Czech Technical University in Prague
Faculty of Nuclear Sciences and Physical Engineering

Department of Physics

Collective measurement and system disturbance

Master's thesis

Author: Bc. Elisabeth Andriantsarazo
Supervisor: Ing. Václav Potoček, Ph.D.
Academic year: 2019/2020

Prohlášení

Prohlašuji, že jsem svou diplomovou práci vypracovala samostatně a použila jsem pouze podklady (literaturu, projekty) uvedené v přiloženém seznamu.

Nemám závažný důvod proti použití tohoto školního díla ve smyslu § 60 Zákona č. 121/200 Sb., o právu autorském, o právech souvisejících s právem autorským a o změně některých zákonů (autorský zákon).

V Praze dne 7. 1. 2020

....
Bc. Elisabeth Andriantsarazo

Acknowledgement

I would like to express my very great appreciation to my supervisor, Ing. Václav Potoček, Ph.D., for his kind guidance and for all his help in what was not an easy task for me. His encouragement and dedication really inspired me and his suggestions were incredibly valuable during the writing process. I would also like to thank my friends and family for their unending support. I am especially thankful to my partner and my son, who provided me with their never ending encouragement.

Bc. Elisabeth Andriantsarazo

Název:
Kolektivní měření a porucha systému

Autorka: Bc. Elisabeth Andriantsarazo

Druh práce: Diplomová práce

Vedoucí práce: Ing. Václav Potoček, Ph.D.
Department of Physics, Faculty of Nuclear Sciences and Physical Engineering, Czech Technical University in Prague

Abstrakt:
V této práci bude prozkoumán vztah mezi množstvím poruchy kvantového systému a ziskem informace. Na začátku bude představena základní teorie kvantového měření a teorie informace, která byla vyvinuta Shannonem. Vlastní práce začíná ukázáním možných způsobů snížení celkové poruchy systému vhodným zadefinováním měřících operátorů na dvoučásticových systémech. Tento výsledek bude následně zobecněn v páté kapitole dalšími měřeními, jež nebudou rozlišovat mezi jednotlivými částicemi v systému. Prozkoumané metody budou poté použity k prozkoumání N částicového případu na němž ukážeme jak vhodným výběrem měřících operátorů snížit poruchu v našem systému.

Klíčová slova: Kvantová měření, šum a porucha, Heisenbergova relace neurčitosti, kvantové operace

Title:
Collective measurement and system disturbance

Author: Bc. Elisabeth Andriantsarazo

Supervisor: Ing. Václav Potoček, Ph.D.
Department of Physics, Faculty of Nuclear Sciences and Physical Engineering, Czech Technical University in Prague

Abstract:
In this work we will investigate a trade-off between the amount of quantum system disturbance and information gain. Basic theory of quantum measurement and classical information theory developed by Shannon will be presented. Our own work starts with showing possible ways of reducing the overall system disturbance by conveniently defining the measurement operators on two-particle systems. This result will be generalized by several schemes of non-distinguishable measurement presented in the fith chapter. Methods explored in these chapters will then be used to examine the case, where we have N particles in our system, on which we will show how by appropriately chosing the measurement operator set we can decrease the disturbance on our system.

Key words: Quantum measurement, noise and disturbance, Heisenberg uncertainty relation, quantum operations

Contents

Contents **IX**

Introduction **XI**

1 Generalized measurements **1**
1.1 von Neumann measurements . . . 1
1.2 Generalized measurements . . . 3
1.3 Positive operator-valued measure . . . 4
1.4 Operations . . . 7

2 Fundamentals of information theory **9**

3 Noise and disturbance in measurement **13**
3.1 Noise and disturbance using statistics . . . 13
3.2 Noise and disturbance using entropy . . . 15

4 Zero disturbance on systems with two identical particles **19**
4.1 Zero disturbance using statistical noise and disturbance . . . 19
4.2 Zero disturbance using entropy . . . 21

5 Reducing disturbance without addressing specific particles **23**
5.1 Measuring total spin . . . 23
5.2 Projection with replacement . . . 25
5.3 Calculating disturbance for both types of projection . . . 26
5.3.1 Disturbance for total spin measurement without replacement . . . 27
5.3.2 Disturbance for total spin measurement with replacement . . . 29
5.4 Weakening of the projectors . . . 31
5.4.1 Disturbance for projector weakening . . . 34
5.4.2 Classical channels as functions of μ, ν . . . 36

6 N-particle system disturbance **37**
6.1 Projection with replacement for N particles . . . 37
6.2 Projector weakening for N particles . . . 39

Conclusion **43**

Bibliography **43**

Introduction

The main aim of this thesis is to study quantum measurement protocols with respect to the concepts of noise and disturbance. When the notions of noise and disturbance are mentioned, the first thing that comes to mind is usually the famous relation from Heisenberg, which is commonly written in the following form:

$$\sigma_x \sigma_p \geq \frac{\hbar}{2}, \tag{1}$$

where σ_i denotes standard deviation related to the position and momentum measurement, respectively. The problem of noise in a measurement and disturbance of a quantum system is much more intricate though. First thing complicating the whole situation is the fact that Heisenberg never gave any physical interpretations to σ_i. The original Heisenberg relation takes the following form:

$$\Delta x \Delta p \geq h, \tag{2}$$

where $h = 2\pi\hbar$. The meaning behind equation (2) is that the measurement precision of one physical property on wave-like systems necessarily influences measurement precision of another physical property. A great example of this phenomenon can be a tone pitch and its duration. The shorter we measure certain tone pitch, the harder it is to exactly tell what its frequency is. This behaviour is natural for all systems that can be described using wave equation, which means that it must be observed in quantum systems. This principle is described in great detail in the third chapter.

Both relations (1) and (2) are not what we mean by measurement disturbance, since, as we already mentioned, the principle described above applies to all wave-like systems, not only the quantum ones. What does apply only to the quantum world is the observer effect, a concept, which tells us that every measurement necessarily influences the system being measured. This influence then causes some loss of information about the system and there is no way of completely getting rid of it. In order to quantify this mainly negative influence and the information loss it causes, we need to correctly and rigorously define joint quantum measurements. We can then introduce functions which, using information theory and statistics, give us the amount of correlation between the outcomes we measured and outcomes that can be considered as "real". Quantifying this link between the two means calculating the noise and disturbance on the system by some joint measurement.

Chapter 1

Generalized measurements

In this chapter we shall briefly explore the basics of measurement theory in quantum mechanics. In the first section we start with von Neumann measurements, where we represent a state by a vector. We then make a generalization and switch to a better way of state representation – density matrices, denoted throughout this thesis as $\hat{\rho}$, which serve as a very convenient way to describe composite systems and their evolution. We end this chapter with the theory of quantum operations and show the superiority of this notation.

For a quantum system, whose states can be, up to a global phase factor $e^{i\varphi}$, represented by a set of state vectors $|\psi_i\rangle$ with respective probabilities p_i is the density matrix, or density operator, defined as follows:

$$\hat{\rho} \equiv \sum_i p_i |\psi_i\rangle \langle\psi_i| \,. \tag{1.1}$$

The evolution of a state can then be very conveniently described using unitary operator $\hat{U}$:

$$\hat{\rho} \to \hat{U}\hat{\rho}\hat{U}^\dagger, \tag{1.2}$$

where $\hat{U}^\dagger$ represents the adjoint operator to $\hat{U}$. The density operator completely describes a quantum state and is always a positive operator with $\mathrm{Tr}(\hat{\rho}) = 1$.

1.1 von Neumann measurements

The von Neumann measurement, introduced by Hungarian-American mathematician John von Neumann, is also called ideal or projective measurement. It does not take effects of noise into account, meaning the measured system is understood as a combined system of the apparatus and the quantum state we seek to measure. Let us have an observable quantity represented by a Hermitian operator $\hat{A}$ with a set of eigenstates (now represented by vectors) $\{|\lambda_n\rangle\}$ and a set of eigenvalues $\sigma_A = \{\lambda_n\}$, where $\lambda_n \in \mathbb{C}$ and $\{|\lambda_n\rangle\}$ form an orthonormal basis on our Hilbert space $\mathcal{H}$. We start with the following equation:

$$\hat{A} |\lambda_n\rangle = \lambda_n |\lambda_n\rangle \,. \tag{1.3}$$

The set of eigenvalues $\{\lambda_n\}$ represents the measurement outcomes. An operator $\hat{A}$ with non-degenerate spectrum can be decomposed using both eigenvalues and eigenstates $\{\lambda_n\}$ as follows:

$$\hat{A} = \sum_n \lambda_n \left|\lambda_n\right\rangle \left\langle\lambda_n\right|. \tag{1.4}$$

Let us also have a state represented by a density matrix $\hat{\rho}$, then we can calculate the probability of measuring a value λ_n from the spectrum σ_A of the operator $\hat{A}$ as:

$$P(\lambda_n) = \langle\lambda_n|\hat{\rho}|\lambda_n\rangle. \tag{1.5}$$

If we assume the eigenstates can be degenerate the operator $\hat{A}$ can be decomposed as follows:

$$\hat{A} = \sum_n \sum_j^{d(n)} \lambda_n \left|\lambda_n^{(j)}\right\rangle \left\langle\lambda_n^{(j)}\right|, \tag{1.6}$$

where $d(n)$ denotes the multiplicity of eigenvalues λ_n and $\{|\lambda_n^{(j)}\rangle \,|1 \geq j \geq d(n)\}$ forms an orthonormal basis if the eigenstate $|\lambda_n\rangle$ correspond with the eigenvalue λ_n. The probability of measuring such degenerate eigenvalue (measurement outcome) is given by:

$$P(\lambda_n^{(j)}) = \sum_j \langle\lambda_n^{(j)}|\rho|\lambda_n^{(j)}\rangle \tag{1.7}$$

Here we can make an abstraction by transitioning from observables to outcomes, which will change our description of the measurement act. So instead of the measurement outcome being labeled as λ_n we shall label it as n and use a projector defined as:

$$\hat{P}_n = \left|\lambda_n\right\rangle \left\langle\lambda_n\right| = \left|n\right\rangle \left\langle n\right|, \tag{1.8}$$

or, for degenerate states, as:

$$\hat{P}_n = \sum_j \left|\lambda_n^{(j)}\right\rangle \left\langle\lambda_n^{(j)}\right| = \sum_j \hat{P}_n^{(j)}. \tag{1.9}$$

Projectors are always positive $\hat{P}_n \geq 0$ and have the following properties:

$$\begin{gathered} \sum_n \hat{P}_n = \hat{I}, \\ \hat{P}_m \hat{P}_n = \delta_{mn} \hat{P}_n, \\ \hat{P}_n^\dagger = \hat{P}_n \end{gathered} \tag{1.10}$$

where $\hat{P}_n^\dagger$ is the Hermitian conjugate of $\hat{P}_n$, $\hat{I}$ represents the identity operator on our Hilbert space and δ_{mn} is the Kronecker delta. The respective probabilities, defined previously in (1.5) and (1.7) can now be written as follows:

$$P(\lambda_n) = P(n) = \mathrm{Tr}\left(\langle n|\hat{\rho}|n\rangle\right) = \mathrm{Tr}\left(\hat{\rho}\left|n\right\rangle\left\langle n\right|\right) = \mathrm{Tr}(\hat{\rho}\hat{P}_n) = \mathrm{Tr}(\hat{P}_n\hat{\rho}\hat{P}_n^\dagger) \tag{1.11}$$

thanks to properties of trace and projectors.

Another important thing we need when studying quantum measurements is the state right after the measurement took place. Since we wish to represent our quantum states by density matrices, we ought to know the new density matrix of the n-th outcome $\hat{\rho}'_n$ given $\hat{\rho}$ was the state prior to a measurement. We determine $\hat{\rho}'_n$ by acting on $\hat{\rho}$ with projector $\hat{P}_n$ on both sides:

$$\hat{\rho} \to \hat{\rho}'_n = \frac{\hat{P}_n \hat{\rho} \hat{P}_n^\dagger}{\mathrm{Tr}(\hat{P}_n \hat{\rho} \hat{P}_n^\dagger)} = \frac{\hat{P}_n \hat{\rho} \hat{P}_n^\dagger}{P(\lambda_n)}, \tag{1.12}$$

where the denominator is there for normalization of the post-measurement state[1]. The state right after projective measurement, according to the Copenhagen interpretation, collapsed into one of n possible outcomes and revealed all the information that was hidden prior to the measurement. This collapse also means that the state was irreversibly changed and applying the same projective measurement on such state $\hat{\rho}'_n$ gives no new information. In the next section we will take into account the effects of noise and induced errors in the von Neumann measurement.

1.2 Generalized measurements

Generalized quantum measurement can be described by a set of so called measurement operators $\hat{M}_n$, where the index n refers, again, to different outcomes of our experiment. In the von Neumann case, this set is represented by a set of projectors, but this is not always the case (as we will see later). The measurement operators, as was the case for projectors, also need to satisfy a condition called the completeness relation:

$$\sum_n \hat{M}_n^\dagger \hat{M}_n = \hat{I}. \tag{1.13}$$

The probability of obtaining the n-th outcome is then given by:

$$P(n) = \mathrm{Tr}\left(\hat{M}_n \hat{\rho} \hat{M}_n^\dagger\right). \tag{1.14}$$

The post-measurement state is defined just as it was defined previously for the projective measurement, only now $\hat{M}_n^\dagger \neq \hat{M}_n$:

$$\hat{\rho}'_n = \frac{\hat{M}_n \hat{\rho} \hat{M}_n^\dagger}{P(n)}. \tag{1.15}$$

Introducing generalized measurement operators allows for weakening of the destructing effect a projector has on a state. We can demonstrate such weakening on a specific example by defining the following measurement operators acting on a state $|\psi\rangle = \begin{pmatrix} \alpha \\ \beta \end{pmatrix}$ on $\mathcal{H} = \mathbb{C}^2$:

$$\hat{M}_0 = \begin{pmatrix} \sqrt{\frac{2}{3}} & 0 \\ 0 & \sqrt{\frac{1}{3}} \end{pmatrix}, \quad \hat{M}_1 = \begin{pmatrix} \sqrt{\frac{1}{3}} & 0 \\ 0 & \sqrt{\frac{2}{3}} \end{pmatrix}. \tag{1.16}$$

[1]We choose to write the post-measurement state with the hermitian conjugate so the form is the same as forms we use later in our text.

Calculating $\hat{M}_0 |\psi\rangle$ and $\hat{M}_0 |\psi\rangle$ yields the following expressions:

$$\hat{M}_0 |\psi\rangle = \begin{pmatrix} \sqrt{\frac{2}{3}}\alpha \\ \sqrt{\frac{1}{3}}\beta \end{pmatrix} \quad \hat{M}_1 |\psi\rangle = \begin{pmatrix} \sqrt{\frac{1}{3}}\alpha \\ \sqrt{\frac{2}{3}}\beta \end{pmatrix} \tag{1.17}$$

and as we can see this type of measurement did not "destroy" the state $|\psi\rangle$ completely. Probabilities of such measurement would then be:

$$\begin{aligned} P(0) &= \frac{2}{3}|\alpha|^2 + \frac{1}{3}|\beta|^2 \\ P(1) &= \frac{1}{3}|\alpha|^2 + \frac{2}{3}|\beta|^2 \end{aligned} \tag{1.18}$$

instead of $P(0) = |\alpha|^2$ and $P(1) = |\beta|^2$ which would be the case for projectors. Such measurement can then be repeated in order for probabilities to be more defined causing better post-measurement state discrimination. This process of projector weakening is discussed in great detail in section 5.4.

1.3 Positive operator-valued measure

The positive operator-valued measure, or POVM for short, is a mathematical tool designed specifically for quantum measurement analysis. It is defined as a set of operators $\hat{\Pi}_n$ whose expectational values are equal to probabilities of corresponding outcomes. For this reason they are often called probability operators. We can define them for the von Neumann measurement by a relation with measurement operators defined in the section 1.2 as follows:

$$\hat{\Pi}_n \equiv \hat{M}_n^\dagger \hat{M}_n. \tag{1.19}$$

The completeness relation for probability operators now reads:

$$\sum_i \Pi_i = \hat{I}. \tag{1.20}$$

The equation (1.19) makes $\hat{\Pi}_n$ a positive operator. The number of elements in the POVM set can be greater or even smaller than the dimension of the state space they operate on. The probability of getting the n-th outcome is given by:

$$P(n) = \mathrm{Tr}(\hat{\rho}\hat{\Pi}_n). \tag{1.21}$$

It is widely known (for reference, see [1] on page 90) that any set of operators satisfying the completeness relation (1.13) is utilisable for describing any generalized measurement[2], which is then conducted as follows. We start by preparing an ancillary quantum system P, or probe, in a known quantum state, here denoted by a vector $|\xi\rangle \in \mathcal{H}_P$. We then cause this state to interact in a controlled way with the system we seek to measure, which we denote as $|\psi\rangle \in \mathcal{H}$. This interaction, as we mentioned in the beginning of this chapter,

[2]This statement is a direct corollary of Naimark's theorem, see [2] page 31.

can be represented by a unitary operator $\hat{U}$, that acts on the entangled state of the probe $|\xi\rangle$ and our system $|\psi\rangle$, meaning $\hat{U} \in \mathcal{H} \otimes \mathcal{H}_P$. We can write this interaction down as follows:

$$|\xi\rangle \rightarrow \hat{U}\left(|\psi\rangle \otimes |\xi\rangle\right). \tag{1.22}$$

After this interaction has taken place we perform a von Neumann measurement on the composed system. This measurement is equivalent to projecting the now entangled state onto a complete set of states in an extended state space[3] and our POVM then describes a von Neumann measurement on the whole system. The resultant extended state is then partially traced to get rid of the probe.

An important statement needs to be added here. As we already indicated earlier, one can describe any measurement by POVM, but the implication also goes the other way – if a measurement cannot be described by the POVM formalism then it cannot be performed. This is a great way of determining which processes have any relation to reality.

We can show the practicality of POVM in practice on the following example. Let us have a device that measures a qubit[4], meaning the measurement outcome is going to be either 0 or 1. In the case of non-ideal measurement we induce errors on the ideal case, meaning the probability distribution we had previously was distorted by a source of noise in our apparatus. This, from the practical point of view, means that a part of the ideally accessible information was distorted. We start with the ideal measurement constructed in the following form:

$$\begin{aligned} \hat{M}_0 &= \hat{P}_0 = |0\rangle\langle 0| \\ \hat{M}_1 &= \hat{P}_1 = |1\rangle\langle 1|. \end{aligned} \tag{1.23}$$

Let m be different real outcome of our measurement and n be an outcome of the respective ideal measurement. Then the probability of getting m-th outcome is given by:

$$P(m) = \sum_i P(m|n)\mathrm{Tr}(\hat{\rho}\hat{P}_i). \tag{1.24}$$

So in our case the respective probabilities would be:

$$\begin{aligned} P(0) &= (1-p)\mathrm{Tr}(\hat{\rho}\hat{P}_0) + p\mathrm{Tr}(\hat{\rho}\hat{P}_1) \\ P(1) &= (1-p)\mathrm{Tr}(\hat{\rho}\hat{P}_1) + p\mathrm{Tr}(\hat{\rho}\hat{P}_0) \end{aligned} \tag{1.25}$$

in the case of symmetric noisy channel[5]. The probabilities in (1.25) changed since we accounted for some source of noise that was causing statistical errors. These errors can be added into the act of measurement itself. There we had probabilities of getting either

[3]The extended state space can be represented by extra states of the original unknown quantum state not used in prior notation.

[4]Here we mean the term qubit as a two level system, such as polarization of a photon or spin of an electron. We give a brief explanation of both terms at the end of the second chapter.

[5]A symmetric noisy channel measures two possible outcomes $m, n \in \{0, 1\}$ with probabilities $P(m = n) = 1 - p$ and $P(m \neq n) = p$. Further description of classical information channels can be found in [3].

0 or 1 as an outcome in equations (1.25). Since POVM elements have their expectational values equal to probabilities of corresponding outcomes we can rewrite (1.25) as follows:

$$\begin{aligned} P(0) &= (1-p)\langle 0|\hat{\rho}|0\rangle + p\langle 1|\hat{\rho}|1\rangle \\ P(1) &= (1-p)\langle 1|\hat{\rho}|1\rangle + p\langle 0|\hat{\rho}|0\rangle \end{aligned} \tag{1.26}$$

and we can determine the form of both $\hat{\Pi}_0$ and $\hat{\Pi}_1$ as follows:

$$\begin{aligned} \mathrm{Tr}(\hat{\rho}\hat{\Pi}_0) = (1-p)\langle 0|\hat{\rho}|0\rangle + p\langle 1|\hat{\rho}|1\rangle = (1-p)\mathrm{Tr}(\hat{P}_0\hat{\rho}) + p\mathrm{Tr}(\hat{P}_1\hat{\rho}) = \\ = \mathrm{Tr}\left[((1-p)\hat{P}_0 + p\hat{P}_1)\hat{\rho}\right] \end{aligned} \tag{1.27}$$

$$\begin{aligned} \mathrm{Tr}(\hat{\rho}\hat{\Pi}_1) = (1-p)\langle 1|\hat{\rho}|1\rangle + p\langle 0|\hat{\rho}|0\rangle = (1-p)\mathrm{Tr}(\hat{P}_1\hat{\rho}) + p\mathrm{Tr}(\hat{P}_0\hat{\rho}) = \\ = \mathrm{Tr}\left[((1-p)\hat{P}_1 + p\hat{P}_0)\hat{\rho}\right]. \end{aligned} \tag{1.28}$$

From equations (1.27) and (1.28) we can extract the form of both POVM elements:

$$\begin{aligned} \hat{\Pi}_0 &= (1-p)\hat{P}_0 + p\hat{P}_1 \\ \hat{\Pi}_1 &= (1-p)\hat{P}_1 + p\hat{P}_0. \end{aligned} \tag{1.29}$$

What does a post-measurement state look like? One would naively consider the square root of probability operator $\hat{M}_n = \sqrt{\Pi_n}$ which satisfies the equation (1.19), or even more generally $\hat{M}_n = \hat{U}\sqrt{\Pi_n}$ for any unitary operator $\hat{U}$. We could, however, calculate the post-measurement state directly. Since we are using symmetric noisy channel we know we can either get the desired outcome or we can get the other one with probability p. This makes the post-measurement state a statistical mixture of both states $\hat{P}_0\hat{\rho}\hat{P}_0^\dagger$ and $\hat{P}_1\hat{\rho}\hat{P}_1^\dagger$:

$$\begin{aligned} \hat{\rho}_0 &= \frac{(1-p)\hat{P}_0\hat{\rho}\hat{P}_0^\dagger + p\hat{P}_1\hat{\rho}\hat{P}_1^\dagger}{P(0)} \\ \hat{\rho}_1 &= \frac{p\hat{P}_0\hat{\rho}\hat{P}_0^\dagger + (1-p)\hat{P}_1\hat{\rho}\hat{P}_1^\dagger}{P(1)}. \end{aligned} \tag{1.30}$$

As we can see, this is not equal to the form (1.15) for any $\hat{M}_m^\dagger$ [6] so a further generalization is in order. We devote the next section to it.

[6] In a language of generalized measurement there is a possibility of constructing a set of Kraus operators in the following forms. For the first outcome 0 we can write: $\hat{M}_0^{(1)} = \sqrt{1-p}\hat{P}_0, \hat{M}_0^{(2)} = \sqrt{p}\hat{P}_1$ and for the second outcome 1 we can write: $\hat{M}_1^{(1)} = \sqrt{1-p}\hat{P}_1, \hat{M}_1^{(2)} = \sqrt{p}\hat{P}_0$. The POVM set would then be constructed as follows: $\hat{\Pi}_0 = \hat{M}_0^{(1)\dagger}\hat{M}_0^{(1)} + \hat{M}_0^{(2)\dagger}\hat{M}_0^{(2)} = (1-p)\hat{P}_0 + p\hat{P}_1$ and $\hat{\Pi}_1 = \hat{M}_1^{(1)\dagger}\hat{M}_1^{(1)} + \hat{M}_1^{(2)\dagger}\hat{M}_1^{(2)} = (1-p)\hat{P}_1 + p\hat{P}_0$. We could also define another set as follows: $\hat{\tilde{M}}_0 = \sqrt{(1-p)\hat{P}_0 + p\hat{P}_1} = \sqrt{(1-p)}\hat{P}_0 + \sqrt{p}\hat{P}_1$, where $\hat{\tilde{M}}\hat{\tilde{M}}^{(\dagger)} = \hat{\Pi}_0$ and $\hat{\tilde{M}}_1 = \sqrt{(1-p)\hat{P}_1 + p\hat{P}_0} = \sqrt{(1-p)}\hat{P}_1 + \sqrt{p}\hat{P}_0$, where $\hat{\tilde{M}}_0\hat{\tilde{M}}_0^{(\dagger)} = \hat{\Pi}_0$ and $\hat{\tilde{M}}_1\hat{\tilde{M}}_1^{(\dagger)} = \hat{\Pi}_1$. Upon investigating the post-measurement state we find out, that $\hat{\tilde{\rho}}_0 = \left(\hat{\tilde{M}}_0\hat{\rho}\hat{\tilde{M}}_0^{(\dagger)}\right) \neq \hat{\rho}_0$, where $\hat{\rho}_0$ is the post-measurement state of the first set of measurement operators without tilde. This ambiguity need correction that we provide in the next section.

1.4 Operations

In previous sections we focused on how a generalized measurement can be performed and how to determine probabilities for each possible outcome. In this section we will present a more versatile tool for describing changes that a system undergoes during measurement. Quantum operations represent a great mechanism for describing other than destructive (projective) measurements and to handle unwanted interactions with the environment, which we identify with noise. They are also practical and useful for describing sequential measurement.

A quantum operation can be described as a completely positive, trace non-increasing map[7] $\mathcal{E}$ that takes a density matrix $\hat{\rho}$ representing the initial state and, without an explicit reference for the passage of time, gives us the final state density matrix $\hat{\rho}'$:

$$\hat{\rho} \to \mathcal{E}(\hat{\rho}). \tag{1.31}$$

As we can see the map denoted here as $\mathcal{E}$ represents a state transformation that can be repeated or even composed. The post-measurement state will take the following form:

$$\hat{\rho}'_m = \frac{\mathcal{E}_m(\hat{\rho})}{\mathrm{Tr}(\mathcal{E}_m(\hat{\rho}))} = \frac{\mathcal{E}_m(\hat{\rho})}{P(m)}, \tag{1.32}$$

where $P(m)$ is the probability of measuring the m-th outcome. Quantum operations can be represented by a set of measurement operators $\{\hat{A}_m\}$ commonly known as effects:

$$\mathcal{E}_m(\hat{\rho}) = \sum_i \hat{A}_m^{(i)} \hat{\rho} \hat{A}_m^{(i)\dagger}. \tag{1.33}$$

There also exists a dual map to $\mathcal{E}_m$ denoted as $\mathcal{E}_m^*(\hat{\rho})$:

$$\mathcal{E}_m^*(\hat{\rho}) = \sum_i \hat{A}_m^{(i)\dagger} \hat{\rho} \hat{A}_m^{(i)} \tag{1.34}$$

thanks to which we can define a POVM set as follows:

$$\hat{\Pi}_m = \mathcal{E}_m^*(\hat{I}) = \sum_i \hat{A}_m^{(i)\dagger} \hat{A}_m^{(i)}. \tag{1.35}$$

As we can see, $\mathcal{E}_m$ determines $\hat{\Pi}_m$ uniquely. How about the inverse statement? Unfortunately, there are infinitely many sets of $\{\hat{A}_m^{(i)}\}$ satisfying the equation (1.35) corresponding to different physical interpretations. So the set of effects can't be reconstructed from the POVM alone.

Let's assume the result of a generalized measurement performed with POVM $\{\hat{\Pi}_m\}$ in the form (1.35) is m. Post-measurement states are described by effects as:

$$\hat{\rho} \to \hat{\rho}'_m = \frac{\sum_i \hat{A}_m^{(i)} \hat{\rho} \hat{A}_m^{(i)\dagger}}{\mathrm{Tr}(\hat{A}_m^{(i)} \hat{\rho} \hat{A}_m^{(i)\dagger})} = \frac{\sum_i \hat{A}_m^{(i)} \hat{\rho} \hat{A}_m^{(i)\dagger}}{P(m)} \tag{1.36}$$

[7] For more information on quantum operations see [4].

with probabilities given by:

$$P(m) = \mathrm{Tr}(\mathcal{E}_m(\hat{\rho})) = \mathrm{Tr}(\hat{\Pi}_m\hat{\rho}) = \sum_i \mathrm{Tr}(\hat{A}_m^{(i)}\hat{\rho}\hat{A}_m^{(i)\dagger}). \tag{1.37}$$

As we can see the equation (1.15) is a special case for $\{\hat{A}_m^{(i)}\} = \{\hat{M}_m\}$.

Effects can be chained to represent joint measurement. If we measure the state (1.36) with a set of operators $\{\hat{B}_j\}$ with outcomes labeled as j, then the joint probability of measuring the outcome j after measuring i can be calculated using the conditional probability formula $P(b_j|a_i)P(a_i) = P(a_i, b_j)$:

$$P(j|i) = \frac{\mathrm{Tr}\left(\hat{B}_j\hat{A}_i\hat{\rho}\hat{A}_i^\dagger\hat{B}_j^\dagger\right)}{\mathrm{Tr}\left(\hat{A}_i\hat{\rho}\hat{A}_i^\dagger\right).} \tag{1.38}$$

So the joint probability of measuring states i and j sequentially is:

$$P(i,j) = P(j|i)P(i) = \mathrm{Tr}\left(\hat{B}_j^\dagger\hat{B}_j\hat{A}_i\hat{\rho}\hat{A}_i^\dagger\right) = \mathrm{Tr}\left(\hat{A}_i^\dagger\hat{B}_j^\dagger\hat{B}_j\hat{A}_i\hat{\rho}\right). \tag{1.39}$$

As we can see, by utilizing properties of tracing, we can understand this sequential measurement as a new POVM and determine the combined probability operator as:

$$\hat{\Pi}_{ij} = \hat{A}_i^\dagger\hat{B}_j^\dagger\hat{B}_j\hat{A}_i. \tag{1.40}$$

If we define $\hat{C}_{ij} = \hat{B}_j\hat{A}_i$ then:

$$\hat{\Pi}_{ij} = \hat{C}_{ij}^\dagger\hat{C}_{ij}. \tag{1.41}$$

In this case the operator $\hat{C}_{ij}$ acts as a measurement operator for obtaining the combined results (i,j) for the sequential measurement and the initial state described by $\hat{\rho}$ changes as follows:

$$\hat{\rho} \to \sum_{ij} \hat{A}_i^\dagger\hat{B}_j^\dagger\hat{\rho}\hat{B}_j\hat{A}_i = \hat{C}_{ij}^\dagger\hat{\rho}\hat{C}_{ij}. \tag{1.42}$$

The operations themselves can also be composed without the use of effects in a rather elegant way. If we took a density matrix $\hat{\rho}$ with a post-measurement state defined as:

$$\hat{\rho}_i' = \frac{\mathcal{E}_i(\hat{\rho})}{P(i)}, \tag{1.43}$$

we can take this new state $\hat{\rho}_i'$ and apply a new operation, this time denoted as $\mathcal{F}_j$. The new post-measurement state can then be written as:

$$\hat{\rho}_{i\to j}' = \hat{\rho}_{ij}' = \frac{\mathcal{F}_j(\hat{\rho}_i')}{P(i,j)} = \frac{1}{P(j|i)}\mathcal{F}_j\left(\frac{\mathcal{E}_i(\hat{\rho})}{P(i)}\right) = \frac{1}{P(j|i)P(i)}\mathcal{F}_j(\mathcal{E}_i(\hat{\rho})), \tag{1.44}$$

where $P(i) = \mathrm{Tr}(\mathcal{E}_i(\hat{\rho}))$ and $P(j|i)P(i) = P(i,j) = \mathrm{Tr}(\mathcal{F}_j(\mathcal{E}_i(\hat{\rho})))$. The new composed operation can then be denoted as $\mathcal{G}_{i,j}$:

$$\mathcal{G}_{i,j} = \mathcal{F}_j \circ \mathcal{E}_i. \tag{1.45}$$

Chapter 2

Fundamentals of information theory

In this chapter we will present basic theorems from classical information theory, or Shannon theory, needed in this work for proper understanding of the next chapter. We define information in a mathematically correct way for classical systems and discuss how the situation changes for quantum systems. The main source for this chapter was [5] and [6] together with the paper A Mathematical Theory of Communication by C. E. Shannon [3].

For correct definition of information in the mathematical sense we first need to introduce the concept of entropy. Let X be a discrete random variable with realizations denoted as x. Entropy can then be understood as its measure of uncertainty or as a measure of the amount of information needed to describe the random variable, or information content for short. Let $\mathfrak{X}$ be an alphabet of X, $\mathfrak{X}$ finite, and let $P(x) = P(X = x), x \in \mathfrak{X}$ be its probability function.

Definition 1. *The entropy, denoted $H(X)$, of a random discrete variable X is defined as*

$$H(X) = -\sum_{x\in\mathfrak{X}} P(x) \log_2 P(x). \tag{2.1}$$

We shall omit the subscript in the logarithm henceforward since the entropy is expressed in bits.

One could ask why the logarithm - and not some other type of function - is used to define the measure of information. The reasons are quite simple. First, we require higher uncertainty for lower probability, but the most important argument is that logarithm has useful properties, such as additivity or positivity in the whole range of the function. Shannon himself gives several reasons in his famous paper why logarithm is the most convenient choice. First and foremost it is very close to our intuitive understanding of measure (logarithmic scale is used in many different sciences as a scale, for example pH, sound level or Richter magnitude scale) but it is also mathematically suitable, as we suggested earlier. This gives the entropy H properties that very much make the concept intuitive. $H = 0$ only if the probabilities are all equal to one meaning there is no uncertainty in things that happen for sure. For all the other cases, when the probability takes a values $P(x) \in (0, 1)$,

H stays positive. In the next definition we extend the number of discrete variables and define joint and conditional entropy, which will be needed to study collective systems.

Definition 2. *The joint entropy, denoted $H(X,Y)$, of a pair of random discrete variables (X,Y) with a joint probability distribution $P(x,y)$ is defined as*

$$H(X,Y) = -\sum_{x\in\mathfrak{X}}\sum_{y\in\mathfrak{Y}} P(x,y)\log P(x,y). \tag{2.2}$$

If the two discrete variables (X,Y) are independent, then:

$$H(X,Y) = H(X) + H(Y). \tag{2.3}$$

Definition 3. *The conditional entropy, denoted $H(Y|X)$, of a pair of random discrete variables (X,Y) with a joint probability distribution $P(x,y)$ is defined as*

$$H(Y|X) = -\sum_{x\in\mathfrak{X}}\sum_{y\in\mathfrak{Y}} P(x,y)\log P(y|x). \tag{2.4}$$

Theorem 1. *Chain rule for the joint entropy:*

$$H(X,Y) = H(X) + H(Y|X). \tag{2.5}$$

The chain rule can be proved using the following equation[1]:

$$P(y|x) = \frac{P(x,y)}{P(x)}. \tag{2.6}$$

A direct corollary to Theorem 1 is the following equation:

$$H(X,Y|Z) = H(X|Z) + H(Y|X,Z). \tag{2.7}$$

We now introduce relative entropy, the measure of distance between two distributions, and mutual information:

Definition 4. *The relative entropy $D(p||q)$ is a measure of distance between two probability distributions $p(x)$ and $q(x)$ and it can be calculated as:*

$$D(p||q) = \sum_{x\in\mathfrak{X}} p(x)\log\frac{p(x)}{q(x)}. \tag{2.8}$$

The mutual information is then defined as follows:

Definition 5. *For two random variables X and Y, with probability distributions $p(x)$ and $p(y)$ and joint probability distribution $p(x,y)$, the mutual information si expressed by the relative entropy between the joint and the product distribution $p(x)p(y)$:*

$$I(X;Y) = \sum_{x\in\mathfrak{X}}\sum_{y\in\mathfrak{Y}} p(x,y)\log\frac{p(x,y)}{p(x)p(y)} = D(p(x,y)||p(x)p(y)). \tag{2.9}$$

[1]For more information about discrete probability distribution see [7].

Another way we can express the mutual information is in terms of entropy:

$$I(X;Y) = H(X) - H(X|Y) = H(Y) - H(Y|X), \tag{2.10}$$

and as we can see, the mutual information reduces the uncertainty of X thanks to the knowledge of Y and vice versa. This also means that X gives the same amount of information about Y as Y gives about X. A corollary to the symmetric properties of entropy is the following statement:

$$I(X;Y) = H(X) + H(Y) - H(X,Y). \tag{2.11}$$

Now that we defined all the necessary concepts for two random variables, we should extend the previous definitions for N random variables, $N \in \mathbb{N}$.

Theorem 2. *Let us have random variables $X_1, X_2, \cdots, X_N$ with probability distribution $p(x_1, x_2, \cdots, x_N)$, then*

$$H(X_1, X_2, \cdots, X_N) = \sum_{i=1}^{N} H(X_i|X_{i-1}, \cdots, X_1). \tag{2.12}$$

Using the result of 2 , we can write the next statement:

$$I(X_1, X_2, \cdots, X_N; Y) = \sum_{i=1}^{N} I(X_i; Y|X_{i-1}, \cdots, X_1). \tag{2.13}$$

Another thing that requires a brief written summary is the notion of discrete channels. Shannon defines discrete channel as a system which transmits a sequence of symbols picked from a discrete set between two parties. He offers as an example any stochastic process producing discrete sequence of symbols $X \in \mathfrak{X}$, such as human languages or quantized signals. One channel in particular, which was mentioned in the previous chapter in section 1.3, was the symmetric noisy channel, which took into account the possibility of information distortion during the transmission. Another type of discrete channels will be used later in this thesis for mapping classical outcomes $m \in M$ from a measurement to a set containing two elements, namely 0 and 1.

We conclude this chapter by a brief discussion about how the situation changes for quantum systems. The outcome of Shannon's information theory is the concept of bit as a measure of entropy which is connected to information via equation (2.1). 1 bit represents the amount of information we get when answering a yes-or-no type of question and can be physically represented by a device with two stable positions, for example by a transistor. This idea of physical representation is crucial for expressing quantum information, where the bit changes to quantum bit, or qubit for short. This measure of quantum information is then represented by a two-level quantum system, such as the polarization of a photon or an atom in ground and excited energy level and by measuring such quantum state we get at most one bit of information.

Chapter 3

Noise and disturbance in measurement

In this chapter we discuss in detail how to mathematically define the terms noise and disturbance in quantum measurement and find a trade-off between those two, which will help us understand the concept of joint measurements and what happens when we try to measure two observables on one system. We use mainly two sources, see [8] and [9].

3.1 Noise and disturbance using statistics

We will start this section by examination of the Heisenberg uncertainty principle which will naturally lead to definitions of noise and disturbance. The Heisenberg relation is usually demonstrated by the Robertson uncertainty relation, which is defined for two non-commuting observables as:

$$\sigma(\hat{A},\psi)\sigma(\hat{B},\psi) \geq \frac{|\langle\psi|[\hat{A},\hat{B}]|\psi\rangle|}{2}, \tag{3.1}$$

where $\sigma(\hat{X},\psi)$ is the standard deviation of an observable $\hat{X}$ and a state ψ defined as:

$$\sigma(\hat{X},\psi)^2 = \langle\psi|\hat{X}^2|\psi\rangle - \langle\psi|\hat{X}|\psi\rangle^2. \tag{3.2}$$

One has to be careful though because there is a certain ambiguity in what we generally call the Heisenberg principle. The general interpretation of relation (3.1) refers to the physically possible accuracy of quantum state preparation and has no relation to disturbance introduced by a measurement. It gives us the fundamental limit of knowledge precision of two general non-commuting complementary variables. On the other hand, the observer effect is what gives us the limit on measuring one observable and causing the other observable to be disturbed by this measurement. It tells us that no measurement in quantum mechanics can be done without somehow affecting the system. So for a measuring apparatus M and an observables $\hat{A},\hat{B}$ we get what is called Heisenberg noise-disturbance uncertainty relation:

$$\epsilon(M,\hat{A})\eta(M,\hat{B}) \geq \frac{|\langle\psi|[\hat{A},\hat{B}]|\psi\rangle|}{2}, \tag{3.3}$$

where the function $\epsilon(M, \hat{A})$ refers to noise and $\eta(M, \hat{B})$ represents disturbance. We can also formulate what is called the Heisenberg uncertainty relation for joint measurements:

$$\epsilon(M, \hat{A})\epsilon(M, \hat{B}) \geq \frac{|\langle\psi|[\hat{A}, \hat{B}]|\psi\rangle|}{2}, \tag{3.4}$$

which describes a joint measurement of observables $\hat{A}, \hat{B}$ by an apparatus M. The relationship between equations (3.3) and (3.4) can be described as follows: if the measurement of $\hat{A}$ by an apparatus A is immediately followed by measuring $\hat{B}$ by a noiseless apparatus B, we can put A and B together and get a new apparatus denoted as C. This new apparatus C then performs joint measurement on $\hat{A}$ and $\hat{B}$. Since B is noiseless, the noise caused by C on $\hat{B}$ is equal to the disturbance of $\hat{B}$ caused by A. In other words:

$$\epsilon(A, \hat{A}) = \epsilon(C, \hat{A}) \tag{3.5}$$

$$\eta(A, \hat{B}) = \epsilon(C, \hat{B}). \tag{3.6}$$

Equation (3.3) cannot be generally reduced to (3.4) due to reasons explained in a great detail in [9]. The noise can be conveniently defined as a root-mean-square deviation, which gives us the distance between a system observable and our measurement outcome. We first deduce the real value of the classical outcome, meaning the value one would measure without noise in one's apparatus. We say that M measures $\hat{A}$ precisely if

$$P(m) = \langle\psi|\hat{E}(m)|\psi\rangle, \tag{3.7}$$

where $\hat{E}(m)$ is the projector for an outcome m. Equation (3.7) is called Born statistical formula and it can also be understood in terms of how much the experimental probability distribution coincides with theoretical prediction. Now, if M does not satisfy the condition (3.7), we say the apparatus M measures $\hat{A}$ with noise.

To correctly determine the noise and the disturbance of our measurement we first write down the evolution of our system in time. The input state is given by entangling the object we seek to measure with a probe denoted as P prepared in some referential state $|\xi\rangle$. We denote the state of such system as $\psi \otimes \xi$. Then:

$$\hat{A}^{in} = \hat{A} \otimes \hat{I}, \tag{3.8}$$

$$\hat{P}^{in} = \hat{I} \otimes \hat{P}, \tag{3.9}$$

$$\hat{A}^{out} = \hat{U}^{\dagger}(\hat{A} \otimes \hat{I})\hat{U}, \tag{3.10}$$

$$\hat{P}^{out} = \hat{U}^{\dagger}(\hat{I} \otimes \hat{P})\hat{U} \tag{3.11}$$

where $\hat{U}$ represents the time evolution of our composite system $\psi \otimes \xi$ and $\hat{P}$ denotes some observable to be measured on P, chosen suitably to reveal information about $\hat{A}$. The output is then recorded by m with probability distribution given by:

$$P(m) = \langle\psi \otimes \xi|\hat{E}(p^{out})|\psi \otimes \xi\rangle, \tag{3.12}$$

where $\{p^{out}\}$ is the set of classical outcomes of $\hat{P}^{out}$. The noise $\epsilon(M, \hat{A})$ is then defined as the root-mean-square deviation of $\hat{P}^{out}$ from the theoretical $\hat{A}^{in}$:

$$\epsilon(M, \hat{A}) = \langle(\hat{P}^{out} - \hat{A}^{in})^2\rangle^{1/2}. \tag{3.13}$$

The disturbance is defined as the degree to which one can correct such noise for a given observable and it can, as well, be defined as root-mean-square of the change happening during measurement:

$$\eta(M, \hat{B}) = \langle(\hat{B}^{out} - \hat{B}^{in})^2\rangle^{1/2}. \tag{3.14}$$

It can then be proved that $\eta = 0$ if and only if:

$$\langle E(b^{in})\rangle = \langle E(b^{out})\rangle \tag{3.15}$$

for any input state.

The last paragraph of this section will be devoted to an important result of the paper [9], where the authors corrected the noise-disturbance uncertainty relation (3.3) by adding an additional term that made the relation universally valid for all types of measurement. They called this additional term the correlation term and ensured that those experiments which found a violation of the original Heisenberg principle[1], obeyed this new inequality. We now define several new terms for this new inequality as follows:

$$\hat{P}^{out} = \hat{A}^{in} + N(A) \tag{3.16}$$
$$\hat{B}^{out} = \hat{B}^{in} + D(B), \tag{3.17}$$

where $N(A)$ is called the noise operator and $D(B)$ is the disturbance operator. Also, $[\hat{P}^{out}, \hat{B}^{out}] = 0$, which tells us that $\hat{M}$ and $\hat{B}$ are observables in different systems. Then after some steps, which are described in great detail in [9], we get the universally valid noise-disturbance relation:

$$\epsilon(\hat{A})\eta(\hat{B}) + \frac{|\,\langle[N(A), \hat{B}^{in}]\rangle + \langle[\hat{A}^{in}, D(B)]\rangle\,|}{2} \geq \frac{|\,\langle\psi|[\hat{A}, \hat{B}]|\psi\rangle\,|}{2} \tag{3.18}$$

for any pair of observables $(\hat{A}, \hat{B})$. This equation can also be written in terms of standard deviations $\sigma(\hat{A})$ and $\sigma(\hat{B})$ as follows:

$$\epsilon(\hat{A})\eta(\hat{B}) + \epsilon(\hat{A})\sigma(\hat{B}) + \sigma(\hat{A})\eta(\hat{B}) \geq \frac{|\,\langle\psi|[\hat{A}, \hat{B}]|\psi\rangle\,|}{2}. \tag{3.19}$$

3.2 Noise and disturbance using entropy

Another approach to define the noise and disturbance can be done using entropy. First we need to specify the measurement model. Let's consider a quantum system S with two

[1]For more information see [10]

observables $\hat{A}$, whose set of eigenvalues is $\{|\psi^a\rangle\}$, and B with set of eigenvalues $\{|\phi^b\rangle\}$. We subject S to some measurement denoted as M and get an output labeled m. This observed outcome can then be compared to eigenvalues of measured observable denoted as $\{a\}$ and $\{b\}$ for the observables $\hat{A}$ and $\hat{B}$, respectively. Since correlation measurements can be expressed by a conditional probability distribution, quantifying noise means finding this distribution $P(a|m)$, where by a we mean the input state taken from the set $\{|\psi^a\rangle\}$. Assuming no prior information is available we have:

$$P(a) = \frac{1}{d}, \tag{3.20}$$

where $d = \dim \mathscr{H}$ is dimension of our Hilbert space. We can thus provide the joint input-output probability:

$$P(m,a) = P(a)P(m|a) = \frac{P(m|a)}{d}. \tag{3.21}$$

We quantify both noise and disturbance using correlations between input and output of some measurement device. The noise can be hence defined as follows:

Definition 6. *The noise of the instrument M as a measurement of $\hat{A}$ is defined as*

$$N(M,\hat{A}) \equiv H(\hat{A}|M), \tag{3.22}$$

where $H(\hat{A}|M)$ denotes the conditional entropy.

The conditional entropy $H(\hat{A}|M)$ can be computed from $P(m,a)$ and $P(a|m)$ using the Bayes rule.[2]

We can see the noise decreases as the precision of our correctly guessing the eigenvalue ξ^a from m increases. The situation is somewhat more complicated for disturbance since some kinds of measured back-action can be corrected by post-processing. Unlike noise we have an option to use any possible correction procedures that will help us restore more information about the input state. Such procedure can be mathematically defined by a completely positive trace-preserving map $\mathcal{E}$ that attempts at reconstructing the pre-measurement quantum system S from S'. The disturbance will then depend on both $m = b$ and a guess given by $\mathcal{E}S' = S$, denoted here as b', and can be characterized using correlation between b and b'. This correlation is then given by a joint probability distribution between b and b':

$$P(b',b) = P(b)P(b'|b) = \frac{P(b'|b)}{d} \tag{3.24}$$

Using this result we can define disturbance properly.

[2]The Bayes rule gives us conditional probability $P(X|Y)$ for two discrete variables X, Y with respective probabilities $P(X), P(Y)$ and conditional probability $P(Y|X)$ as follows:

$$P(X|Y) = \frac{P(Y|X)P(X)}{P(Y)}, \quad \forall y : P(Y=y) \neq 0. \tag{3.23}$$

More on probability theory can be found for example in [7].

Definition 7. *The disturbance any measuring device introduces on any subsequent measurement of $\hat{B}$ is defined as:*

$$D(M, \hat{B}) \equiv \min_{\mathcal{E}} H(\hat{B}|B'), \tag{3.25}$$

where the entropy $H(\hat{B}|B')$ can be computed from $P(b', b)$.

Definitions 6 and 7 then lead to a noise-disturbance relation defined for any measuring apparatus M and any non-degenerate observables A and B as:

$$N(M, \hat{A}) + D(M, \hat{B}) \geq -\log_2 c, \tag{3.26}$$

where $c \equiv \max_{a,b} | \langle \psi^a | \phi^b \rangle |^2$. The equation (3.26) tells us that when $\hat{A}$ and $\hat{B}$ are incompatible, meaning $c < 1$, it is impossible to measure one with no noise without, at the same time, disturbing the other. It also tells us how any measurement apparatus must disturb one observable to gain information about the other observable.

We should add an important note before proceeding further. The disturbance defined by equation (3.25) is minimized over the full set of operations $\mathcal{E}$ which doesn't always present the most convenient way for calculations. This means that in some cases we will skip the opportunity for the correction, fixing $\mathcal{E}$ at the identity map meaning the following simplification will be used instead:

$$D(M, \hat{B}) \equiv H(\hat{B}|B'). \tag{3.27}$$

The inequality (3.26) remains the same.

Chapter 4

Zero disturbance on systems with two identical particles

In this chapter we will explore how to use the language of joint measurements to cause as little disturbance as possible on our system made out of two copies of the same particle. The main motivation behind this chapter is to show that disturbance can be influenced by choosing the measurement operators correctly and using some tricks.

We will start with a state that can be described as follows:

$$|\phi\rangle = |\psi\rangle \otimes |\psi\rangle \,, \quad |\psi\rangle = \alpha\,|0\rangle + \beta\,|1\rangle \tag{4.1}$$

where $\alpha, \beta \in \mathbb{C}$. In the system of this state we consider two single-qubit observables, which we will denote as $\hat{A}$ and $\hat{B}$. These will be extended to $\hat{A} \otimes \hat{I}$ and $\hat{B} \otimes \hat{I}$ so that they can act on a two-qubit system $|\phi\rangle$ while retaining commutation properties of $\hat{A}$ and $\hat{B}$. The question now is how to make a measurement with as little as possible noise and disturbance.

There is always a trivial way to accomplish no disturbance whatsoever by achieving in some way that $\hat{A}$ is measured on the first qubit and $\hat{B}$ is measured on the second qubit. This makes the measurement absolutely free from any disturbance because it basically describes two independent measurement with no relationship between the two observables $\hat{A}$ and $\hat{B}$, such as in (3.3). Here noise and disturbance are zero and this trick represents a great motivation for us to try ”squeezing” the disturbance in other situations. In the following sections we show how this trick formally fits into both statistical and entropy approach described in sections 3.1 and 3.2.

4.1 Zero disturbance using statistical noise and disturbance

As we showed in the third chapter, both noise and disturbance can be quantified using different operators on input and output states, see relations (3.8) – (3.11). If we choose these operators "cleverly" we can show that the disturbance defined in (3.14) will be zero. So let us have the system S of two qubits with the input state $|\phi\rangle$. We entangle this state with a probe system also described by a qubit in some initial state $|\xi\rangle$ and define the unitary operator as the SWAP operator[1] which we let act on the last two qubits of our composite system. We also have to redefine the operators $\hat{A}^{in}, \hat{B}^{in}, \hat{P}^{in}$ so that they can act on $2 \times 2 \times 2$ dimensional systems. The reason is that the four-dimensional input state is now entangled with two-dimensional probe[2]. We label these new operators with a subscript $8D$ and define them as follows:

$$\hat{A}^{in}_{8D} = \hat{A} \otimes \hat{I} \otimes \hat{I}, \tag{4.3}$$

$$\hat{B}^{in}_{8D} = \hat{B} \otimes \hat{I} \otimes \hat{I}, \tag{4.4}$$

$$\hat{P}^{in}_{8D} = \hat{I} \otimes \hat{I} \otimes \hat{P}. \tag{4.5}$$

If we now apply the SWAP operator on equations (4.3), (4.4) and (4.5), we get:

$$\hat{A}^{out}_{8D} = \hat{A} \otimes \hat{I} \otimes \hat{I}, \tag{4.6}$$

$$\hat{B}^{out}_{8D} = \hat{B} \otimes \hat{I} \otimes \hat{I}, \tag{4.7}$$

$$\hat{P}^{out}_{8D} = \hat{I} \otimes \hat{P} \otimes \hat{I}. \tag{4.8}$$

Now we can use the equation (3.14) to calculate the disturbance this measurement makes on our composite system:

$$\eta(M, \hat{B}_{8D}) = \langle (\hat{B}^{out}_{8D} - \hat{B}^{in}_{8D})^2 \rangle^{1/2} = \langle 0 \rangle^{1/2} = 0. \tag{4.9}$$

An interesting point can be brought up – even though we found a combination of operators so that the measurement has zero disturbance, it doesn't have zero noise. If we try to calculate the noise according to equation (3.13) we actually find out that it is non-zero:

$$\epsilon(M, \hat{A}_{8D}) = \langle (\hat{P}^{out}_{8D} - \hat{A}^{in}_{8D})^2 \rangle^{1/2} = \langle (\hat{I} \otimes \hat{P} \otimes \hat{I} - \hat{A} \otimes \hat{I} \otimes \hat{I})^2 \rangle^{1/2}.$$

Clearly on a state of the form (4.1) a measurement of $\hat{A} \otimes \hat{I} \otimes \hat{I}$ can be substituted by that of $\hat{I} \otimes \hat{A} \otimes \hat{I}$, therefore it is natural to choose $\hat{A} \equiv \hat{P}$:

$$\epsilon(M, \hat{A}_{8D}) = \langle (\hat{I} \otimes \hat{A} \otimes \hat{I} - \hat{A} \otimes \hat{I} \otimes \hat{I})^2 \rangle^{1/2} = \langle (\hat{I} \otimes \hat{A}^2 \otimes \hat{I} - 2\hat{A} \otimes \hat{I} \otimes \hat{I} + \hat{A}^2 \otimes \hat{I} \otimes \hat{I})^2 \rangle^{1/2} =$$

$$= (2 \langle \hat{A}^2 \rangle - 2 \langle \hat{A} \rangle^2)^{1/2} = \sqrt{2}\sigma(\hat{A}) \neq 0.$$

We can see that this choice of measurement does not satisfy the noise-disturbance uncertainty relation (3.3) for generally non-commuting $\hat{A}$ and $\hat{B}$. This case is one of the ones for which we had to define the new uncertainty relation (3.19).

[1] Such operator can be represented in matrix form on a two qubit system as:

$$\hat{U} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{4.2}$$

[2] The probe can, in general, have as many dimensions as we want. We chose two in this particular case for practical reasons.

4.2 Zero disturbance using entropy

Similarly to the previous section we will show that by choosing $\mathcal{E} = \text{SWAP}$ and observables $\hat{A} \otimes \hat{I}, \hat{B} \otimes \hat{I}$ we achieve zero disturbance. This is one of those cases where using the simplified version (3.27) doesn't necessarily mean it will simplify our calculation. Also, the operation $\mathcal{E}$ was defined, so it is easier to make the minimization.

The relation (3.25) in Definition 7 is a function of eigenvalues $\{b\}$ of an observable $\hat{B}$ and our set of guesses $\{b'\}$. There is no need for the relation (3.24) though, since the guess made by $\mathcal{E} = \text{SWAP}$ gives us exactly the value b, meaning $b' = b$ in this particular case. Using relation (3.25) we get:

$$H(\hat{B}|B')|_{\mathcal{E}=SWAP} = H(\hat{B}|\hat{B}). \tag{4.10}$$

Using the chain rule from the joint entropy[3] we know that $H(X|X)$ for a discrete random variable X is zero, which can't be further minimized, meaning:

$$D(M, \hat{B}) = H(\hat{B}|\hat{B}) = 0. \tag{4.11}$$

This concludes the chapter discussing how to achieve zero disturbance on a system containing two identical particles. In the next chapter we shall discuss how to reduce disturbance in a more restrictive scenario with three particles and, finally, with N particles, $N \in \mathbb{N}$.

[3] See Theorem 1 in the second chapter.

Chapter 5

Reducing disturbance without addressing specific particles

In this chapter we present three different types of measurement, which no longer distinguish between particles, together with disturbance calculations. The measured system now contains three particles, instead of two, which could for example be indistinguishable bosons. Assuming particle indistinguishability for constructed measurements means that the trick used in the previous chapter, where we applied the measurement operators in a way that one measurement did had no relation to the other one, is no longer realizable. Another implemented change is that the number of measurement operators differs with the number of particles in our system. There were two copies of the same system in the previous chapter. This number is now increased to three meaning we can illustrate our problem on more general examples. Three approaches on how to construct measurements under said conditions will be discussed in great detail and the main differences will be shown. The aim of this chapter is to primarily show these differences in terms of overall system disturbance. We will also discuss possible approaches of disturbance reduction.

Addressing an ensemble of particles in a non-distinguishing manner, two basic approaches are easily identifiable: either treating every particle the same way, or choosing one at random and averaging over the choice. Both approaches will be discussed in following sections.

5.1 Measuring total spin

Let us first consider subjecting every particle of the ensemble to the same binary measurement described by projectors $\{\hat{P}_0, \hat{P}_1\}$. For indistinguishable particles, we only learn the number of $\hat{P}_0$ and $\hat{P}_1$ outcomes. If the measurement was a measurement of spin component in a given axis, this situation would arise in measuring the total spin of the system in this direction, so we will use this both to illustrate the principle and to refer to this scheme.

Let us have a system of three identical spin-1/2 particles described as follows:

$$|\phi\rangle = |\psi\rangle \otimes |\psi\rangle \otimes |\psi\rangle \,, \tag{5.1}$$

where $|\psi\rangle = \alpha |0\rangle + \beta |1\rangle$, and measure the total spin in the axis which fixes basis $|0\rangle, |1\rangle$. The possible outcomes are:

$$\left\{-\frac{3}{2}, -\frac{1}{2}, \frac{1}{2}, \frac{3}{2}\right\} \tag{5.2}$$

which we want to map to the values 0 and 1 using some classical information channel so that the the single-particle probabilities $P(0) = |\alpha|^2 = p$ and $P(1) = |\beta|^2 = 1 - p$ are reconstructed for any input state. Respective measurement operators can then be found as follows:

$$\begin{aligned}
\hat{A}_0 &= \hat{P}_0 \otimes \hat{P}_0 \otimes \hat{P}_0 \\
\hat{A}_1 &= \hat{P}_1 \otimes \hat{P}_0 \otimes \hat{P}_0 + \hat{P}_0 \otimes \hat{P}_1 \otimes \hat{P}_0 + \hat{P}_0 \otimes \hat{P}_0 \otimes \hat{P}_1 \\
\hat{A}_2 &= \hat{P}_1 \otimes \hat{P}_1 \otimes \hat{P}_0 + \hat{P}_1 \otimes \hat{P}_0 \otimes \hat{P}_1 + \hat{P}_0 \otimes \hat{P}_1 \otimes \hat{P}_1 \\
\hat{A}_3 &= \hat{P}_1 \otimes \hat{P}_1 \otimes \hat{P}_1.
\end{aligned} \tag{5.3}$$

This measurement is designed in such a way that we're not able to distinguish between particles. This indicates invariance towards permutation of our set of particles or, in quantum mechanical language, the quantum operations assigned to a measurement commute with permutation. The overall probability for the first outcome is p^3 since only three spins down give the overall value of spin equal to $-3/2$. Similarly, the second equation tells us that the channel maps two outcomes of $-1/2$ and one outcome $1/2$ to 0 in three different ways meaning the overall probability in terms of p would be $3p^2(1-p)$. This way we can assign a polynomial function of p to every outcome $m \in M$.:

$$\begin{aligned}
P\left[m = -\frac{3}{2}\right] &= p^3 = p_0 \\
P\left[m = -\frac{1}{2}\right] &= 3p^2(1-p) = p_1 \\
P\left[m = \frac{1}{2}\right] &= 3p(1-p)^2 = p_2 \\
P\left[m = \frac{3}{2}\right] &= (1-p)^3 = p_3.
\end{aligned} \tag{5.4}$$

The next step is to retrodict the single particle spin $-1/2, 1/2$, or the binary outcome $0, 1$, with the correct probabilities. Using equations in (5.4) we can calculate the conditional probabilities $P(0|m)$ and $P(1|m)$ from the fact that for random discrete variables X, Y the following is true[1]:

$$P(X = x) = \sum_{y \in \mathfrak{Y}} P(X = x | Y = y) P(Y = y). \tag{5.5}$$

We get the following equations from (5.5):

$$\begin{aligned}
P(0) = P(\downarrow) = P(0|m = -3/2)p^3 + P(0|m = -1/2)3p^2(1-p) + \\
+ P(0|m = 1/2)3p(1-p)^2 + P(0|m = 3/2)(1-p)^3 = p
\end{aligned} \tag{5.6}$$

[1]This law is called the law of total probability in probability theory.

and

$$\begin{aligned} P(1) = P(\uparrow) = P(1|m = 3/2)p^3 + P(1|m = 1/2)3p^2(1-p) + \\ + P(1|m = -1/2)3p(1-p)^2 + P(1|m = -3/2)(1-p)^3 = 1 - p. \end{aligned} \tag{5.7}$$

Using symmetry of the two equations with respect to the replacement $P(0) \leftrightarrow P(1), p \leftrightarrow 1-p$ we can rewrite these equations in a more convenient way using real constants a, b, c, d:

$$\begin{aligned} & a \cdot p_0 + b \cdot p_1 + c \cdot p_2 + d \cdot p_3 = \\ & = a \cdot p^3 + 3b \cdot p^2(1-p) + 3c \cdot p(1-p)^2 + d \cdot (1-p)^3 = p \end{aligned} \tag{5.8}$$

After comparing both sides we get the following result:

$$a = 1, \quad b = \frac{2}{3}, \quad c = \frac{1}{3}, \quad d = 0 \quad \forall p \in (0, 1). \tag{5.9}$$

Measuring the spin of all three particles gives us the following equations for probability of measuring the down spin:

$$P(0) = P(\downarrow) = \frac{3p_0 + 2p_1 + p_2}{3} = p. \tag{5.10}$$

For the probability of measuring the up spin we get:

$$P(1) = P(\uparrow) = \frac{3p_3 + 2p_2 + p_1}{3} = 1 - p. \tag{5.11}$$

This channel is depicted in the figure 5.1:

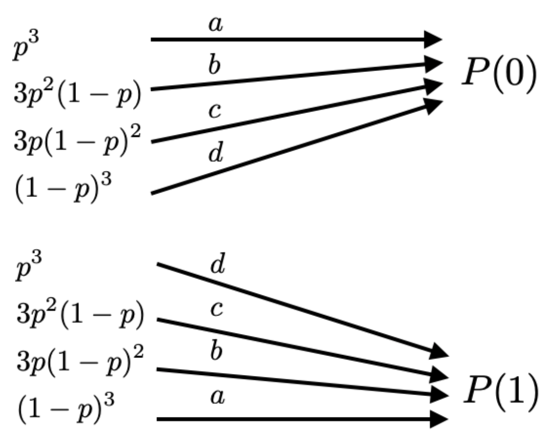

Figure 5.1: Depiction of the channel that maps outcomes $m \in M$ to either $P(0)$ or $P(1)$.

The observation that a, b, c, d form an arithmetic sequence will be generalized to N copies and prove analytically in the next chapter.

5.2 Projection with replacement

The previous case did not offer anything new compared to the measurement of the two-particle system; in fact it compared even worse in terms of the number of measurements one had to perform in order to get the same amount of information. Our second model will be designed differently. We will pick a particle from our system at random, but this time, after measurement, we put it back to the system. Only after that will a new measurement take place. This way we have only two outcomes, instead of four, so there is no need for any channel that maps the outcomes to a two-element set. The measurement operators would take the following form:

$$\begin{aligned}\hat{A}_0^{(1,2,3)} &= \left\{ \frac{1}{\sqrt{3}} \left(\hat{P}_0 \otimes \hat{I} \otimes \hat{I} \right), \frac{1}{\sqrt{3}} \left(\hat{I} \otimes \hat{P}_0 \otimes \hat{I} \right), \frac{1}{\sqrt{3}} \left(\hat{I} \otimes \hat{I} \otimes \hat{P}_0 \right) \right\} \\ \hat{A}_1^{(1,2,3)} &= \left\{ \frac{1}{\sqrt{3}} \left(\hat{P}_1 \otimes \hat{I} \otimes \hat{I} \right), \frac{1}{\sqrt{3}} \left(\hat{I} \otimes \hat{P}_1 \otimes \hat{I} \right), \frac{1}{\sqrt{3}} \left(\hat{I} \otimes \hat{I} \otimes \hat{P}_1 \right) \right\}.\end{aligned} \tag{5.12}$$

We can see that the individual operators no longer commute with permutation of particles. Nevertheless, the combined quantum operation assigned to outcome 0 or 1 via (1.33) does, which makes this a physically realizable measurement.

We know that the probabilities for each outcome of such measurement are the same compared to the previous case since both POVM sets made from (5.3) and (5.12) give the same results, as shown in the following equations. First we will calculate the POVM elements similarly to the situation shown in the first chapter in (1.29). First for the previous case (5.3):

$$\begin{aligned}\hat{\Pi}_0 &= \frac{3}{3}\hat{A}_0^\dagger \hat{A}_0 + \frac{2}{3}\hat{A}_1^\dagger \hat{A}_1 + \frac{1}{3}\hat{A}_2^\dagger \hat{A}_2 + \frac{0}{3}\hat{A}_3^\dagger \hat{A}_3 = \hat{A}_0 + \frac{2}{3}\hat{A}_1 + \frac{1}{3}\hat{A}_2 \\ \hat{\Pi}_1 &= \frac{3}{3}\hat{A}_3^\dagger \hat{A}_3 + \frac{2}{3}\hat{A}_2^\dagger \hat{A}_2 + \frac{1}{3}\hat{A}_1^\dagger \hat{A}_1 + \frac{0}{3}\hat{A}_0^\dagger \hat{A}_0 = \hat{A}_3 + \frac{2}{3}\hat{A}_2 + \frac{1}{3}\hat{A}_1\end{aligned} \tag{5.13}$$

and second for the case of projection with replacement (5.12):

$$\begin{aligned}\hat{\Pi}_0 &= \hat{A}_0^{(1)\dagger}\hat{A}_0^{(1)} + \hat{A}_0^{(2)\dagger}\hat{A}_0^{(2)} + \hat{A}_0^{(3)\dagger}\hat{A}_0^{(3)} = \frac{1}{\sqrt{3}}\left(\hat{A}_0^{(1)} + \hat{A}_0^{(2)} + \hat{A}_0^{(3)}\right) \\ \hat{\Pi}_1 &= \hat{A}_1^{(1)\dagger}\hat{A}_1^{(1)} + \hat{A}_1^{(2)\dagger}\hat{A}_1^{(2)} + \hat{A}_1^{(3)\dagger}\hat{A}_1^{(3)} = \frac{1}{\sqrt{3}}\left(\hat{A}_1^{(1)} + \hat{A}_1^{(2)} + \hat{A}_1^{(3)}\right).\end{aligned} \tag{5.14}$$

As mentioned in the first chapter, the expectational values of POVM elements are equal to probabilities of the outcomes, as seen in equations in (1.26). Since $\hat{\Pi}_0$ and $\hat{\Pi}_1$ are the same[2] in both (5.13) and (5.14) we get the same probabilities we had in the previous case: $P(0) = p$ and $P(1) = 1 - p$.

5.3 Calculating disturbance for both types of projection

We will now show calculation of system disturbance for both measurements proposed in equations (5.3) and (5.12). First step is to define operators $\hat{A}$ from equation (3.22) and

[2] The equality of both POVM sets can be easily shown by writing the elements in matrix form, which is possible to be extended for the case with N particles.

$\hat{B}$ from equation (3.25). We will make the following choice:

$$\hat{A} = \hat{\sigma}_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \hat{B} = \hat{\sigma}_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{5.15}$$

where $\hat{\sigma}_x, \hat{\sigma}_z$ denotes the Pauli matrices. This means we choose to measure spin in x and z direction respectively. Matrices in (5.15) have the following eigenstates:

$$\begin{aligned} \hat{\sigma}_x: &\qquad |+\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad |-\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix} \\ \hat{\sigma}_z: &\qquad |0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad |1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \end{aligned} \tag{5.16}$$

We also define projectors, needed for both measurement schemes (5.3) and (5.12), on the respective eigenstates as follows:

$$\begin{aligned} \hat{P}_+ &= \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \quad \hat{P}_- = \frac{1}{2}\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \\ \hat{P}_0 &= \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \qquad \hat{P}_1 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}. \end{aligned} \tag{5.17}$$

5.3.1 Disturbance for total spin measurement without replacement

In the beginning of disturbance calculation, operators of the first measurement, that will be applied on the input state, need to be defined. We will denote them as $\hat{A}_i$. These will be constructed correspondingly to (5.3), which means we can write the set $\hat{A}_i$ as follows:

$$\begin{aligned} \hat{A}_0 &= \hat{P}_- \otimes \hat{P}_- \otimes \hat{P}_- \\ \hat{A}_1 &= \hat{P}_+ \otimes \hat{P}_- \otimes \hat{P}_- + \hat{P}_- \otimes \hat{P}_+ \otimes \hat{P}_- + \hat{P}_- \otimes \hat{P}_- \otimes \hat{P}_+ \\ \hat{A}_2 &= \hat{P}_+ \otimes \hat{P}_+ \otimes \hat{P}_- + \hat{P}_+ \otimes \hat{P}_- \otimes \hat{P}_+ + \hat{P}_- \otimes \hat{P}_+ \otimes \hat{P}_+ \\ \hat{A}_3 &= \hat{P}_+ \otimes \hat{P}_+ \otimes \hat{P}_+. \end{aligned} \tag{5.18}$$

We now choose the input state as $|\phi\rangle = |000\rangle$; this choice will be denoted as $b = 0$. If we were to choose $|\phi\rangle = |111\rangle$ we would denote it as $b = 1$. After calculating $\hat{A}_i\,|000\rangle$ we get the following outcomes:

$$\begin{aligned} \hat{A}_0\,|000\rangle &= \frac{1}{\sqrt{2}^3}\,|- - -\rangle \\ \hat{A}_1\,|000\rangle &= \frac{1}{\sqrt{2}^3}\,(|+ - -\rangle + |- + -\rangle + |- - +\rangle) \\ \hat{A}_2\,|000\rangle &= \frac{1}{\sqrt{2}^3}\,(|+ + -\rangle + |+ - +\rangle + |- + +\rangle) \\ \hat{A}_3\,|000\rangle &= \frac{1}{\sqrt{2}^3}\,|- + ++\rangle\,. \end{aligned} \tag{5.19}$$

with respective probabilities for the choice $b = 0$ being equal to the vector norms:

$$P(i=0) = \frac{1}{8}, \quad P(i=1) = \frac{3}{8}, \quad P(i=2) = \frac{3}{8}, \quad P(i=3) = \frac{1}{8}. \tag{5.20}$$

In the next step we define the second measurement, which will be labeled as $\hat{B}_j$:

$$\begin{aligned}
\hat{B}_0 &= \hat{P}_0 \otimes \hat{P}_0 \otimes \hat{P}_0 \\
\hat{B}_1 &= \hat{P}_1 \otimes \hat{P}_0 \otimes \hat{P}_0 + \hat{P}_0 \otimes \hat{P}_1 \otimes \hat{P}_0 + \hat{P}_0 \otimes \hat{P}_0 \otimes \hat{P}_1 \\
\hat{B}_2 &= \hat{P}_1 \otimes \hat{P}_1 \otimes \hat{P}_0 + \hat{P}_1 \otimes \hat{P}_0 \otimes \hat{P}_1 + \hat{P}_0 \otimes \hat{P}_1 \otimes \hat{P}_1 \\
\hat{B}_3 &= \hat{P}_1 \otimes \hat{P}_1 \otimes \hat{P}_1
\end{aligned} \tag{5.21}$$

which is going to be used for measuring the first normalized state from (5.19):

$$|---\rangle \tag{5.22}$$

The respective outcomes of $\hat{B}_j\,|---\rangle$ are summarized in the table 5.1. After repeating the same calculation on all intermediate states (5.19), we get the outcomes for $P(j|i)$ which are summarized in the table 5.2.

	States after measuring $\hat{B}_j$	Respective probabilities $P(j\|b=0, i=0)$
$\hat{B}_0\,\|---\rangle$	$\|000\rangle$	$\frac{1}{8}$
$\hat{B}_1\,\|---\rangle$	$\frac{1}{\sqrt{3}}(\|100\rangle + \|010\rangle + \|001\rangle)$	$\frac{3}{8}$
$\hat{B}_2\,\|---\rangle$	$\frac{1}{\sqrt{3}}(\|110\rangle + \|101\rangle + \|011\rangle)$	$\frac{3}{8}$
$\hat{B}_3\,\|---\rangle$	$\|111\rangle$	$\frac{1}{8}$

Table 5.1: Table of post-measurement states $\hat{B}_j\,|---\rangle\,, j \in \{0,1,2,3\}$ and their respective probabilities $P(j|b=0, i=0)$.

$P(j\|i, b=0)$	$i=0$	$i=1$	$i=2$	$i=3$
$j=0$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$
$j=1$	$\frac{3}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{3}{8}$
$j=2$	$\frac{3}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{3}{8}$
$j=3$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

Table 5.2: Table of probabilities $P(j|i)$ for all values $i \in \{0,1,2,3\}$ and $j \in \{0,1,2,3\}$.

As we can see, probabilities in table 5.2 are symmetrical and the diagonal elements have the same value.

The next step is to process outcomes of the second measurement using the classical channel we applied earlier for mapping outcomes of our measurement (5.3) to a two-element set $\{0,1\}$. We get the following equations:

$$\begin{aligned}
&P(b'=0|b=0)= \\
&=[a\cdot P(j=0|b)+b\cdot P(j=1|b)+c\cdot P(j=2|b)+d\cdot P(j=3|b)]|_{b=0}=\frac{1}{2} \\
&P(b'=1|b=0)= \\
&=[d\cdot P(j=0|b)+b\cdot P(j=1|b)+c\cdot P(j=2|b)+a\cdot P(j=3|b)]|_{b=0}=\frac{1}{2}
\end{aligned} \tag{5.23}$$

where a,b,c,d are taken from (5.9. Thanks to the symmetry of the problem we can deduce $P(b'=0|b=0)=P(b'=0|b=1)$ and $P(b'=1|b=0)=P(b'=1|b=1)$. Those values are independent of i.

The paper [8] suggests using the prior $P(b)=\frac{1}{d}$, in our case $d=2$. Knowing this and having conditional probabilities $P(b'|b)$, we can calculate $P(b',b)$ and $P(b|b')$, which incidentally coincides with $P(b'|b)$ due to the symmetry. The overall disturbance then takes the following form:

$$D(M,\hat{B}_j)=H\left(\frac{1}{2},\frac{1}{2}\right)=-\left(\frac{1}{2}\log\frac{1}{2}+\frac{1}{2}\log\frac{1}{2}\right)=1\,\text{bit}. \tag{5.24}$$

This outcome tells us that in this measurement protocol, measuring $\hat{A}=\sigma_x$ with subsequent measurement of $\hat{B}=\sigma_z$, disturbs our system maximally, leaving no information about the input state. There is thus no improvement compared to a single-particle measurement of the same observables. For these reasons we will, from now on, focus more on the second measurement (5.12), although similar steps will be taken when calculating disturbance for projector weakening, presented at the end of this chapter.

5.3.2 Disturbance for total spin measurement with replacement

The steps leading to the outcome of system disturbance in projection with replacement are analogous to the previous case. We first define operators $\hat{A}_i$ as follows:

$$\begin{aligned}
\hat{A}_0^{(1,2,3)}&=\left\{\frac{1}{\sqrt{3}}\left(\hat{P}_-\otimes\hat{I}\otimes\hat{I}\right),\frac{1}{\sqrt{3}}\left(\hat{I}\otimes\hat{P}_-\otimes\hat{I}\right),\frac{1}{\sqrt{3}}\left(\hat{I}\otimes\hat{I}\otimes\hat{P}_-\right)\right\} \\
\hat{A}_1^{(1,2,3)}&=\left\{\frac{1}{\sqrt{3}}\left(\hat{P}_+\otimes\hat{I}\otimes\hat{I}\right),\frac{1}{\sqrt{3}}\left(\hat{I}\otimes\hat{P}_+\otimes\hat{I}\right),\frac{1}{\sqrt{3}}\left(\hat{I}\otimes\hat{I}\otimes\hat{P}_+\right)\right\}
\end{aligned} \tag{5.25}$$

and take the same input state as previously: $|\phi\rangle=|000\rangle$. Applying $\hat{A}_i^{(1,2,3)}$ on the state $|000\rangle$ leads to the following unnormalized intermediate states:

$$\begin{aligned}
\hat{A}_0^{(1)}|000\rangle&=\frac{1}{\sqrt{6}}|{-00}\rangle\,, \quad \hat{A}_0^{(2)}|000\rangle=\frac{1}{\sqrt{6}}|0-0\rangle\,, \quad \hat{A}_0^{(3)}|000\rangle=\frac{1}{\sqrt{6}}|00-\rangle \\
\hat{A}_1^{(1)}|000\rangle&=\frac{1}{\sqrt{6}}|{+00}\rangle\,, \quad \hat{A}_1^{(2)}|000\rangle=\frac{1}{\sqrt{6}}|0+0\rangle\,, \quad \hat{A}_1^{(3)}|000\rangle=\frac{1}{\sqrt{6}}|00+\rangle
\end{aligned} \tag{5.26}$$

with probabilities for each outcome being equal to $\frac{1}{6}$. We now apply the second set of measurement operators $\hat{B}_j^{(1,2,3)}$, on the first normalized intermediate outcome $|{-00}\rangle$, in the following form:

$$\begin{aligned}\hat{B}_0^{(1,2,3)} &= \left\{\frac{1}{\sqrt{3}}\left(\hat{P}_0\otimes\hat{I}\otimes\hat{I}\right), \frac{1}{\sqrt{3}}\left(\hat{I}\otimes\hat{P}_0\otimes\hat{I}\right), \frac{1}{\sqrt{3}}\left(\hat{I}\otimes\hat{I}\otimes\hat{P}_0\right)\right\} \\ \hat{B}_1^{(1,2,3)} &= \left\{\frac{1}{\sqrt{3}}\left(\hat{P}_1\otimes\hat{I}\otimes\hat{I}\right), \frac{1}{\sqrt{3}}\left(\hat{I}\otimes\hat{P}_1\otimes\hat{I}\right), \frac{1}{\sqrt{3}}\left(\hat{I}\otimes\hat{I}\otimes\hat{P}_1\right)\right\}\end{aligned} \tag{5.27}$$

which yield post-measurement states with respective probabilities summarized in the table 5.3.

	States after measuring $\hat{B}_j$	Respective probabilities $P(j\|b=0,i=0)$
$\hat{B}_0^{(1)}\left\|-00\right\rangle$	$\left\|000\right\rangle$	$\frac{1}{6}$
$\hat{B}_0^{(2)}\left\|-00\right\rangle$	$\left\|-00\right\rangle$	$\frac{1}{3}$
$\hat{B}_0^{(3)}\left\|-00\right\rangle$	$\left\|-00\right\rangle$	$\frac{1}{3}$
$\hat{B}_1^{(1)}\left\|-00\right\rangle$	$\left\|100\right\rangle$	$\frac{1}{6}$
$\hat{B}_1^{(2)}\left\|-00\right\rangle$	–	0
$\hat{B}_1^{(3)}\left\|-00\right\rangle$	–	0

Table 5.3: Table of post-measurement states $\hat{B}_j^{(1,2,3)}\left|-00\right\rangle, j\in\{0,1\}$ and their respective probabilities $P(j|b=0,i=0)$. The last two lines are empty since we are measuring $\hat{P}_1\left|0\right\rangle$.

The previous case was more complicated compared to this one because we had 4 outcomes that needed to be mapped on either 0 or 1 using (5.9). This time there is no need for such procedure because summing first three probabilities from the table 5.3 gives the conditional probability $P(b'=0|b=0,i=0)=\frac{5}{6}$ and summing the last three probabilities from the table 5.3 gives $P(b'=1|b=0=\frac{1}{6}$. After more calculations we can put together table 5.4 which summarizes all conditional probabilities $P(b'|b)$.

$P(b'\|b)$	$b=0$	$b=1$
$b'=0$	$\frac{5}{6}$	$\frac{1}{6}$
$b'=1$	$\frac{1}{6}$	$\frac{5}{6}$

Table 5.4: Table of probabilities $P(b'|b)$ for all values $b\in\{0,1\}$ and $b'\in\{0,1\}$.

As we can see, the table came out symmetrical, just as in the previous case. The disturbance can then be calculated as follows:

$$D(M,\hat{B}_i) = H\left(\frac{5}{6},\frac{1}{6}\right) = 0.65\,\mathrm{bit}. \tag{5.28}$$

So by changing the measurement scheme slightly we managed to significantly reduce disturbance.[3]

[3]Concerning noise: observables $\hat{A}_i$ are noiseless according to Definition 6.

5.4 Weakening of the projectors

We presented two similar ways to conduct a measurement on a set of three identical particles. The main differences between both measurements shown in equations (5.3) and (5.12) were discussed, mainly in terms of disturbance.

Another approach to the measurement construction would be to keep the first scheme but weaken the overall influence of measurement, resulting in smaller disturbance. This means that the outcomes are going to be harder to distinguish, compared with the previous cases where we used projectors, but since we posses copies of the original state we can carry out the measurement several times and try to compensate in terms of the overall information gain. This approach requires using effects instead of projectors. Let us first show how such measurement would work on one particle and then extend the problem to a three-particle case. The N-particle case will be explored in the next chapter.

The measurement operators, or effects, for the one-particle case can be constructed as follows:

$$\begin{aligned}\hat{Q}_0 &= \mu\hat{I} + \nu\hat{P}_0 \\ \hat{Q}_1 &= \mu\hat{I} + \nu\hat{P}_1,\end{aligned} \tag{5.29}$$

where w.l.o.g. $\mu, \nu \in \langle 0, 1\rangle$. Measurement operators can then be defined as the effects themselves:

$$\begin{aligned}\hat{A}_0 &= \hat{Q}_0 \\ \hat{A}_1 &= \hat{Q}_1.\end{aligned} \tag{5.30}$$

Calculating the respective probabilities first requires a POVM set, which can be written as follows:

$$\begin{aligned}\hat{\Pi}_0 &= \hat{A}_0^\dagger\hat{A}_0 = u\hat{I} + v\hat{P}_0 \\ \hat{\Pi}_1 &= \hat{A}_1^\dagger\hat{A}_1 = u\hat{I} + v\hat{P}_1,\end{aligned} \tag{5.31}$$

where we used the following substitution:

$$\begin{aligned}u &= \mu^2 \\ v &= \nu^2 + 2\mu\nu\end{aligned} \tag{5.32}$$

so that the we can use a shorter form for the following calculation. In order for the measurement operators to be defined correctly we need to determine u, v so that $\sum_i \hat{A}_i^\dagger\hat{A}_i = \hat{I}$:

$$\hat{Q}_0^\dagger\hat{Q}_0 + \hat{Q}_1^\dagger\hat{Q}_1 = u\hat{I} + v\hat{P}_0 + u\hat{I} + v\hat{P}_1 \tag{5.33}$$

from which we can make the following constraints on u, v:

$$2u + v \stackrel{!}{=} 1. \tag{5.34}$$

Using the substitution defined in (5.32), we can write:

$$2\mu^2 + 2\mu\nu + \nu^2 = 1. \tag{5.35}$$

Measurement operators (5.30) under the constraint (5.35) are constructed correctly, since the POVM elements are both positive operators and they sum to $\hat{I}$, meaning we can now write:

$$q = u + vp, \quad 1 - q = u + v(1 - p) \tag{5.36}$$

for an input state in the form $|\psi\rangle = \alpha |0\rangle + \beta |1\rangle$.

We are no longer using projectors so there is no way for us to assign zero or unitary probability to either of $\{0, 1\}$ outcome[4]. In other words, the distribution (5.36) contains slightly less information than $\{p, 1-p\}$. How to at least approach the original probability distribution? We will demonstrate on a three particle case with measurement operators constructed as follows:[5]

$$\begin{aligned}
&\hat{A}_0 = \hat{Q}_0 \otimes \hat{Q}_0 \otimes \hat{Q}_0 \\
&\hat{A}_1^{(1,2,3)} = \left\{ \left(\hat{Q}_1 \otimes \hat{Q}_0 \otimes \hat{Q}_0\right), \left(\hat{Q}_0 \otimes \hat{Q}_1 \otimes \hat{Q}_0\right), \left(\hat{Q}_0 \otimes \hat{Q}_0 \otimes \hat{Q}_1\right) \right\} \\
&\hat{A}_2^{(1,2,3)} = \left\{ \left(\hat{Q}_1 \otimes \hat{Q}_1 \otimes \hat{Q}_0\right), \left(\hat{Q}_1 \otimes \hat{Q}_0 \otimes \hat{Q}_1\right), \left(\hat{Q}_0 \otimes \hat{Q}_1 \otimes \hat{Q}_1\right) \right\}. \\
&\hat{A}_3 = \hat{Q}_1 \otimes \hat{Q}_1 \otimes \hat{Q}_1.
\end{aligned} \tag{5.38}$$

We again have to correctly define both μ, ν. If it weren't for the previous one-particle case we would need to do a tedious calculation, but the following implication is going to simplify our problem:

$$\hat{Q}_0^\dagger \hat{Q}_0 + \hat{Q}_1^\dagger \hat{Q}_1 = \hat{I} \implies \sum_i \sum_j \hat{A}_i^{(j)\dagger} \hat{A}_i^{(j)} = \hat{I} \tag{5.39}$$

which is true for the measurement operators $\hat{A}_i^{(j)}$ in (5.38). Now that we know what values can be assigned to both μ, ν we can calculate the respective probabilities which take the following form:

$$\begin{aligned}
P(0) &= q^3 \\
P(1) &= 3q^2(1 - q) \\
P(2) &= 3q(1 - q)^2 \\
P(3) &= (1 - q)^3
\end{aligned} \tag{5.40}$$

[4]This statement is true unless we assign zero value to u. Such case would then lead to the previous cases of projection.

[5]The reason we chose a form similar to (5.12) instead of (5.3) is the following: the effects constructed as

$$\begin{aligned}
\hat{A}_0 &= \hat{Q}_0 \otimes \hat{Q}_0 \otimes \hat{Q}_0 \\
\hat{A}_1 &= \hat{Q}_1 \otimes \hat{Q}_0 \otimes \hat{Q}_0 + \hat{Q}_0 \otimes \hat{Q}_1 \otimes \hat{Q}_0 + \hat{Q}_0 \otimes \hat{Q}_0 \otimes \hat{Q}_1 \\
\hat{A}_2 &= \hat{Q}_1 \otimes \hat{Q}_1 \otimes \hat{Q}_0 + \hat{Q}_1 \otimes \hat{Q}_0 \otimes \hat{Q}_1 + \hat{Q}_0 \otimes \hat{Q}_1 \otimes \hat{Q}_1 \\
\hat{A}_3 &= \hat{Q}_1 \otimes \hat{Q}_1 \otimes \hat{Q}_1.
\end{aligned} \tag{5.37}$$

do not satisfy the completeness relation under (1.13).

Since the measurement operators take a similar form to those in (5.3) we will try to solve for a, b, c, d in substituted equations (5.8):

$$\begin{aligned} a \cdot q^3 + 3b \cdot q^2(1-q) + 3c \cdot q(1-q)^2 + d \cdot (1-q)^3 &= p = \frac{q-u}{v} \\ d \cdot q^3 + 3c \cdot q^2(1-q) + 3b \cdot q(1-q)^2 + a \cdot (1-q)^3 &= 1-p = \frac{1-q+u}{v}. \end{aligned} \tag{5.41}$$

We get the following result:

$$a = \frac{1-u}{v} = \frac{3-3u}{3v}, \quad b = \frac{2-3u}{3v}, \quad c = \frac{1-3u}{3v}, \quad d = -\frac{u}{v} = \frac{0-3u}{3v}. \tag{5.42}$$

As we can see, the results take form of an arithmetic sequence, just as a, b, c, d in section 5.1 did. This is no coincidence and the proof of this statement will be provided in the next chapter.

We now need to assign (5.42) specific values in order to proceed with our calculations. After solving a, b, c, d in terms of u we can conclude that they can not be considered an information channel unless we make some adjustments, since all four constants need to be at least equal to zero and at most equal to one, which is not the case. The situation is depicted in figure 5.2.

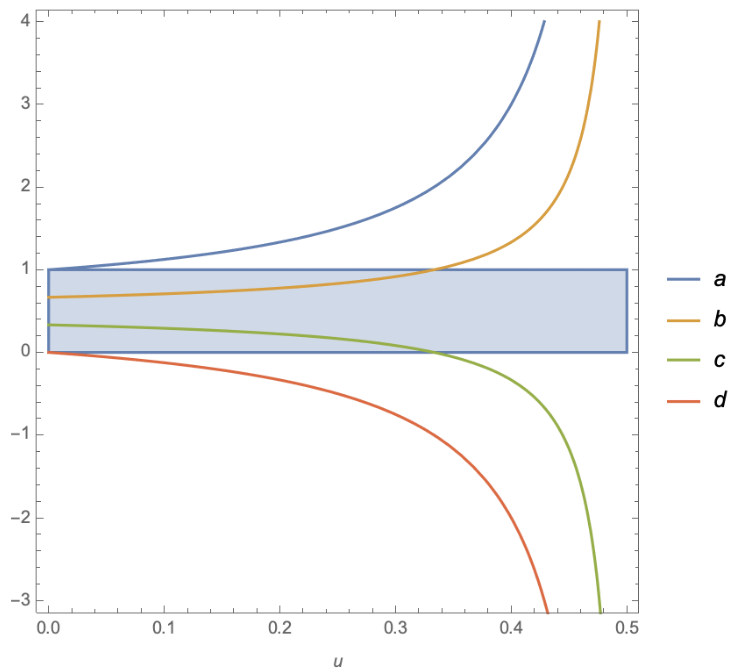

Figure 5.2: Constants a, b, c, d as functions of u. The blue frame depicts the values of $a, b, c, d \in \langle 0, 1\rangle$ where they can be considered a description of an information channel.

The adjustments of a, b, c, d in (5.42) in order for them to become a realistic information channel are discussed at the end of this chapter, section 5.4.2. For now we can assume that they are assigned some values such as to at least approximate an inverse to (5.36).

5.4.1 Disturbance for projector weakening

We will now calculate disturbance for measurement proposed in (5.38). The first step, just as in previous cases, is to write down the measurement operators $\hat{A}_i$ with substituted $\hat{Q}_0 = \mu\hat{I} + \nu\hat{P}_0$ for $\hat{Q}_- = \mu\hat{I} + \nu\hat{P}_-$ and $\hat{Q}_1 = \mu\hat{I} + \nu\hat{P}_1$ for $\hat{Q}_+ = \mu\hat{I} + \nu\hat{P}_+$:

$$\begin{aligned}
&\hat{A}_0 = \hat{Q}_- \otimes \hat{Q}_- \otimes \hat{Q}_- \\
&\hat{A}_1^{(1,2,3)} = \left\{ \left(\hat{Q}_+ \otimes \hat{Q}_- \otimes \hat{Q}_-\right), \left(\hat{Q}_- \otimes \hat{Q}_+ \otimes \hat{Q}_-\right), \left(\hat{Q}_- \otimes \hat{Q}_- \otimes \hat{Q}_+\right) \right\} \\
&\hat{A}_2^{(1,2,3)} = \left\{ \left(\hat{Q}_+ \otimes \hat{Q}_+ \otimes \hat{Q}_-\right), \left(\hat{Q}_+ \otimes \hat{Q}_- \otimes \hat{Q}_+\right), \left(\hat{Q}_- \otimes \hat{Q}_+ \otimes \hat{Q}_+\right) \right\}. \\
&\hat{A}_3 = \hat{Q}_+ \otimes \hat{Q}_+ \otimes \hat{Q}_+.
\end{aligned} \tag{5.43}$$

Applying operators (5.43) on the state $|\phi\rangle = |000\rangle$ yields the following outcomes:

$$\begin{aligned}
\hat{A}_0\,|000\rangle &= \left(\mu\,|0\rangle + \frac{\nu}{\sqrt{2}}\,|-\rangle\right) \otimes \left(\mu\,|0\rangle + \frac{\nu}{\sqrt{2}}\,|-\rangle\right) \otimes \left(\mu\,|0\rangle + \frac{\nu}{\sqrt{2}}\,|-\rangle\right) = \frac{1}{\sqrt{2}^3}\,|\chi_-\chi_-\chi_-\rangle \\
\hat{A}_1^{(1)}\,|000\rangle &= \left(\mu\,|0\rangle + \frac{\nu}{\sqrt{2}}\,|+\rangle\right) \otimes \left(\mu\,|0\rangle + \frac{\nu}{\sqrt{2}}\,|-\rangle\right) \otimes \left(\mu\,|0\rangle + \frac{\nu}{\sqrt{2}}\,|-\rangle\right) = \frac{1}{\sqrt{2}^3}\,|\chi_+\chi_-\chi_-\rangle \\
\hat{A}_1^{(2)}\,|000\rangle &= \left(\mu\,|0\rangle + \frac{\nu}{\sqrt{2}}\,|-\rangle\right) \otimes \left(\mu\,|0\rangle + \frac{\nu}{\sqrt{2}}\,|+\rangle\right) \otimes \left(\mu\,|0\rangle + \frac{\nu}{\sqrt{2}}\,|-\rangle\right) = \frac{1}{\sqrt{2}^3}\,|\chi_-\chi_+\chi_-\rangle \\
\hat{A}_1^{(3)}\,|000\rangle &= \left(\mu\,|0\rangle + \frac{\nu}{\sqrt{2}}\,|-\rangle\right) \otimes \left(\mu\,|0\rangle + \frac{\nu}{\sqrt{2}}\,|-\rangle\right) \otimes \left(\mu\,|0\rangle + \frac{\nu}{\sqrt{2}}\,|+\rangle\right) = \frac{1}{\sqrt{2}^3}\,|\chi_-\chi_-\chi_+\rangle \\
\hat{A}_2^{(1)}\,|000\rangle &= \left(\mu\,|0\rangle + \frac{\nu}{\sqrt{2}}\,|+\rangle\right) \otimes \left(\mu\,|0\rangle + \frac{\nu}{\sqrt{2}}\,|+\rangle\right) \otimes \left(\mu\,|0\rangle + \frac{\nu}{\sqrt{2}}\,|-\rangle\right) = \frac{1}{\sqrt{2}^3}\,|\chi_+\chi_+\chi_-\rangle \\
\hat{A}_2^{(2)}\,|000\rangle &= \left(\mu\,|0\rangle + \frac{\nu}{\sqrt{2}}\,|+\rangle\right) \otimes \left(\mu\,|0\rangle + \frac{\nu}{\sqrt{2}}\,|-\rangle\right) \otimes \left(\mu\,|0\rangle + \frac{\nu}{\sqrt{2}}\,|+\rangle\right) = \frac{1}{\sqrt{2}^3}\,|\chi_+\chi_-\chi_+\rangle \\
\hat{A}_2^{(3)}\,|000\rangle &= \left(\mu\,|0\rangle + \frac{\nu}{\sqrt{2}}\,|-\rangle\right) \otimes \left(\mu\,|0\rangle + \frac{\nu}{\sqrt{2}}\,|+\rangle\right) \otimes \left(\mu\,|0\rangle + \frac{\nu}{\sqrt{2}}\,|+\rangle\right) = \frac{1}{\sqrt{2}^3}\,|\chi_-\chi_+\chi_+\rangle \\
\hat{A}_3\,|000\rangle &= \left(\mu\,|0\rangle + \frac{\nu}{\sqrt{2}}\,|+\rangle\right) \otimes \left(\mu\,|0\rangle + \frac{\nu}{\sqrt{2}}\,|+\rangle\right) \otimes \left(\mu\,|0\rangle + \frac{\nu}{\sqrt{2}}\,|+\rangle\right) = \frac{1}{\sqrt{2}^3}\,|\chi_+\chi_+\chi_+\rangle
\end{aligned} \tag{5.44}$$

where we used the following notation:

$$\begin{aligned}
|\chi_-\rangle &= \sqrt{2}\left(\mu\,|0\rangle + \frac{\nu}{\sqrt{2}}\,|-\rangle\right) \\
|\chi_+\rangle &= \sqrt{2}\left(\mu\,|0\rangle + \frac{\nu}{\sqrt{2}}\,|+\rangle\right)
\end{aligned} \tag{5.45}$$

for simplicity. The condition (5.35) guarantees that these kets are normalized. The second set of measurement operators $\hat{B}_j$ takes the following form:

$$\begin{aligned}
&\hat{B}_0 = \hat{Q}_0 \otimes \hat{Q}_0 \otimes \hat{Q}_0 \\
&\hat{B}_1^{(1,2,3)} = \left\{ \left(\hat{Q}_1 \otimes \hat{Q}_0 \otimes \hat{Q}_0\right), \left(\hat{Q}_0 \otimes \hat{Q}_1 \otimes \hat{Q}_0\right), \left(\hat{Q}_0 \otimes \hat{Q}_0 \otimes \hat{Q}_1\right) \right\} \\
&\hat{B}_2^{(1,2,3)} = \left\{ \left(\hat{Q}_1 \otimes \hat{Q}_1 \otimes \hat{Q}_0\right), \left(\hat{Q}_1 \otimes \hat{Q}_0 \otimes \hat{Q}_1\right), \left(\hat{Q}_0 \otimes \hat{Q}_1 \otimes \hat{Q}_1\right) \right\}. \\
&\hat{B}_3 = \hat{Q}_3 \otimes \hat{Q}_3 \otimes \hat{Q}_3.
\end{aligned} \tag{5.46}$$

Applying the operators $\hat{B}_j$ in (5.46) on the first intermediate state $|\chi_-\chi_-\chi_-\rangle$ yields outcomes which are summarized in the table 5.5. In this case we chose not to include the exact form of the states after measuring $\hat{B}_j\,|\chi_-\chi_-\chi_-\rangle)$ in the table 5.5, simply because their form was too long and they are not needed for the next steps of disturbance calculation.

States after measuring $\hat{B}_j$	Respective probabilities $P(j\|b=0,i=0)$
$\hat{B}_0\,\|\chi_-\chi_-\chi_-\rangle$	$\left(\frac{1+2\mu\nu+2(\mu\nu)^2}{2}\right)^3 = r^3$
$\hat{B}_1^{(1)}\,\|\chi_-\chi_-\chi_-\rangle$	$\left(\frac{1+2\mu\nu+2(\mu\nu)^2}{2}\right)^2\left(\frac{1-2\mu\nu-2(\mu\nu)^2}{2}\right) = r^2(1-r)$
$\hat{B}_1^{(2)}\,\|\chi_-\chi_-\chi_-\rangle$	$\left(\frac{1+2\mu\nu+2(\mu\nu)^2}{2}\right)^2\left(\frac{1-2\mu\nu-2(\mu\nu)^2}{2}\right) = r^2(1-r)$
$\hat{B}_1^{(3)}\,\|\chi_-\chi_-\chi_-\rangle$	$\left(\frac{1+2\mu\nu+2(\mu\nu)^2}{2}\right)^2\left(\frac{1-2\mu\nu-2(\mu\nu)^2}{2}\right) = r^2(1-r)$
$\hat{B}_2^{(1)}\,\|\chi_-\chi_-\chi_-\rangle$	$\left(\frac{1+2\mu\nu+2(\mu\nu)^2}{2}\right)\left(\frac{1-2\mu\nu-2(\mu\nu)^2}{2}\right)^2 = r(1-r)^2$
$\hat{B}_2^{(2)}\,\|\chi_-\chi_-\chi_-\rangle$	$\left(\frac{1+2\mu\nu+2(\mu\nu)^2}{2}\right)\left(\frac{1-2\mu\nu-2(\mu\nu)^2}{2}\right)^2 = r(1-r)^2$
$\hat{B}_2^{(3)}\,\|\chi_-\chi_-\chi_-\rangle$	$\left(\frac{1+2\mu\nu+2(\mu\nu)^2}{2}\right)\left(\frac{1-2\mu\nu-2(\mu\nu)^2}{2}\right)^2 = r(1-r)^2$
$\hat{B}_3\,\|\chi_-\chi_-\chi_-\rangle$	$\left(\frac{1-2\mu\nu-2(\mu\nu)^2}{2}\right)^3 = r^3$

Table 5.5: Table of post-measurement states $\hat{B}_j\,|\chi_-\chi_-\chi_-\rangle), j \in \{0,1,2,3\}$ and their respective probabilities $P(j|b=0,i=0)$.

The next step is to create a table of probabilities $P(j=0,1,2,3|i=0,b=0)$. The only difference, compared to the case of projection covered in section 5.1, is that we first have to sum probabilities of the states $\hat{B}_1^{(1,2,3)}$ and $\hat{B}_2^{(1,2,3)}$. Calculation of these overall probabilities can be found in (5.47).

$$
\begin{aligned}
||\hat{B}_1^(j)\,|\chi_-\chi_-\chi_-\rangle\,|| &= 3r^2(1-r) \\
||\hat{B}_2^(j)\,|\chi_-\chi_-\chi_-\rangle\,|| &= 3r(1-r)^2
\end{aligned}
\tag{5.47}
$$

These calculations leads us to the table 5.6 of conditional probabilities $P(j|i)$.

$P(j\|i)$	$i=0$	$i=1$	$i=2$	$i=3$
$j=0$	r^3	r^3	r^3	r^3
$j=1$	$3r^2(1-r)$	$3r^2(1-r)$	$3r^2(1-r)$	$3r^2(1-r)$
$j=2$	$3r(1-r)^2$	$3r(1-r)^2$	$3r(1-r)^2$	$3r(1-r)^2$
$j=3$	$(1-r)^3$	$(1-r)^3$	$(1-r)^3$	$(1-r)^3$

Table 5.6: Table of probabilities $P(j|i)$ for all values $i \in \{0,1,2,3\}$ and $j \in \{0,1,2,3\}$.

In order to calculate disturbance we need to use a,b,c,d from (5.42) to map outcomes from table 5.6 to a two-element set $\{0,1\}$, which would yield us conditional probabilities $P(b'|b)$ as functions of u and v.[6]. There is a catch though – as we indicated earlier, for example in figure 5.2, for most values u,v is $a>1$ and $d<0$, meaning a,b,c,d can not represent an information channel. We dedicate the next section to this problem.

[6] $P(b'|b)$ would take the following form:

5.4.2 Classical channels as functions of μ, ν

In the ending discussion of this chapter we are going to examine several possible cases, when values of a, b, c, d (5.42) can be considered as realizable classical information channel.

There are two trivial choices. First is to assign the following values to u, v:

$$u = 0, \quad v = 1 \tag{5.49}$$

which corresponds to the previous case of projection seen in (5.3). The other trivial case would be:

$$u = \frac{1}{2}, \quad v = 0. \tag{5.50}$$

Assigning such values of u, v to operators in (5.37) would have no effect on the input state and would also yield no information whatsoever, but mainly such channel (5.42) would be ill defined since we would divide by 0. All the interesting and relevant channels hence lie between those two cases. The correction of values in (5.42) can be done as follows. We will first fix the value of a to the maximal one while fixing the value of d to the minimum and assign a value which lies in $\langle 0, 1\rangle$ to b and c such that $b + c = 1$. For every u we can there exists this corrected channel, but also μ, ν and r can be calculated from it. The values in table 5.5 will then have a fixed value from which disturbance can be calculated, a general calculation can be seen in figure 5.3.

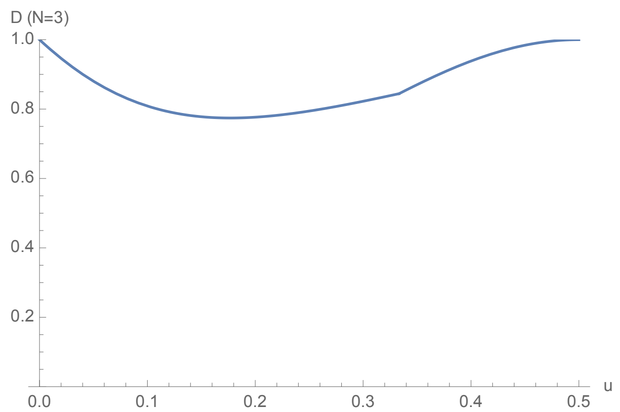

Figure 5.3: Disturbance for the three-particle case as a function of u.

$$\begin{aligned}
&P(b' = 0|b = 0) = P(b' = 1|b = 1) \\
&= [a \cdot P(j = 0|b) + 3b \cdot P(j = 1|b) + 3c \cdot P(j = 2|b) + d \cdot P(j = 3|b)]|_{b=0} = \\
&= \frac{1-u}{v} r^3 + 3\frac{2-3u}{3v} r^2(1-r) + 3\frac{1-3u}{3v} r(1-r)^2 - \frac{u}{v}(1-r)^3 = \frac{r-u}{v} \\
&P(b' = 1|b = 0) = P(b' = 1|b = 0) \\
&[d \cdot P(j = 0|b) + 3b \cdot P(j = 1|b) + 3c \cdot P(j = 2|b) + a \cdot P(j = 3|b)]|_{b=0} = \\
&= -\frac{u}{v} r^3 + 3\frac{1-3u}{3v} r^2(1-r) + 3\frac{2-3u}{3v} r(1-r)^2 + \frac{1-u}{v}(1-r)^3 = \frac{1-r+u}{v}.
\end{aligned} \tag{5.48}$$

Chapter 6

N-particle system disturbance

In this chapter we are going to finalize our work and extend the previous discussions, namely the measurement proposed in projection with replacement and projector weakening, to an N-particle case.

6.1 Projection with replacement for N particles

The measurement operators $A_i^{(1,2,3)}$ presented in (5.12) would extend to the N-particle case as follows:

$$\begin{aligned} \hat{A}_0^{(1,2,\ldots,N)} &= \left\{ \frac{1}{\sqrt{N}} \left(\hat{P}_- \otimes \hat{I} \otimes \cdots \otimes \hat{I} \right), \ldots, \frac{1}{\sqrt{N}} \left(\hat{I} \otimes \hat{I} \otimes \cdots \otimes \hat{P}_- \right) \right\} \\ \hat{A}_1^{(1,2,\ldots,N)} &= \left\{ \frac{1}{\sqrt{N}} \left(\hat{P}_+ \otimes \hat{I} \otimes \cdots \otimes \hat{I} \right), \ldots, \frac{1}{\sqrt{N}} \left(\hat{I} \otimes \hat{I} \otimes \cdots \otimes \hat{P}_+ \right) \right\}. \end{aligned} \tag{6.1}$$

We take the input state in the following form:

$$|\phi\rangle = \underbrace{|\psi\rangle \otimes |\psi\rangle \otimes \ldots |\psi\rangle}_{N \text{ copies}} = |00\ldots 0\rangle \tag{6.2}$$

and apply measurement operators from (6.1). Such operation would yield the following outcomes with probability for each outcome being equal to $\frac{1}{2N}$:

$$\begin{aligned} &\hat{A}_0^{(1)} |\phi\rangle = |{-0}\ldots 0\rangle \quad \hat{A}_0^{(2)} |\phi\rangle = |0 - \ldots 0\rangle \quad \ldots \quad \hat{A}_0^{(N)} |\phi\rangle = |00\ldots -\rangle \\ &\hat{A}_1^{(1)} |\phi\rangle = |{+0}\ldots 0\rangle \quad \hat{A}_1^{(2)} |\phi\rangle = |0 + \ldots 0\rangle \quad \ldots \quad \hat{A}_1^{(N)} |\phi\rangle = |00\cdots +\rangle. \end{aligned} \tag{6.3}$$

In analogy to the three-particle case we now present the second set of measurement operators $\hat{B}_j^{(1,2,\ldots,N)}$:

$$\begin{aligned} \hat{B}_0^{(1,2,\ldots,N)} &= \left\{ \frac{1}{\sqrt{N}} \left(\hat{P}_0 \otimes \hat{I} \otimes \cdots \otimes \hat{I} \right), \ldots, \frac{1}{\sqrt{N}} \left(\hat{I} \otimes \hat{I} \otimes \cdots \otimes \hat{P}_0 \right) \right\} \\ \hat{B}_1^{(1,2,\ldots,N)} &= \left\{ \frac{1}{\sqrt{N}} \left(\hat{P}_1 \otimes \hat{I} \otimes \cdots \otimes \hat{I} \right), \ldots, \frac{1}{\sqrt{N}} \left(\hat{I} \otimes \hat{I} \otimes \cdots \otimes \hat{P}_1 \right) \right\}. \end{aligned} \tag{6.4}$$

which will be used to measure the normalized states (6.3), yielding probabilities summarized in the table 6.1. Using these outcomes we can calculate conditional probabilities $P(b'|b)$, which are summarized in the table 6.2 and are no longer dependent on i.

$\hat{B}_j^{(1,2,\dots,N)}$	Respective probabilities $P(j\|b=0)$
$\hat{B}_0^{(1)} \|{-0\dots0}\rangle$	$\frac{1}{2N}$
$\hat{B}_0^{(2)} \|{-0\dots0}\rangle$	$\frac{1}{N}$
$\vdots$	$\vdots$
$\hat{B}_0^{(N)} \|{-0\dots0}\rangle$	$\frac{1}{N}$
$\hat{B}_1^{(1)} \|{-0\dots0}\rangle$	$\frac{1}{2N}$
$\hat{B}_1^{(2)} \|{-0\dots0}\rangle$	0
$\vdots$	$\vdots$
$\hat{B}_1^{(N)} \|{-0\dots0}\rangle$	0

Table 6.1: Table of post-measurement states $\hat{B}_j^{(1,2,\dots,N)} \left|-0\dots0\right\rangle, j \in \{0,1\}$ and their respective probabilities $P(j|b=0, i=0)$.

$P(b'\|b)$	$b=0$	$b=1$
$b'=0$	$\frac{2N-1}{2N}$	$\frac{1}{2N}$
$b'=1$	$\frac{1}{2N}$	$\frac{2N-1}{2N}$

Table 6.2: Table of probabilities $P(b'|b)$ for all values $b \in \{0,1\}$ and $b' \in \{0,1\}$.

The disturbance can then, in analogy to (5.28) in the previous chapter, be calculated as follows:

$$D(M, \hat{B}_j^{(1,2,\dots,N)}) = H\left(\frac{2N-1}{2N}, \frac{1}{2N}\right). \tag{6.5}$$

We can rewrite $D(M, \hat{B}_j^{(1,2,\dots,N)})$ as follows:

$$D(M, \hat{B}_j^{(1,2,\dots,N)}) = \frac{\log(2N-1) - 2N\log(1-1/N)}{N\log 4}. \tag{6.6}$$

Calculating the disturbance for macroscopic systems means calculating the limit $N \to \infty$ from (6.6):

$$\lim_{N\to\infty} D(M, \hat{B}_j^{(1,2,\dots,N)}) = \lim_{N\to\infty} \frac{\log(2N-1) - 2N\log(1-1/N)}{N\log 4} = 0. \tag{6.7}$$

Just as we expected, the system disturbance in macroscopic world can be neglected. We conclude this chapter by plotting $D(M, \hat{B}_j^{(1,2,\dots,N)})$ as a function of N.

6.2 Projector weakening for N particles

We start again by extending $\hat{A}_i$ from the three-particle case in (5.43) to N:

$$\begin{aligned}
\hat{A}_0 &= \left[\hat{Q}_- \otimes \hat{Q}_- \otimes \cdots \otimes \hat{Q}_-\right] \\
\hat{A}_1^{(1)} &= \left[\hat{Q}_+ \otimes \hat{Q}_- \otimes \cdots \otimes \hat{Q}_-\right] \\
\hat{A}_1^{(2)} &= \left[\hat{Q}_- \otimes \hat{Q}_+ \otimes \cdots \otimes \hat{Q}_-\right] \\
&\vdots \\
\hat{A}_1^{(N)} &= \left[\hat{Q}_- \otimes \hat{Q}_- \otimes \cdots \otimes \hat{Q}_+\right] \\
\hat{A}_2^{(1)} &= \left[\hat{Q}_+ \otimes \hat{Q}_+ \otimes \cdots \otimes \hat{Q}_-\right] \\
&\vdots \\
\hat{A}_2^{(N)} &= \left[\hat{Q}_- \otimes \cdots \otimes \hat{Q}_+ \otimes \hat{Q}_+\right] \\
\hat{A}_3 &= \binom{N}{N} \left[\hat{Q}_+ \otimes \hat{Q}_+ \otimes \cdots \otimes \hat{Q}_+\right].
\end{aligned} \tag{6.8}$$

Applying (6.8) on the input state $|\phi\rangle$ yields the following outcomes:

$$\begin{aligned}
\hat{A}_0 |\phi\rangle &= \frac{1}{\sqrt{2}^N} |\chi_-\chi_- \dots \chi_-\rangle \\
\hat{A}_1^{(1)} |\phi\rangle &= \frac{1}{\sqrt{2}^N} |\chi_+\chi_- \dots \chi_-\rangle \\
\hat{A}_1^{(2)} |\phi\rangle &= \frac{1}{\sqrt{2}^N} |\chi_-\chi_+ \dots \chi_-\rangle \\
&\vdots \\
\hat{A}_1^{(N)} |\phi\rangle &= \frac{1}{\sqrt{2}^N} |\chi_-\chi_- \dots \chi_+\rangle \\
\hat{A}_2^{(1)} |\phi\rangle &= \frac{1}{\sqrt{2}^N} |\chi_+\chi_+ \dots \chi_-\rangle \\
&\vdots \\
\hat{A}_2^{(N)} |\phi\rangle &= \frac{1}{\sqrt{2}^N} |\chi_- \dots \chi_+\chi_+\rangle \\
\hat{A}_3 |\phi\rangle &= \frac{1}{\sqrt{2}^N} |\chi_+ \dots \chi_+\chi_+\rangle .
\end{aligned} \tag{6.9}$$

We again pick the first intermediate state $|\chi_- \chi_- \dots \chi_-\rangle$ as the state on which we apply the following set of measurement operators $\hat{B}_j^{(l)}$, $l \in \{1, 2, \dots N\}$:

$$\begin{aligned}
\hat{B}_0 &= \left[\hat{Q}_0 \otimes \hat{Q}_0 \otimes \cdots \otimes \hat{Q}_0\right] \\
\hat{B}_1^{(1)} &= \left[\hat{Q}_1 \otimes \hat{Q}_0 \otimes \cdots \otimes \hat{Q}_0\right] \\
\hat{B}_1^{(2)} &= \left[\hat{Q}_0 \otimes \hat{Q}_1 \otimes \cdots \otimes \hat{Q}_0\right] \\
&\vdots \\
\hat{B}_1^{(N)} &= \left[\hat{Q}_0 \otimes \hat{Q}_0 \otimes \cdots \otimes \hat{Q}_1\right] \\
\hat{B}_2^{(1)} &= \left[\hat{Q}_1 \otimes \hat{Q}_1 \otimes \cdots \otimes \hat{Q}_0\right] \\
&\vdots \\
\hat{B}_2^{(N)} &= \left[\hat{Q}_0 \otimes \cdots \otimes \hat{Q}_1 \otimes \hat{Q}_1\right] \\
\hat{B}_3 &= \left[\hat{Q}_1 \otimes \hat{Q}_1 \otimes \cdots \otimes \hat{Q}_1\right].
\end{aligned} \tag{6.10}$$

which yield the outcomes summarized in the table 6.3, where we again used the substitution $r = \frac{1+2\mu\nu+2(\mu\nu)^2}{2}$.

$j \in \{0, 1, 2, \dots N\}$	$P(j\|i = 0)$
$j = 0$	$\binom{N}{0} r^N$
$j = 1$	$\binom{N}{1} r^{(N-1)}(1 - r)$
$\vdots$	$\vdots$
$j = N$	$\binom{N}{N}(1 - r)^N$

Table 6.3: Table of probabilities $P(j|i = 0)$ for all values $j \in \{0, 1, 2, \dots N\}$.

For disturbance calculation we first need a mapping of the outcomes in table 6.3 onto the set $\{0, 1\}$, just as we did in the three-particle case. Here we present a proof of the statement that a, b, c, d in (5.9) and (5.42) are indeed in a form of arithmetic sequence. We will do so only for the case of projector weakening since the measurement scheme presented in section 5.1 is only a special case for $u = 0$ and $v = 1$.

Measuring the intermediate state $|\chi_- \chi_- \dots \chi_-\rangle$ with $\hat{B}_j^{(1,2,\dots,N)}$ each yield either 0 or 1 in terms of the binary outcome with the following probability distribution:

$$\begin{aligned}
&j = 0 \quad \text{with probability} \quad r^N \\
&j = 1 \quad \text{with probability} \quad N \cdot r^{N-1}(1 - r) \\
&\quad\vdots \\
&j = k \quad \text{with probability} \quad \binom{N}{k} \cdot r^{N-k}(1 - r)^k \\
&\quad\vdots \\
&j = N \quad \text{with probability} \quad (1 - r)^N
\end{aligned} \tag{6.11}$$

where $j \in N$. Our hypothesis based on (5.42) is that the conditional probabilities $P(0|j)$ and $P(1|j)$ should take the following form:

$$\begin{aligned} P(0|j) &= \frac{(N-j) - N \cdot u}{N \cdot v} \\ P(1|j) &= \frac{j - N \cdot u}{N \cdot v}. \end{aligned} \tag{6.12}$$

If our assumption about the channel mapping taking form of an arithmetic sequence is true, then the outcome should yield p and $1-p$ respectively for an arbitrary q when we write them in a form of (5.5):

$$\begin{aligned} \sum_{j=0}^{N} P(0|j)P(j) &= \sum_{j=0}^{N} \frac{(N-j) - N \cdot u}{N \cdot v} \binom{N}{j} \cdot q^{N-j} (1-q)^j = \\ = \sum_{j=0}^{N} \frac{N-j}{N \cdot v} \binom{N}{j} q^{N-j} (1-q)^j &- \frac{u}{v} \sum_{j=0}^{N} \binom{N}{j} q^{N-j} (1-q)^j = \\ = \sum_{j=0}^{N} \frac{N-j}{N \cdot v} \frac{N!}{j!(N-j)!} q^{N-j} (1-q)^j &- \frac{u}{v}(q+1-q)^N = \\ = \frac{1}{v} \sum_{j=0}^{N} \frac{(N-1)!}{j!(N-j-1)!} q^{N-j} (1-q)^j &- \frac{u}{v} = \\ = \frac{1}{v} \sum_{j=0}^{N} \binom{N-1}{j} q^{N-j} (1-q)^j &- \frac{u}{v} = \\ = \frac{q}{v}(q+1-q)^{N-1} - \frac{u}{v} &= \frac{q-u}{v} = p. \end{aligned}$$

We can now move on to calculating the disturbance. Since we know that the channel (6.12) used for the previous calculation is an idealized one and can not be achieved in reality, we need to make several adjustment. A real channel would take the following form for $P(0|j)$:

$$\begin{aligned} P(0|j) &= \frac{N - j - N \cdot u}{N \cdot v} \quad \text{if} \in \langle 0, 1 \rangle \\ P(0|j) &= 0 \quad \text{if} < 0 \\ P(0|j) &= 1 \quad \text{if} > 1. \end{aligned} \tag{6.13}$$

$P(1|j)$ can be constructed in analogy to (6.13). As we can see, (6.13) is a function of both N and u, from which we can calculate all other variables used in previous calculations:

$$u \in \left(0, \frac{1}{2}\right), \quad v = 1 - 2u, \quad \mu = \sqrt{u}, \quad \nu = \sqrt{u+v} - \sqrt{u}. \tag{6.14}$$

Calculating the disturbance then means putting the probabilities from table 6.2 through the above channel, obtaining $P(b' = 0|b = 0)$ and $P(b' = 1|b = 0$ and calculating the conditional entropy $H(B'|B)$ as before. We show this in the two limit cases of section 5.4.2. The limit $u \to 0$ would mean, as we have shown in section 5.4.2, the first case of simple projection, for which we calculated $D(M, \hat{B}_j) = 1$, see (5.24).

The other limit, $u \to \frac{1}{2}$, would mean, that r gets "worse", meaning disturbance would approach 1 bit. On the other hand, $\frac{r-u}{v} \to 1$, which would cause the disturbance to approach 0 bits. This is the value that would be obtained if a channel described by (6.12) were available, but it coincides with the result of the modified channel in the limit $N \to \infty$. This then results in both disturbance and error to approach 0 for $N \to \infty$. Using the values of the "unreal" channel (6.12), we can plot disturbance as a function of N, which is depicted in figure 6.1.

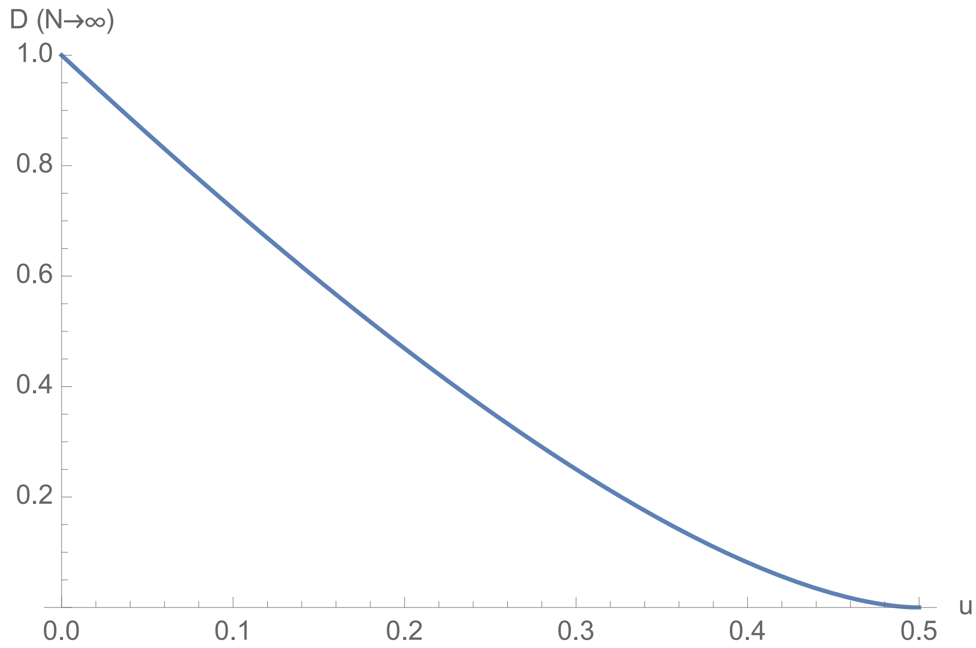

Figure 6.1: The N-particle system disturbance as a function of N.

To conclude this chapter, with $N \to \infty$ copies we can expand the $\hat{A}, \hat{B}$ measurements such that $\hat{A}$ is measured without error and $\hat{B}$ without disturbance in either scheme. For the section 6.1, the results we obtained are exact. For the section 6.2, cases with $N < \infty$ have to be computed numerically.

Conclusion

The main aim of this work was to present the theory behind quantum measurements and study how noise and disturbance affects quantum systems with respect to different measurement schemes. We studied several cases and discussed how the situation with respect to disturbances changes for each of them, starting with measurements on one-particle systems. The one-particle case was then extended for three-particles in order to show the process of measuring several copies of one particle. The measurement schemes presented in the three-particle case was discussed in great detail in chapter five. We conclude our work with case of having N copies of a quantum state and investigating how the disturbance changes with regards to sending $N \to \infty$. We found that for the measurement scheme which uses replacement the disturbance has a specific value as a function of N and goes to zero with N approaching ∞. For the second important case, the measurement scheme where we used projector weakening, the disturbance is a rather complicated function of N which nonetheless approaches 0 for $N \to \infty$.

Bibliography

[1] Michael A Nielsen and Isaac Chuang. Quantum computation and quantum information, 2002.

[2] Paul Busch, Marian Grabowski, and Pekka J Lahti. *Operational quantum physics*, volume 31. Springer Science & Business Media, 1997.

[3] Claude Elwood Shannon. A mathematical theory of communication. *Bell system technical journal*, 27(3):379–423, 1948.

[4] Karl Kraus. *States, effects and operations: fundamental notions of quantum theory*. Springer, 1983.

[5] Thomas M Cover and Joy A Thomas. *Elements of information theory*. John Wiley & Sons, 2012.

[6] Mark M Wilde. *Quantum information theory*. Cambridge University Press, 2013.

[7] Robert B Ash. *Basic probability theory*. Courier Corporation, 2008.

[8] Francesco Buscemi, Michael JW Hall, Masanao Ozawa, and Mark M Wilde. Noise and disturbance in quantum measurements: an information-theoretic approach. *Physical review letters*, 112(5):050401, 2014.

[9] Masanao Ozawa. Universally valid reformulation of the heisenberg uncertainty principle on noise and disturbance in measurement. *Physical Review A*, 67(4):042105, 2003.

[10] Masanao Ozawa. Position measuring interactions and the heisenberg uncertainty principle. *Physics Letters A*, 299(1):1–7, 2002.