantum Theory: From Uncertainty to Entanglement.* Cambridge University Press, 2011.

[11] Y. Kuramochi. Construction of the least informative observable conserved by a given quantum instrument. *J. Math. Phys.*, 56, 05 2015.

[12] M. Ozawa. Quantum measuring processes of continuous observables. *J. Math. Phys*, 25(1):79–87, 1984.

[13] V. Paulsen. *Completely Bounded Maps and Operator Algebras.* Cambridge University Press, 2003.

[14] M. Reed and B. Simon. *Methods of modern mathematical physics. 1, Functional analysis.* New York, London: Academic Press, 1972.

[15] J. Renes, R. Blume-Kohout, A. Scott, and C. Caves. Symmetric informationally complete quantum measurements. *J. Math. Phys.*, 45, 2004.

[16] W. Rudin. *Real and Complex Analysis, 3rd Ed.* McGraw-Hill, Inc., New York, NY, USA, 1987.

[17] R. Tumulka. A Kolmogorov extension theorem for POVMs. *Lett. Math. Phys.*, 84(1):41–46, Apr 2008.