



TECHNISCHE  
UNIVERSITÄT  
WIEN  
Vienna University of Technology

## DIPLOMARBEIT

### Entropic Noise-Disturbance Measurements with a Three-Output POVM in Neutron Polarimetry

ausgeführt zum Zwecke der Erlangung des akademischen Grades eines  
Diplom-Ingenieurs in Technischer Physik

Hauptbetreuung

Associate Prof. Dipl.-Ing. Dr. Yuji Hasegawa

Mitbetreuung

Dipl.-Ing. Dr. Stephan Sponar

**E141 – Atominstitut**

eingereicht an der Technischen Universität Wien  
Fakultät für Physik

von

**Vito Fabian Pecile BSc.**

Matr. Nr. 01326485

Schwenkgasse 7/5

1120 Wien

---

(Ort, Datum)

---

(Unterschrift Diplomand)

---

(Unterschrift Betreuer)



Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.  
The approved original version of this thesis is available in print at TU Wien Bibliothek.

The uncertainty principle introduced by *Werner Heisenberg* in 1927 [1] was the first scientific work addressing the inability of measuring two non-commuting observables of a quantum system sequentially with arbitrarily high precision. The principle was shown heuristically by the means of the so called  $\gamma$ -ray microscope Gedankenexperiment. The first mathematically rigorous formulation of the uncertainty principle was introduced by *Earl Hesse Kennard* only some months later, leading to the most common formulation of uncertainty known nowadays, being  $\sigma(Q)\sigma(P) \geq \frac{\hbar}{2}$  [2]. To describe the tradeoff relation between the two observables, he used the standard deviation  $\sigma$ .

The usage of the *Shannon entropy* as description for uncertainty was introduced by *David Deutsch* in 1983 [3]. He argued, that by using entropic definitions one can achieve a *tighter* bound for the tradeoff relation, meaning that the constant limiting the uncertainty becomes smaller.

This thesis deals with the experimental investigation of entropic noise-disturbance measurement uncertainty relations. The definitions of noise and disturbance follow a modern characterization by means of Shannon entropy, as introduced by *Buscemi et al.* in 2014 [4]. We use neutron polarimetry to show the principle for qubit systems. The spin of the neutron is the quantum system under investigation. A 3-outcome POVM measurement is performed as proposed by *Abbott and Branciard* in 2016 [5]. They argue, that this leads to a tight uncertainty relation for qubit measurements, outperforming the results that can be achieved by projective measurements.

We were able to show that a tight noise-disturbance measurement uncertainty relation for qubits can be achieved with a 3-output POVM using neutron polarimetry. The measurement results are significantly tighter than previous experiments that used projective measurements [6].



Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.  
The approved original version of this thesis is available in print at TU Wien Bibliothek.

Die Unschärferelation war die erste wissenschaftliche Arbeit, welche sich mit der Unmöglichkeit der beliebig exakten Messung zweier nicht-kommutierender Observablen eines quantenmechanischen Systems beschäftigte. Sie wurde von *Werner Heisenberg* im Jahre 1927 eingeführt [1]. Er verwendete das sogenannte  $\gamma$ -Strahlen-Mikroskop-Gedankenexperiment, um das Prinzip heuristisch einzuführen. Die erste mathematisch rigorose Ableitung des Prinzips folgte wenige Monate später von *Earl Hesse Kennard*, welcher die Varianz  $\sigma$  einführte um die Ungleichung zu beschreiben. Dies führte zur heute wohl bekanntesten Formulierung eines Unschärfepinzips,  $\sigma(Q)\sigma(P) \geq \frac{\hbar}{2}$  [2].

Die Verwendung der Shannon Entropie für eine Definition der Unschärfe wurde von *David Deutsch* 1983 eingeführt [3]. Er argumentierte, dass man mit Hilfe entropischer Definitionen eine kleinere Konstante für das Produkt der Unschärfe der beiden Observablen erhalten würde und diese somit geringer sei.

In dieser Diplomarbeit werden entropische Noise-Disturbance Unschärferelationen experimentell untersucht. Die Definitionen von Noise und Disturbance folgen einer modernen Charakterisierung, welche die Shannon-Entropie zur Quantifizierung verwendet. Sie wurde 2014 von *Buscemi et al.* eingeführt [4]. Wir verwenden Neutronenpolarimetrie um das Prinzip für Qubit-Systeme zu zeigen. Konkret ist der Neutronenspin das untersuchte Quantensystem. Eine POVM Messung mit drei Ausgängen wird verwendet. Diese wurde von *Abbott and Branciard* 2016 vorgeschlagen [5]. Sie argumentieren, dass man mit Hilfe dieser Messvorschrift eine optimal geringe Konstante für die Unschärferelation erhält und die Ergebnisse gegenüber projektiven Messvorschriften verbessert werden können.

Es war möglich zu zeigen, dass mit Hilfe einer POVM Messvorschrift mit 3 Ausgängen eine dichte Unschärferelation für Qubits unter Verwendung von Neutronenpolarimetrie erhalten werden kann. Die Messpunkte liegen signifikant unter jenen vorhergehender Messungen, welche eine projektive Messvorschrift verwendet haben [6].



Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.  
The approved original version of this thesis is available in print at TU Wien Bibliothek.

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Theory</b>	<b>3</b>
2.1	Set Theory . . . . .	4
2.2	Probability Theory . . . . .	7
2.3	Information Theory . . . . .	11
2.4	Quantum Formalism . . . . .	14
2.4.1	Basics of Quantum Mechanics . . . . .	15
2.4.2	Representation of Spin in Quantum Mechanics . . . . .	20
2.5	Quantum Measurement . . . . .	22
2.5.1	Projective Measurements . . . . .	22
2.5.2	POVM Measurements . . . . .	23
2.6	Uncertainty Relations . . . . .	24
2.6.1	Information Theoretic Noise and Disturbance in Quantum Measurements . . . . .	27
<b>3</b>	<b>Experimental Setup</b>	<b>31</b>
3.1	The Neutron . . . . .	31
3.1.1	Neutron Sources for Research . . . . .	33
3.2	Elements of Neutron Polarimetry . . . . .	36
3.2.1	Neutron Supermirrors . . . . .	36
3.2.2	Guidefield . . . . .	40
3.2.3	Spinrotator Coils . . . . .	42
3.2.4	Neutron Detectors . . . . .	46
3.3	Setup for Noise and Disturbance Measurements . . . . .	49
3.3.1	POVM and Correction Procedure . . . . .	49
3.3.2	Setup for Noise Measurement . . . . .	51
3.3.3	Setup for Disturbance Measurement . . . . .	53
<b>4</b>	<b>Results</b>	<b>57</b>
4.1	Noise Measurement . . . . .	57
4.1.1	Raw Data and Data Treatment . . . . .	57
4.1.2	Computation of Probabilities . . . . .	59
4.1.3	Final Noise Plot . . . . .	64
4.2	Disturbance Measurement . . . . .	64
4.2.1	Raw Data and Data Treatment . . . . .	64
4.2.2	Computation of Probabilities . . . . .	66
4.2.3	Final Disturbance Plot . . . . .	70

4.3	Final Noise Disturbance Plot . . . . .	71
4.4	Discussion of the Results . . . . .	72
<b>5</b>	<b>Conclusion and Outlook</b>	<b>77</b>
	<b>References</b>	<b>79</b>
	<b>List of Figures</b>	<b>81</b>
	<b>Acknowledgements</b>	<b>87</b>

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.  
The approved original version of this thesis is available in print at TU Wien Bibliothek.





## Introduction

---

The beginning is the most important part  
of the work.

---

*Plato*

The uncertainty principle introduced by *Werner Heisenberg* in 1927 [1] was the first scientific work addressing the inability of measuring two non-commuting observables of a quantum system sequentially with arbitrarily high precision. The principle was shown heuristically by the means of the so called  $\gamma$ -ray microscope Gedankenexperiment. It is a direct consequence of the *Matrizenmechanik* which was introduced by *Born, Jordan* and *Heisenberg* one year before in 1926 [7]. Matrices do in general not commute. The order of several matrices is relevant in terms of multiplication. Only for the case where the matrices have the same system of eigenvectors a change of order is irrelevant. We say, that those matrices *commute*.

In his Gedankenexperiment Heisenberg argues that in order to measure the position of an electron it must be illuminated by light. A photon interacts with the electron and is reflected back to the observer. The more exact we want to know the position of the electron, the shorter the wavelength of the photon has to be. The energy of the incident photon increases for smaller wavelengths. By interaction with the photon a momentum is passed on to the electron. Mathematically this is a consequence of the non-commuting observables  $\hat{P}$  and  $\hat{Q}$  which are associated with the momentum and position of the electron, respectively. The accuracy of the measurement of the electron position  $q_1$  and the *discontinuous change* of the momentum  $p_1$  caused by the Compton effect can be described by the trade-off relation  $q_1 p_1 \sim h$ . Heisenberg did not specify the indeterminacy but only suggested the correspondance of  $q_1$  and  $p_1$  to some mean error.

The first mathematically rigorous formulation of the uncertainty principle was introduced by *Earl Hesse Kennard* only some months later, leading to the most common formulation of uncertainty known nowadays, being  $\sigma(Q)\sigma(P) \geq \frac{\hbar}{2}$  [2]. To describe the tradeoff relation between the two observables, he used the standard deviation  $\sigma$ . The inequality was soon accepted as an adequate mathematical formulation of the principle. In fact there is a fundamental difference between those two ideas. For the Heisenberg principle the *measurement*, including a force acting on the system under observation is essential. For the Kennard formulation only a wavefunction and corresponding operators are relevant. The formulation can be derived purely mathematically, without the need of a reference to any process of measurement. Nowadays uncertainty relations that describe a measurement, as originally introduced by Heisenberg, are referred to as *measurement uncertainty relations*. The Kennard relation describes

the intrinsic property of quantum systems, that an exact definition of the state of the momentum and the position cannot be acquired with exact precision. Those types of uncertainty principles are called *preparation uncertainty principles* in modern literature.

In the following years the standard deviation was the means of choice for the quantification of the uncertainty principle. Apart from other works on this topic an important milestone was the generalization of the Kennard principle for any non-commuting observables  $\hat{A}$  and  $\hat{B}$  by *Robertson* [8]. The relation was later on tightened by *Schrödinger* [9].

The usage of the *Shannon entropy* as description for uncertainty was introduced by *David Deutsch* in 1983 [3]. He argued, that by using entropic definitions one can achieve a *tighter* bound for the tradeoff relation, meaning that the constant limiting the uncertainty becomes smaller. It was also considered a more appropriate means for the description of uncertainty. Stochastic measures always describe the average behaviour of a system when several independent events are performed consecutively. The assignment of concrete values for observables to single particles is therefore problematic. A detailed review about entropic uncertainty relations and their applications is given in [10]. A review of measurements concerning entropic uncertainty relations using neutron optics is given in [11].

In 2014 *Buscemi et al.* [4] introduced an entropic noise-disturbance measurement uncertainty relation. An investigation of the behaviour of the uncertainty using projective measurements has been done by *Sulyok et al.* [6] in 2015. They were able to confirm the theoretical predictions that conclude from the relation. In 2016 *Abbott and Branciard* proposed a 3-outcome POVM [5] as a more generalized measurement. They argue, that this leads to a tight uncertainty relation for qubit measurements, outperforming the results that can be achieved by projective measurements.

In this thesis, we deal with the experimental investigation of generalized measurements, using the definitions of noise and disturbance from *Buscemi et al.* and the proposed 3-outcome POVM of *Abbott and Branciard*. We use neutron polarimetry to show the principle for qubit systems. The spin of the neutron is the quantum system under investigation. The results are compared with the projective measurements done by *Sulyok et al.* We were able to show that the 3-output POVM outperforms the projective measurements with high statistical significance.

The thesis is structured as follows: At first, the theoretical background relevant for the understanding of the experiment is established. Second, the neutron, with focus on its spin as a quantum system, is introduced and elements of neutron polarimetry are discussed. Finally, the setup used for our measurements is discussed and the measurement results are presented and compared to projective measurements.

It doesn't matter how beautiful your theory is, it doesn't matter how smart you are. If it doesn't agree with experiment, it's wrong.

---

*Richard P. Feynman*

To understand the environment in which we are experimenting, some key elements of mathematical and physical theory need to be introduced and briefly discussed.

As the definitions for noise and disturbance used are based on information theory, an introduction to elements of that theory that are relevant for us will be given. Information theory is linked to probability theory, therefore we need to establish a basic understanding of probabilistic key elements.

Set theory is an axiomatic theory that can be seen as the backbone of not only any probabilistic formulation but also many other formalisms, as quantum formalism. Therefore a need arises to discuss its elements as a foundation for more complex theoretical concepts.

We are dealing with actual measurements of states of a quantum system, therefore we need to introduce a basic quantum formalism that helps us understand what *measuring an observable of a quantum system* actually means. Additionally, different types of measurements of quantum systems will be discussed.

We will introduce uncertainty principles and have a detailed look at an information theoretic definition of noise and disturbance in quantum measurements, as introduced by *Buscemi et al.* in 2014 [4]. This formulation of an entropic measurement uncertainty relation is the one used to quantify noise and disturbance in our measurements and is therefore essential.

## 2.1 Set Theory

The following sections will give an introduction to set theory, providing a base for Bayesian probability, which both are elemental to most quantum mechanical formalisms.

A set  $S$  is a collection of distinct objects  $x_i$ , which are called the elements of a set. If an object  $x$  is an element of a set  $S$ , then we can write  $x \in S$ . When an object  $y$  is not an element of  $S$ , we write  $y \notin S$ . Sets can either be defined extensionally or intensionally.

When a set  $S$  contains a finite number of elements  $x_1, \dots, x_n \mid n \in \mathbb{N}$  we can denote it as a list of all elements of the set in curly brackets. This is a so-called *extensional definition*, e.g.:

$$S = \{x_1, \dots, x_n\} .$$

Alternatively sets can be defined by describing the properties of its elements. This notation is called *intensional definition*, e.g.:

$$S = \{x \mid x \in \mathbb{N} \wedge x \leq 10\} .$$

The two examples for sets above both describe *finite sets* and are most relevant for understanding the theoretic background of this work.

Other examples for sets are the *empty set*

$$S = \{\} := \emptyset ,$$

*countable infinite sets*, e.g.

$$S = \{x \mid x \in \mathbb{N}\} ,$$

and *uncountable sets*, e.g.

$$S = \{x \mid x \in \mathbb{R} \wedge x \geq 10\} .$$

When a set  $S$  is finite, the number of its elements is given by its *cardinality*<sup>†</sup>, which is written as  $|S|$ .

A set typical for the use of a qubit system is the set containing the two possible outcomes of the measurement of the spin of a neutron relative to a quantization axis  $q$ :

$$S = \{|\uparrow\rangle_q, |\downarrow\rangle_q\} \quad (2.1)$$

To define operations and relations between sets we firstly introduce the so called *universal set*  $\Omega$  which can be defined as a set that contains all objects that are relevant in a particular context. A simple example would be the set

$$\Omega_{die} = \{\square, \square, \square, \square, \square, \square\} , \quad (2.2)$$

which contains all possible outcomes – and therefore all relevant objects – in the context of a die throw.

The order of the elements of the sets has not been relevant, even though they sometimes seem to incorporate a natural order. We do not want to introduce ordered sets in this thesis due to irrelevance, but we need to shortly introduce *ordered pairs*. An ordered pair is denoted in round brackets and both elements are members of the same set  $S$ :

$$(a, b) \in S \times S; a, b \in S .$$

<sup>†</sup>In our discussion here, the original definition of the cardinality according to *Cantor*, which is only defined for finite sets, is sufficient. Actually the definition of cardinality has been generalized for infinite sets.

Using Eq. (2.2), which defines a universal set for a die throw, we can illustrate a simple example for an ordered set: We throw two dice simultaneously and want to know the single results. If we would just write down the number of the pips for the dice, we would lose the information which die had which number of pips on it. Therefore we use an ordered set  $\Omega_{die} \times \Omega_{die} := (a, b)$  and refer to  $a$  as the result for die 1 and  $b$  for die 2. With the incorporated order of this concept we don't lose the information about which die had which result, as

$$(a, b) \neq (b, a) ,$$

but

$$\{a, b\} = \{b, a\} .$$

### Set Relations

The *complement* of a set  $S$  is always given in respect to another set, usually, if not defined differently, the universal set  $\Omega$ . It is denoted with the superscript  $c$  and contains all objects that are not in  $S$  but  $\Omega$ . We write

$$S^c = \{x \in \Omega \wedge x \notin S\} .$$

When every element of a set  $S$  is also an element of a set  $T$  we call  $S$  a *subset* of  $T$  and denote it as

$$S \subset T .$$

We can also say that  $T$  is a *superset* of  $S$ , denoted as

$$T \supset S .$$

In the case  $S \subset T \wedge T \subset S$  the sets are called *equal* and we write

$$S = T \Leftrightarrow S \subset T \wedge T \subset S .$$

The *union* of two sets  $S, T$  is defined as a set of objects that belong to  $S, T$  or both, we define

$$S \cup T = \{x | x \in S \vee x \in T\} ,$$

for the union of several, up to infinite sets, we may write

$$\bigcup_{i=1}^N S_i = S_1 \cup S_2 \cup \dots = \{x | x \in S_n \text{ for some } n\} .$$

The *intersection* of two sets  $S, T$  is defined as a set of objects that belong to  $S$  and  $T$ , denoted as

$$S \cap T = \{x | x \in S \wedge x \in T\} ,$$

for the intersection of several, up to infinite sets, we may write

$$\bigcap_{i=1}^N S_i = S_1 \cap S_2 \cap \dots = \{x | x \in S_n \text{ for all } n\} .$$

When the intersection of any number of sets is the empty set (meaning they don't have common elements), those sets are called *disjoint*. Intersections and unions are *symmetric*, meaning  $S \cup T = T \cup S$  respectively  $S \cap T = T \cap S$ .

A very interesting property of the relations of sets are the *De Morgan laws*, which constitute a connection between intersections and unions by the means of complementation. They are denoted as

$$\left( \bigcap_{i=1}^N S_i \right)^C \equiv \bigcup_{i=1}^N S_i^c, \quad (2.3)$$

$$\left( \bigcup_{i=1}^N S_i \right)^C \equiv \bigcap_{i=1}^N S_i^c. \quad (2.4)$$

### Power Sets and $\sigma$ -algebras

A *power set*  $\mathcal{P}(S)$  of a set  $S$  is a set containing all possible subsets of  $S$ , including the empty set  $\emptyset$  and  $S$  itself. Exemplarily, for  $S = \{a, b, c\}$  we would get

$$\mathcal{P}(S) = \{\emptyset, \{a\}, \{b\}, \{c\}, \{a, b\}, \{a, c\}, \{b, c\}, \{a, b, c\}\}.$$

For a set  $S$  with  $|S| = n$  we get  $|\mathcal{P}(S)| = 2^n$ .

A class of sets called  $\sigma$ -*algebra* which is fundamental for a mathematical definition of a measurement is defined in the following manner:

#### Definition 2.1. $\sigma$ -algebra:

Given a set  $S$  and its power set  $\mathcal{P}(S)$ , we call  $\Sigma \subseteq \mathcal{P}(S)$  a  $\sigma$ -algebra when it has the following properties:

1.  $S \subset \Sigma$ :  $S$  is in  $\Sigma$ ; for the following properties  $S$  will be related to as the universal set  $S \equiv \Omega$ .
2.  $\Sigma$  is *closed under complementation*: When  $a \subset S$ , so is its complement  $a^C \subset S$ .
3.  $\Sigma$  is *closed under countable unions*: When  $a, b, c, \dots \subset S$ , so is  $a \cup b \cup c \cup \dots \subset S$ .

□

A *measure on a set* is a systematic way to assign a property to each relevant subset on that set. Typically, in measure theory, the property assigned is a number, often referred to as its *size*. A measure can be seen as a generalization of the assignment for any property that can be represented by a real number or  $+\infty$ . A typical example is the so called *Lebesgue measure*. It assigns the size of suitable subsets of an  $n$ -dimensional Euclidean space  $\mathbb{R}^n$ . For  $n = 1, 2, 3$  it corresponds with the intuitive definitions of size, being *length*, *area* and *volume*.

#### Definition 2.2. Measure on a set:

Given a set  $S$  and a  $\sigma$ -algebra  $\Sigma$  over  $S$ , we call a function  $\mu(\Sigma) : x \in \Sigma \rightarrow \mathbb{R} \cup \{-\infty, +\infty\}$  a measure, when the following properties are fulfilled:

1. *Non-negativity*:  $\mu(x) \leq 0 \quad \forall x \in \Sigma$
2. *Null empty set*:  $\mu(\emptyset) = 0$
3. *Countable additivity*:  $\mu\left(\bigcup_{i=1}^{\infty} E_i\right) = \sum_{i=1}^{\infty} \mu(E_i)$  for all countable collections  $\{E_i\}_{i=1}^{\infty}$  of pairwise disjoint sets in  $\Sigma$ .

A pair  $(S, \Sigma)$  is called a *measurable space*, the members  $x \in \Sigma$  are called *measurable sets*. A triple  $(S, \Sigma, \mu)$  is called *measure space*. □

†As not formally introduced,  $\subseteq$  means *being a subset of or equal to*.

## 2.2 Probability Theory

The probabilistic nature of quantum mechanics was not always undisputed among physicists, especially in its beginnings in the mid 1920s. The doubt about the completeness of quantum mechanical descriptions was specified in the so called *Einstein-Podolsky-Rosen-paradox (EPR-paradox)* in 1935 [12]. Einstein suspected so called *hidden variables* behind the probabilistic predictions in quantum mechanics. That approach was put into mathematical clothing by *Bell* in 1964 [13], nowadays known as the so called *Bell inequalities*. Any violation of these equations would contradict the idea of the EPR-paradox, that a physical theory cannot be complete when relying on probabilistic wave functions.

The first experiment that shows a violation of the Bell inequalities, measuring the linear polarization correlation of the photons emitted in an atomic cascade of Calcium, was done by *Freedman and Clauser* [14] in 1972. Since then many additional experiments provided strong evidence for the irreducibility of probabilistic descriptions in quantum mechanics, e.g., the follow-up experiments by *Aspect et al.* [15, 16].

Modern experiments (2015 and later) are considered *loophole free*, meaning that the experiment is set up in a way that the validity of the experimental findings may not be questioned [17, 18, 19]. Therefore probability as optimal description of quantum systems can be considered as almost certain.

A probabilistic model is a description of a situation that can have mutual outcomes. The information about possible outcomes (so called *events*) and their likelihood is described by that model.

Any probabilistic model consists of two basic elements:

- The *sample space*  $\Omega$  which can be interpreted as a set containing all possible events.
- The *probability law*, that assigns a certain *probability*  $p(x)$  to each possible event  $x$ .

The probability law is based on the following axioms, the *Kolmogorov* or *Probability axioms*:

**Definition 2.3.** *Probability axioms:*

1.  $0 \leq p(x) \leq 1$ ,  $p(x) \in \mathbb{R}$ : The probability of any event  $x \in \Omega$  has to be a non-negative real number.
2.  $\sum_i p(x_i) := 1$ : The probability that any event happens has to be 1 (also denoted as  $p(\Omega) = 1$ ).
3.  $p\left(\bigcup_{i=1}^N x_i\right) = \sum_{i=1}^N p(x_i)$ : Any countable sequence of disjoint sets (defining mutually exclusive events)  $x_1, x_2, \dots$  have additive probability.

□

In the following, examples for probabilistic models will be given:

**Example 2.1.** *Coin toss:*

A coin with two different faces, lets call them heads and tails, is tossed. The sample space containing all possible outcomes can therefore be defined as

$$\Omega_{\text{coin}} = \{\text{heads}, \text{tails}\} .$$

The corresponding probability law assumes that both outcomes are equally probable, meaning

$$p(\text{heads}) = p(\text{tails}) = \frac{1}{2} .$$

This results in a total probability of  $p(\Omega_{\text{coin}}) = 1$ .

◇

This example can be considered as the simplest form of a *Laplace experiment*, in which the probability of all possible results is equal and therefore  $p(x) = 1/n$  for  $|\Omega| = n$ . We call that type of distribution a *uniform distribution*. Therefore the probability law can be written as the *Laplace function*

$$p(x_i) = \frac{\# \text{ of results that produce event } x_i}{\# \text{ of possible results}} = \frac{|x_i|}{|\Omega|}. \quad (2.5)$$

Note the distinction between *result* and *event*, which is relevant for more complex situations and will be explained exemplarily in the next example:

**Example 2.2.** *Two dice thrown simultaneously:*

Two typical dice  $a, b \in \Omega_{die} = \{\square, \blacksquare, \blacklozenge, \blacktriangle, \blacktriangleright, \blacktriangleleft\}$ , are thrown simultaneously, defining a sample space with 36 possible *results* (36 possible ordered pairs of pips):

$$\Omega_{diceSim} = \{(a, b) | a, b \in \Omega_{die}\}$$

We now want to calculate the possibility for the *event*  $x = a + b = 9$  using Eq. (2.5):

$$p(x) = \frac{|x|}{|\Omega|} = \frac{4}{36} = \frac{1}{9}$$

The possible results, denoted as ordered pairs, whose sums are equal to 9 are  $(\blacksquare, \blacklozenge), (\blacklozenge, \blacksquare), (\blacktriangle, \blacktriangleright), (\blacktriangleright, \blacktriangle), (\blacktriangleleft, \blacktriangleright), (\blacktriangleright, \blacktriangleleft)$ . ◇

For the simple case of a single die throw the distinction of result and event is not necessary, as they are directly related.

### Probability Mass Functions

In the examples we got to know so far the probabilities are describable with the help of the Laplace function (Eq. (2.5)) and therefore a probability function that is always static and member of the rational number space:  $p(A_i) \in \mathbb{Q}$ . More generally we talk about a so called *probability mass function* when we want to describe the probability of a *discrete random variable*, meaning, that the outcome of our experiment always has a discrete result (the sample space  $\Omega$  is a countable set). Now we want to properly define a probability mass function:

**Definition 2.4.** *Probability mass function*

A function  $f : x \in \Omega \rightarrow [0, 1]$  assigns a real number in the interval  $[0, 1]$  to a member  $x$  of the sample space  $\Omega$ . When the function is normalized in a way that all assigned values add up to 1 we call it a probability mass function:

$$\sum_{x \in \Omega} f(x) = 1 \quad \Rightarrow \quad p(x) := f(x) \quad \forall x \in \Omega. \quad (2.6)$$

□

We will now take a look at an example for the usage of probability mass functions: the measurement of the spin of a qubit.



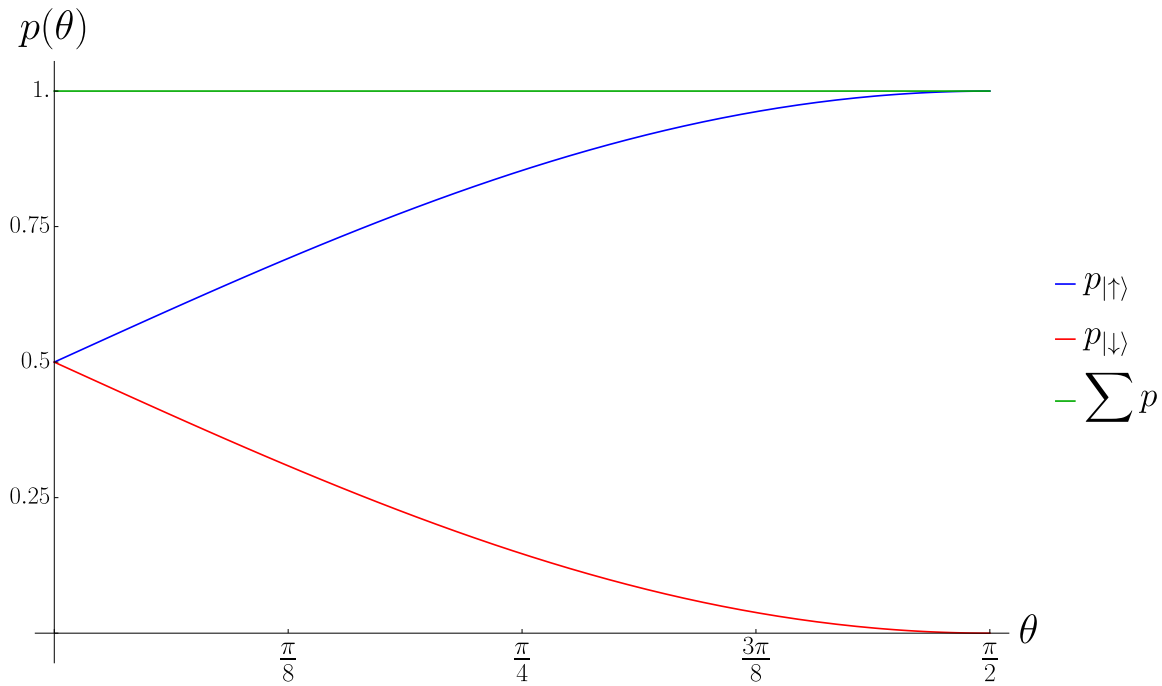


Figure 2.1: Plot of the probability mass function introduced in Example 2.3

**Example 2.3.** *Spin measurement of a neutron*

The spin of a neutron can be interpreted as a Qubit with two possible outcomes. As introduced in Eq. (2.1) we can denote the sample space as

$$\Omega_{spin} = \{|\uparrow\rangle_z, |\downarrow\rangle_z\}$$

when we choose  $z$  as a quantization axis. For simplicity it is common not to write down the quantization axis every single time, meaning  $|\uparrow\rangle \equiv |\uparrow\rangle_z$ ,  $|\downarrow\rangle \equiv |\downarrow\rangle_z$ . Let us now assume that due to a preparation procedure of the spin we can define the following probability mass function for measuring  $|\uparrow\rangle$ :

$$p_{|\uparrow\rangle}(\theta) = \frac{1}{2} \cdot (1 + \sin \theta), \quad \theta \in \left[0, \frac{\pi}{2}\right].$$

As we are using a probability mass function we can directly calculate the probability for the only other outcome  $|\downarrow\rangle$  with

$$\sum_{x \in \Omega} p(x) = p_{|\uparrow\rangle} + p_{|\downarrow\rangle} = 1 \quad \Rightarrow \quad p_{|\downarrow\rangle} = 1 - p_{|\uparrow\rangle} = \frac{1}{2} \cdot (1 - \sin \theta).$$

The probability mass functions for all elements of  $\Omega_{spin}$  are shown in Fig. 2.1.

◇

**Probability Density Functions**

When *continuous random variables* are a possible outcome of the experiment (e.g., the measurement of an ohmic resistance in an electric circuit) we make use of a so called *probability density function* to describe the probability of events in intervals. This means, that the sample space is an uncountable set, e.g.,  $\Omega = \mathbb{R}$ .

We now want to properly define a probability density function:

**Definition 2.5.** *Probability density function*

A function  $f : x \in \Omega \rightarrow [0, 1]$  assigns a real number in the interval  $[0, 1]$  to a continuous random variable  $x$  of the sample space  $\Omega$ . When the function is normalized in a way that all assigned values add up to 1 we call it a probability density function:

$$\int_{\Omega} f(x) dx = 1 \quad \Rightarrow \quad p(x) := f(x) \quad \forall x \in \Omega . \quad (2.7)$$

Even though it is mathematically possible to assign a probability for a *certain* value  $x$ , it only makes sense to define the probability for  $x$  to be in an interval  $[a, b]$ , even though that interval might be infinitesimal. We say that the probability function vanishes for exact values  $x = c$ , we define

$$\int_{\Omega} f(x)\delta(x - c) dx := 0 . \quad (2.8)$$

□

We will now look at an example for the usage of a probability density function, the measurement of the capacity of a capacitor:

**Example 2.4.** *Deviation from the Standard Value of a Capacitor*

We buy a capacitor with the standard value of  $(10 \pm 1) \mu\text{F}$ , provided by the manufacturer. We assume, that the actual value of the capacitor can be described with the help of the *normal distribution*. The normal distribution for an *expectation value*  $\mu$  and a *standard deviation*  $\sigma$  is given by

$$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} . \quad (2.9)$$

We now want to know how likely it is, that the capacitor we bought has a value of  $13 \mu\text{F}$  or more. We set  $\mu = 10$  and  $\sigma = 1$  and receive for the probability

$$p(x \geq 13) = \int_{13}^{\infty} p(x) dx = 0.682689 \approx 68.3\% . \quad (2.10)$$

The probability that the actual value is within  $10 \pm 1$  is given by

$$p(\mu - \sigma < x < \mu + \sigma) = \int_{\mu - \sigma}^{\mu + \sigma} p(x) dx = 0.0013499 \approx 0.13\% . \quad (2.11)$$

◇

### Properties of Probability Laws

Various properties of probability laws can be derived from the probability axioms which were defined in Definition 2.3. They are necessary for further understanding and investigation in the case of more complex probabilistic descriptions. For the following definitions we define a probability law  $p$  and events  $x_1, x_2, x_3$ :

1.  $x_1 \subset x_2 \Rightarrow p(x_1) \leq p(x_2)$
2.  $p(x_1 \cup x_2) = p(x_1) + p(x_2) - p(x_1 \cap x_2)$
3.  $p(x_1 \cup x_2) \leq p(x_1) + p(x_2)$
4.  $p(x_1 \cup x_2 \cup x_3) = p(x_1) + p(x_1^C \cap x_2) + p(x_1^C \cap x_2^C \cap x_3)$

### Conditional and Joint Probabilities

When we discussed Example 2.2 we unknowingly already got in touch with so called joint probabilities. Reformulated we could ask for the probability of *die 1* being  $\boxtimes$  and *die 2* being  $\boxtimes$  instead of looking for ordered pairs as we did before. Thereby we asked for the *joint probability* of two single events  $x_1, x_2$  that occur *at the same time*. The joint probability is denoted as  $p(x_1 \cap x_2) \equiv p(x_1, x_2)$ . We will use the second notation from now on.

*Conditional probability* is a measure for the probability of an event  $x_1$ , assuming that an event  $x_2$  has already occurred. The conditional probability of  $x_1$  *under the condition*  $x_2$  is denoted as  $p(x_1|x_2)$ . When  $p(x_1|x_2) = p(x_1)$  the events  $x_1, x_2$  are called *statistically independent*.

According to *Kolmogorov* the relation between the conditional and joint probability can be defined as

$$p(x_1|x_2) = \frac{p(x_1, x_2)}{p(x_2)} \quad \text{for } p(x_2) > 0 . \quad (2.12)$$

Equation (2.12) can also be reformulated for  $p(x_2|x_1)$  by simply exchanging the arguments  $x_1 \leftrightarrow x_2$  to  $p(x_2|x_1) = \frac{p(x_2, x_1)}{p(x_1)}$ . Using the symmetry of the intersection  $x_1 \cap x_2 = x_2 \cap x_1 = (x_1, x_2) = (x_2, x_1)$ <sup>†</sup> we can deduce the so called *Bayes' theorem*:

$$p(x_1|x_2) = \frac{p(x_2|x_1)p(x_1)}{p(x_2)} \quad \text{for } p(x_2) > 0 . \quad (2.13)$$

## 2.3 Information Theory

Information theory is based on probability theory and statistical means and was introduced by *Claude Shannon* in 1948 in his paper *A Mathematical Theory of Communication* [20]. It assigns a value of the information content  $I(x)$  obtained by the occurrence of an event  $x$  with the probability  $p(x)$ . The information value of an event that is very likely to happen is considered low, as a highly frequent event and its incorporated information may be observed more likely. Opposing, the information value associated with a very improbable event is considered high.

Putting those thoughts into mathematical rigor we can define basic axioms for the information content:

1.  $I(x) \geq 0$ : The information should be non-negative.
2.  $p(x) \rightarrow 1 \Rightarrow I(x) \rightarrow 0$ : Approaching a certain event the information gained vanishes.
3.  $p(x) \rightarrow 0 \Rightarrow I(x)$  increasing: The information should increase when the probability of the associated event approaches zero.
4.  $p(x_1) < p(x_2) \Rightarrow I(x_2) < I(x_1)$  for  $x_1 \neq x_2$ : The more unlikely an event, the more informative it should be.
5.  $p(x_1, x_2) = p(x_1)p(x_2) \Rightarrow I(x_1, x_2) = I(x_1) + I(x_2)$ : The total informational value of two statistically independent events  $x_1, x_2$  should be additive.

<sup>†</sup>Obviously we are not talking about ordered pairs here, even though the notation is the same.

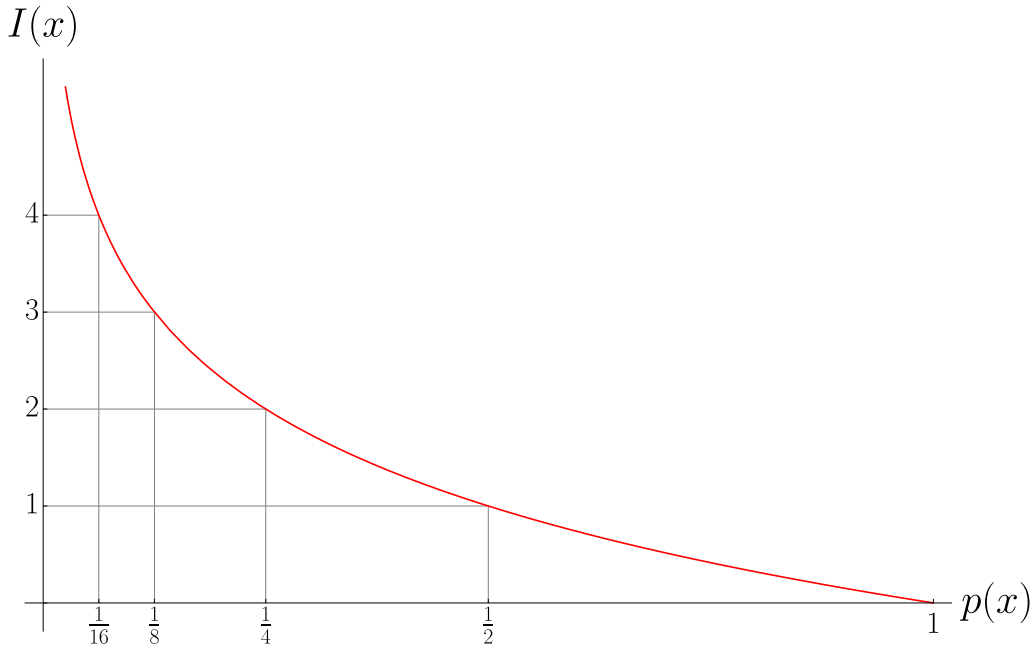


Figure 2.2: The information content  $I(x)$  of an event  $x$  given by its probability  $p(x)$  as defined in Eq. (2.14). One can see the  $2^{-n}$ -dependency of  $p(x)$  from  $I(x)$  that can be derived from its definition.

### Information Content

From the above defined axioms we now derive an expression for  $I(x)$ , which is called *information content*, *self-information* or *surprisal* of a random variable  $x$ .

We are looking for a continuous function of the probability to describe the information content:

$$I(x) = f(p(x)) .$$

With the additivity of the information content of two statistically independent events we have to demand

$$f(p(x_1, x_2)) = \underline{f(p(x_1)p(x_2))} = I(x_1, x_2) = I(x_1) + I(x_2) = \underline{f(p(x_1))} + \underline{f(p(x_2))} ,$$

which is granted by usage of the logarithmic function, as  $\log_x a \cdot b = \log_x a + \log_x b$ .

The base  $b$  of the logarithm can be chosen arbitrarily, we will further on use the most common base of  $b = 2$ , defining the unit of self information as *shannon*[1 Sh], more commonly known as *bit*[1 bit]. With the demand of the information content being a non-negative number we have to negate the logarithm, as  $p(x) \leq 1 \Rightarrow \log_b p(x) \leq 0$  but  $-\log_b p(x) = \log_b \frac{1}{p(x)} \geq 0$ .

Finally, we can define the *information content* of an event  $x$  as

$$I(x) = \log_2 \frac{1}{p(x)} = -\log_2 p(x) . \quad (2.14)$$

A plot of Eq. (2.14) is given in Fig. 2.2.

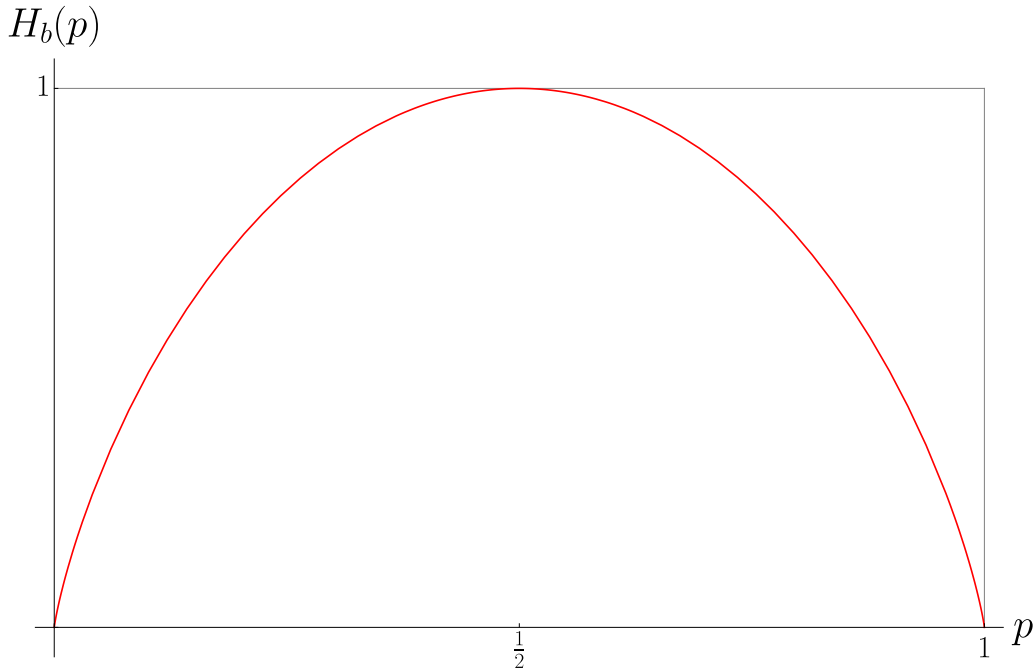


Figure 2.3: The binary entropy function  $H_b(p)$ , defined in Eq. (2.16). The maximal entropy is given at  $p = 1/2 = 1/n$ . It is symmetric around  $p = 1/2$ .

### Shannon Entropy

If we now want to know the average amount of information obtained by a stochastic process  $p(x_i)$  defined on a sample space  $\Omega = \{x_1, x_2, \dots, x_n\}$  we have to find the average of the information content weighted by the probability of its occurrence. We receive the *information entropy* or *Shannon entropy*

$$H(\Omega) = \sum_i p(x_i) I(x_i) = - \sum_i p(x_i) \log_2 p(x_i) . \quad (2.15)$$

In other words we can say that the Shannon entropy gives us a measure for the expected information content acquired by a measurement on a sample space  $\Omega$ . It is maximized when all events are equally probable, representing a Laplace experiment and therefore a uniform distribution. In this case it is also strictly monotonically increasing along with an increasing cardinality of the sample space  $|\Omega|$ .

The second property can be shown very easily, as  $H = \log_2 n$  for an uniform distribution with  $|\Omega| = n$  and  $\log_2 n < \log_2 n + 1$ . We will prove the first one but heuristically, as a mathematical proof would be a little more work: The probability mass function with maximal Shannon entropy seems to be the uniform distribution, as in that case we have the least amount of knowledge about the system we are observing and therefore gain the most amount of information for every measurement.

For the special case of a sample space with only two possible outcomes (mostly referred to as *binary sample space* or *Bernoulli experiment*)  $\Omega = \{a, b\}$  and known probability  $p_a = p$  and therefore  $p_b = 1 - p$  we receive the *binary entropy function*  $H_b(p)$  by calculating the Shannon entropy of the system using Eq. (2.15):

$$H(\Omega) = H_b(p) = -p \log_2 p - (1 - p) \log_2 (1 - p) . \quad (2.16)$$

The binary entropy function is shown in Fig. 2.3.

### Conditional and Joint Entropy

For the *joint entropy*  $H(\Omega_a, \Omega_b)$  we can alter Eq. (2.15) by replacing the probability with the joint probability  $p(a, b)$  of two events  $a \in \Omega_a, b \in \Omega_b$ . Note that we have to use a double sum as we now deal with two sample spaces and need to sum over all possible combinations:

$$H(\Omega_a, \Omega_b) = - \sum_{a \in \Omega_a} \sum_{b \in \Omega_b} p(a, b) \log_2 p(a, b) . \quad (2.17)$$

To define the *conditional entropy*  $H(\Omega_a|\Omega_b)$  we need to average  $H(\Omega_a|b)$  over all possible values of  $b$ :

$$H(\Omega_a|\Omega_b) = \sum_{b \in \Omega_b} p(b) H(\Omega_a|b) = - \sum_{b \in \Omega_b} p(b) \sum_{a \in \Omega_a} p(a|b) \log_2 p(a|b) .$$

We now use Eq. (2.12) to simplify the term to

$$H(\Omega_a|\Omega_b) = - \sum_{b \in \Omega_b} \sum_{a \in \Omega_a} p(a, b) \log_2 p(a|b) = - \sum_{b \in \Omega_b} \sum_{a \in \Omega_a} p(a, b) \log_2 \frac{p(a, b)}{p(b)} ,$$

and finally receive

$$H(\Omega_a|\Omega_b) = \sum_{b \in \Omega_b} \sum_{a \in \Omega_a} p(a, b) \log_2 \frac{p(b)}{p(a, b)} \quad (2.18)$$

for the conditional entropy.

## 2.4 Quantum Formalism

The *bra-ket-notation*, introduced by *Paul Dirac* in 1930, is a standard for the notation of vectors in and linear forms on a Hilbert space in quantum mechanics. A vector  $\psi$  in a Hilbert space  $\mathcal{H}$  over the body of complex numbers  $\mathbb{C}$  is denoted as a *ket*  $|\psi\rangle$ . A *linear form*, also called *dual vector*, on  $|\psi\rangle$  is written as a *bra*  $\langle\psi|$ .

A strict definition of the mathematical concept of the Hilbert space and a derivation of its properties will not be given. For now, the knowledge that the mathematical space in which we operate in quantum mechanics is a Hilbert space over the complex numbers will be sufficient. All properties and concepts that are important for us physicists will be introduced in close accordance with the original work of Dirac, where the bra-ket-notation was introduced [21]. Please note, that there are other formalisms and concepts of denotation in quantum mechanics. The means of the bra-ket-notation are not sufficient for the notation of any problem related to quantum theory. Due to its simplicity and appropriate description of most cases it can still be seen as a notation standard in quantum mechanics and is used throughout most modern publications in the area.

The section on quantum measurement follows chapters 2.2.5 and 2.2.6 of the book *Quantum Computation and Quantum Information* [22]. Here the common types of measurements in a quantum system are introduced and formally defined. They are especially relevant for our measurements on the neutron spin as a qubit system.

### 2.4.1 Basics of Quantum Mechanics

#### Bra and Ket Vectors

For every set of vectors  $\{|\psi_1\rangle, |\psi_2\rangle, \dots, |\psi_n\rangle\}$  it is possible to set up a second set of vectors, so called *dual vectors*  $\{\langle\psi_1|, \langle\psi_2|, \dots, \langle\psi_n|\}$  that is associated to the original ones by the so called *inner product*  $\langle\psi_i|\psi_i\rangle$ . The inner product is a *scalar product*, as the result of multiplying a bra- with a ket-vector is always a scalar of the real numbers, hence

$$\langle\phi|\psi\rangle = c, \quad c \in \mathbb{R}, \quad (2.19)$$

for any vectors  $\langle\phi|, |\psi\rangle$ . In general, the *scalar product* is calculated with the help of an integral function. We receive ultimately

$$\langle\phi|\psi\rangle = \int \phi^\dagger(x_i)\psi(x_i)dx_i \quad (2.20)$$

for any dependency of the coordinates  $x_i$ .

The scalar product of  $\langle\phi|$  and  $|\psi\rangle$  is a linear function of  $|\psi\rangle$  and must therefore fulfill the two basic properties of *linearity*, referred to as *additivity* and *homogeneity*. They are expressed symbolically by

$$\begin{aligned} \langle\phi| \left( |\psi_1\rangle + |\psi_2\rangle \right) &= \langle\phi|\psi_1\rangle + \langle\phi|\psi_2\rangle, \\ \langle\phi| \left( c|\psi\rangle \right) &= c\langle\phi|\psi\rangle, \end{aligned} \quad (2.21)$$

with  $c \in \mathbb{C}$ . The linearity also holds for the first argument of the scalar function, namely the bra-vector.

To keep it general, we used *any vectors* for the definition of the scalar product (Eq. (2.19)). Those vectors have no connection but the existence of a scalar product between a set of a bra- and a ket-vector. Now we want to define a correspondence between a ket-vector and a bra-vector. The operation that ensures the desired correspondence is the *conjugate transpose* or *Hermitian transpose*  $|\psi\rangle^H \equiv |\psi\rangle^\dagger$  of the vector  $|\psi\rangle$ .

For the complex conjugate of a vector we perform the complex conjugation on every single entry, meaning, that the imaginary part changes its sign ( $a + i \cdot b \leftrightarrow a - i \cdot b$ ). When the entry is real, no modification is necessary, as  $a + i \cdot b = a - i \cdot b$  for  $b \rightarrow 0$ . Also we transpose the vector, meaning that a column vector gets written as a row vector and vice versa.

We can now obtain the dual vector corresponding to any  $|\psi\rangle$  with

$$|\psi\rangle^\dagger := \left( \overline{|\psi\rangle} \right)^T = \langle\psi|, \quad (2.22)$$

with the  $\overline{\phantom{x}}$  representing complex conjugation and the *superscripted* T denoting transposition.

The complex conjugate of a bra-vector results in the corresponding ket-vector. Therefore we can state, that the dual vector of a dual vector results in the original vector, symbolically

$$\begin{aligned} |\psi\rangle^\dagger &= \langle\psi|, \\ \langle\psi|^\dagger &= |\psi\rangle. \end{aligned} \quad (2.23)$$

For the physical interpretation of bra- and ket-vectors we define, that *any state of a dynamical system at a particular time can be specified by the direction of a bra- or ket-vector*. This also implies, that the state described by any vector  $|\psi\rangle$  is the same state described by  $c|\psi\rangle$  with  $c \in \mathbb{C}$ ; the relevant information is contained *only by the direction* of the vector.

Explicitly, the sign of the vector is therefore irrelevant, meaning, that  $c|\psi\rangle$  and  $-c|\psi\rangle$  also refer to the same state.

As it is convenient to make all relevant vectors the same size, there is a mathematical process called *normalization*. To *normalize* a vector  $|\psi\rangle$ , we have to divide it through its *length*, also referred to as *norm*  $\|\psi\|$ . The norm of a vector is defined as the square root of its scalar product with its dual vector, symbolically

$$\|\psi\| := \left( \sqrt{\langle\psi|\psi\rangle} \right)^2 = |\langle\psi|\psi\rangle| . \quad (2.24)$$

With

$$\frac{|\psi\rangle}{|\langle\psi|\psi\rangle|} = |\psi\rangle_{norm} , \quad (2.25)$$

we therefore receive a normalized vector  $|\psi\rangle_{norm}$  pointing in the same direction as  $|\psi\rangle$  and  $\|\psi_{norm}\| = 1$ .

The normalization procedure has an important link to probability theory. As introduced in Definition 2.3, the total probability  $p(\Omega)$  of a sample space  $\Omega$  has to be equal to 1. This means, that with every measurement we must receive *any* result. Later on we will make measurements on the states represented by our ket-vectors. By normalization we ensure that the total probability for all possible outcomes of a measurement is equal to 1. Therefore, the integrand of Eq. (2.20) has a direct connection to the *probability density function*, as  $p(x_i) = \phi^\dagger(x_i)\psi(x_i) dx_i$ .

There is another use of the scalar product that is of importance regarding its physical interpretation. When we have two states  $|\psi\rangle$  and  $|\phi\rangle$ , we call the scalar product  $\langle\phi|\psi\rangle$  the *overlap of the states*  $|\phi\rangle$  and  $|\psi\rangle$ . It is interpretable as the probability amplitude for the state  $|\psi\rangle$  to collapse into the state  $|\phi\rangle$ . The overlap is 0, unless  $|\psi\rangle$  is a superposition of  $|\phi\rangle$  and an arbitrary number of other states.

## Linear Operators

So far we have dealt with the handling of mathematical operations of vectors and complex numbers. Now we want to specify a *linear operator*  $\alpha$ , which forms a new vector  $|\phi\rangle$  by multiplication with an initial ket  $|\psi\rangle$ . We say, that we *apply* the operator on our state vector, symbolically

$$|\phi\rangle = \alpha |\psi\rangle . \quad (2.26)$$

It is a rule, that linear operators always have to be written to the left of a state vector. Linear operators also fulfill the properties of linearity, see Eq. (2.21). We can also deduce from Eq. (2.26) that for a  $n$ -dimensional complex vector  $|\psi\rangle \in \mathbb{C}^n$  the operator must be in matrix form  $\alpha \in \mathbb{C}^{n \times n}$ .

In general, when we apply several linear operators on a state, the order of the operators is relevant, therefore we say, that linear operators are in general *not commutative*, meaning

$$\alpha\beta |\psi\rangle \neq \beta\alpha |\psi\rangle , \quad (2.27)$$

for the application of any two linear operators  $\alpha, \beta$  on any state  $|\psi\rangle$ .

For some cases, two linear operators may commute. To measure the deviation from being commutative for two linear operators  $\alpha, \beta$ , we introduce a concept called *commutator*  $[\alpha, \beta]$ . The commutator is defined as

$$[\alpha, \beta] = \alpha\beta - \beta\alpha . \quad (2.28)$$



When the commutator vanishes,  $\alpha$  and  $\beta$  are said to be commutative. The physical interpretation is, that the state resulting of applying  $\alpha$  is not disturbed by the application of  $\beta$  and vice versa, mathematically

$$[\alpha, \beta] = 0 \Leftrightarrow \alpha\beta|\psi\rangle = \beta\alpha|\psi\rangle . \quad (2.29)$$

A linear operator  $\alpha$  may also be applied to a bra  $\langle\phi|$ , here we have the rule that the operator must always be noted on the right. The property of *associativity* holds for a combination of a bra, a ket and a linear operator, mathematically

$$\left(\langle\phi|\alpha\right)|\psi\rangle = \langle\phi|\left(\alpha|\psi\rangle\right) := \langle\phi|\alpha|\psi\rangle . \quad (2.30)$$

When we have a linear operator  $\alpha$  applied on a ket  $|\psi\rangle$  and we want to find the corresponding operator in the dual space, operating on  $\langle\psi|$ , we need to transform  $\alpha$ . The operation that ensures the desired correspondence is the already introduced *conjugate transpose*.

For the conjugate transpose you have to transpose the operator, transforming a row vector into a column vector and vice versa. Also, every single entry is complex conjugated.

We can now obtain the linear operator  $\alpha^\dagger$  corresponding to any  $\alpha$ , we write

$$\alpha^\dagger := \alpha . \quad (2.31)$$

We now can let an operator-vector-pair transform to its corresponding dual space, mathematically formulated as

$$\langle\psi|\alpha^\dagger \xleftrightarrow{\text{dual space}} \alpha|\psi\rangle . \quad (2.32)$$

At first glance, Eqs. (2.30) and (2.32) seem to contradict each other. At first we argue, that there is no difference, whether an operator is applied to a bra or a ket, then we say, that there is a need to transform the operator dependent on whether it is applied to a bra or ket, respectively. To resolve that matter, we point out, that the first equation is corresponding to the mathematical properties of linear operators in Hilbert spaces in general. On the contrary, the latter corresponds to the physical interpretation of the application of a linear operator on a vector.

Therefore it is common procedure, that for  $\langle\phi|\alpha^\dagger|\psi\rangle$  we apply the operator to the left on the bra, for  $\langle\phi|\alpha|\psi\rangle$  we apply it to the right on the ket.

Now we want to conclude with the assumption, that linear operators *correspond* to the dynamical variables of a quantum mechanical system at a certain time. This means, that with the help of a linear operator we can take a measurement of the variables associated with an operator.

Operators which are associated with measuring a certain dynamical variable are normally denoted with a *hat*, e.g. the *Hamilton operator*  $\hat{H}$  which corresponds to the energy of a system or the *position operator*  $\hat{x}$ , corresponding to the position of the system in the associated vector space.

### Eigenvalues and Eigenvectors

We now want to have a look at the so called *eigenvalue equation*, where we make a measurement of the physical observable  $A$  with the help of the linear operator  $\hat{A}$  associated with that observable:

$$\hat{A}|\psi\rangle = a|\psi\rangle . \quad (2.33)$$

When this equation is fulfilled (ignoring the trivial solution  $|\psi\rangle = 0$ ), we call  $a$  an *eigenvalue* of the linear operator  $\hat{A}$  and  $|\psi\rangle$  an *eigenvector* of  $\hat{A}$ . Therefore the terms eigenvalue and eigenvector only make sense when referring to a linear operator or the variable associated with that operator.

As  $a$  is the value of the measurement of a physical observable, it has to be a real number, mathematically  $a \in \mathbb{R}$ . To ensure this property, the associated operator  $\hat{A}$  must be *Hermitian* (also referred to as *self-adjoint*).

An Hermitian operator is defined via the invariance under the operation of the conjugate transpose, symbolically  $\hat{A}^\dagger = \hat{A}$ .

With that information we may define the *measurement of an observable* with the help of Eq. (2.33). We multiply the bra  $\langle\psi|$  on both sides of the equation and get

$$\langle\psi|\hat{A}|\psi\rangle = a\langle\psi|\psi\rangle, \quad (2.34)$$

we can now express  $a$  in terms of the state and the operator, hence

$$a = \frac{\langle\psi|\hat{A}|\psi\rangle}{\langle\psi|\psi\rangle}. \quad (2.35)$$

For normalized eigenstates the equation reduces to

$$a = \langle\psi|\hat{A}|\psi\rangle, \quad (2.36)$$

as  $\langle\psi|\psi\rangle = 1$ .

It is important to note that a linear operator can have several corresponding eigenvectors and eigenvalues. We only consider the case where the eigenvalues for different eigenvectors are *not degenerated*, meaning, that  $a_n \neq a_m \forall n, m$ .

When we apply the operator  $\hat{A}$  on two different eigenvectors  $|\psi_1\rangle, |\psi_2\rangle$  we get

$$\begin{aligned} (1) \quad & \hat{A}|\psi_1\rangle = a_1|\psi_1\rangle, \\ (2) \quad & \hat{A}|\psi_2\rangle = a_2|\psi_2\rangle. \end{aligned}$$

Now we use  $\langle\psi_2|$  on (1) and  $\langle\psi_1|$  on (2), resulting in

$$\begin{aligned} (3) \quad & \langle\psi_2|\hat{A}|\psi_1\rangle = a_1\langle\psi_2|\psi_1\rangle, \\ (4) \quad & \langle\psi_1|\hat{A}|\psi_2\rangle = a_2\langle\psi_1|\psi_2\rangle. \end{aligned}$$

In the next step we conjugate (4) and use  $\hat{A}^\dagger = \hat{A}$  and  $\overline{a_n} = a_n$  and get

$$(5) \quad \langle\psi_2|\hat{A}|\psi_1\rangle = a_2\langle\psi_2|\psi_1\rangle.$$

which is only fulfilled for  $\langle\psi_2|\psi_1\rangle = 0$ , as  $a_2 \neq a_1$ .

When the scalar product of two vectors vanishes, we call them *orthogonal*. Therefore we have shown, that for eigenvectors of different eigenstates, the property of orthogonality holds.

### Orthonormal Bases

Every vector  $|\psi\rangle$  in a Hilbert space can be written as a linear combination of orthogonal basis vectors  $\{|e_n\rangle\}$ . When we write a state as a combination of other states, we call that a *superposition*:

$$|\psi\rangle = \sum_n c_n |e_n\rangle = \sum_n \langle e_n|\psi\rangle |e_n\rangle , \quad (2.37)$$

with  $c_n \in \mathbb{C}$  being the *evolution coefficients*.

Typically, the basis vectors are all normalized, with the help of the *delta function*  $\delta_{mn}$ . An orthogonal base with normalized vectors is called *orthonormal base*. We can therefore write

$$\langle e_m|e_n\rangle = \delta_{mn} . \quad (2.38)$$

The delta function is defined as being 1 for matching indexes and 0 else, mathematically

$$\delta_{mn} = \begin{cases} 1 & \text{for } m = n \\ 0 & \text{for } m \neq n \end{cases} . \quad (2.39)$$

We also want to normalize the state  $|\psi\rangle$ . We form the scalar product  $\langle\psi|\psi\rangle$  with the help of Eq. (2.37) and demand it being equal to 1. We get

$$\langle\psi|\psi\rangle = \sum_n \langle\psi|e_n\rangle \langle e_n|\psi\rangle = \sum_n c_n \bar{c}_n ,$$

and therefore

$$\sum_n |c_n|^2 \stackrel{!}{=} 1 , \quad (2.40)$$

as a normalization condition.

Another possibility of denoting a state which is a composition of base vectors is via the help of a so called *density matrix*  $\rho$ . To properly describe a density matrix, we must first introduce so called *projection operators*, which are a ket followed by a bra representing the same state, symbolically  $|\psi\rangle\langle\psi|$ . Typically we use projectors on the orthonormal base, because we want to see the decomposition of a state into its base vectors and corresponding evolution coefficients. A projector on the  $n$ -th eigenstate of our base is also denoted as  $P_n = |e_n\rangle\langle e_n|$ .

When we project any vector  $|\psi\rangle$  on  $|e_n\rangle$ , we get

$$P_n |\psi\rangle = |e_n\rangle \langle e_n|\psi\rangle = c_n |e_n\rangle . \quad (2.41)$$

Projectors are *idempotent*, meaning, that any power  $m$  of the operator results in the same operator again, symbolically  $P_n^m = P_n$ .

Another very important property of projectors is that we are able to express the unity operator  $\mathbb{1}$  with them by summing up over the complete base, mathematically

$$\sum_n P_n = \mathbb{1} . \quad (2.42)$$

With the help of the projector we can now finally write down our density matrix  $\rho$ , which is nothing else but a matrix representation of a state in terms of an orthonormal base.

The simplest form is denoted as

$$\rho = \sum_n p_n P_n , \quad (2.43)$$

with  $p_n$  being the probability for the system being in a *pure state*  $|e_n\rangle$ . Again, the probabilities need to sum up to 1, therefore we must demand

$$\text{Tr } \rho = 1 , \quad (2.44)$$

with  $\text{Tr } \rho$  being the trace of the density matrix.

We call the state of a system pure, when it cannot be degenerated further, meaning that there is no superposition of states but a *certain probability for that certain state*. When there is a statistical ensemble of several states, representing superposition, we call it a *mixed state*.

If we now want to measure the probability for the state  $|e_n\rangle$ , we have to apply the corresponding projector on the density matrix, mathematically

$$p(e_n) = \text{Tr } P_n \rho . \quad (2.45)$$

If this result is measured the system is now in a pure state and may be described by a new density matrix  $\rho'$ , calculated as

$$\rho' = \frac{P_n \rho P_n}{\text{Tr } P_n \rho} . \quad (2.46)$$

By dividing by  $\text{Tr } P_n \rho$ , the resulting density matrix is renormalized.

## 2.4.2 Representation of Spin in Quantum Mechanics

We now want to have a closer look on how to measure the spin of a quantum system with the help of the mathematical means introduced above. The spin  $\frac{1}{2}$  can be described by usage of a complex vector space of dimension 2, formally  $\mathbb{C}^2$ . This model implies all necessary properties to constitute a Hilbert space. We also define the  $z$ -axis as so called quantization axis.

Any measurement on any spin  $\frac{1}{2}$ -system is performed with the *spin operator*  $\hat{S}$ , which is defined as

$$\hat{S}_i = \frac{\hbar}{2} \hat{\sigma}_i , \quad (2.47)$$

with  $i$  being the measurement axis ( $x, y$  or  $z$ ) and  $\hbar$  the *reduced Planck constant*.

The  $\hat{\sigma}_i$  are called *Pauli matrices* and are defined as follows:

$$\hat{\sigma}_1 = \hat{\sigma}_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} , \quad (2.48)$$

$$\hat{\sigma}_2 = \hat{\sigma}_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} , \quad (2.49)$$

$$\hat{\sigma}_3 = \hat{\sigma}_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} . \quad (2.50)$$

When we look for the eigenvectors of the pauli matrices we get

$$|+\hat{x}\rangle = |+\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad |-\hat{x}\rangle = |-\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix} \quad \text{for } \hat{\sigma}_x, \quad (2.51)$$

$$|+\hat{y}\rangle = |\phi^+\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ i \end{pmatrix} \quad |-\hat{y}\rangle = |\phi^-\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -i \end{pmatrix} \quad \text{for } \hat{\sigma}_y, \quad (2.52)$$

$$|+\hat{z}\rangle = |\uparrow\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad |-\hat{z}\rangle = |\downarrow\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad \text{for } \hat{\sigma}_z, \quad (2.53)$$

with eigenvalues  $+1$  for  $|+\rangle, |\phi^+\rangle, |\uparrow\rangle$  and  $-1$  for  $|-\rangle, |\phi^-\rangle, |\downarrow\rangle$ .

The commutator of any Pauli spin matrices has the property

$$[\hat{\sigma}_i, \hat{\sigma}_j] = 2i\epsilon_{ijk}\hat{\sigma}_k. \quad (2.54)$$

As the Pauli matrices do not commute, we note, that the measurement of the spin in a certain direction disturbs a consecutive measurement. The order of measuring the spin in different directions is therefore relevant.

We used the *Levi-Civita symbol*  $\epsilon_{ijk}$  for the description of the commutator. It is also called *permutation symbol* and is defined as

$$\epsilon_{ijk} = \begin{cases} +1 & \text{if } (i, j, k) \text{ is an even permutation of } (1, 2, 3) \\ -1 & \text{if } (i, j, k) \text{ is an odd permutation of } (1, 2, 3) \\ 0 & \text{if any indices match} \end{cases}. \quad (2.55)$$

## Bloch Sphere

We can geometrically represent any polarization state of our spin system with the help of the so called *Bloch sphere* (see Fig. 2.4). With the help of this representation we can map a unit 2-sphere to our two-dimensional state space  $\mathbb{C}^2$ . Any normalized state  $|\psi\rangle$  can be written as a superposition of the basis vectors  $|\uparrow\rangle$  and  $|\downarrow\rangle$  in the form

$$|\psi\rangle = \cos \frac{\theta}{2} |\uparrow\rangle + e^{i\phi} \sin \frac{\theta}{2} |\downarrow\rangle = \cos \frac{\theta}{2} |\uparrow\rangle + (\cos \phi + i \sin \phi) \sin \frac{\theta}{2} |\downarrow\rangle \quad \text{with } 0 \leq \theta \leq \pi. \quad (2.56)$$

The projections of  $|\psi\rangle$  on the respective axes are given by the vector

$$\vec{n} = \begin{pmatrix} \sin \theta \cos \phi \\ \sin \theta \sin \phi \\ \cos \theta \end{pmatrix}. \quad (2.57)$$

As  $\vec{n}$  is normalized ( $|\vec{n}| = 1$ ), we can measure the spin in any direction with

$$\hat{S} = \frac{\hbar}{2} \vec{n} \begin{pmatrix} \hat{\sigma}_1 \\ \hat{\sigma}_2 \\ \hat{\sigma}_3 \end{pmatrix} = \frac{\hbar}{2} \vec{n} \vec{\sigma}. \quad (2.58)$$

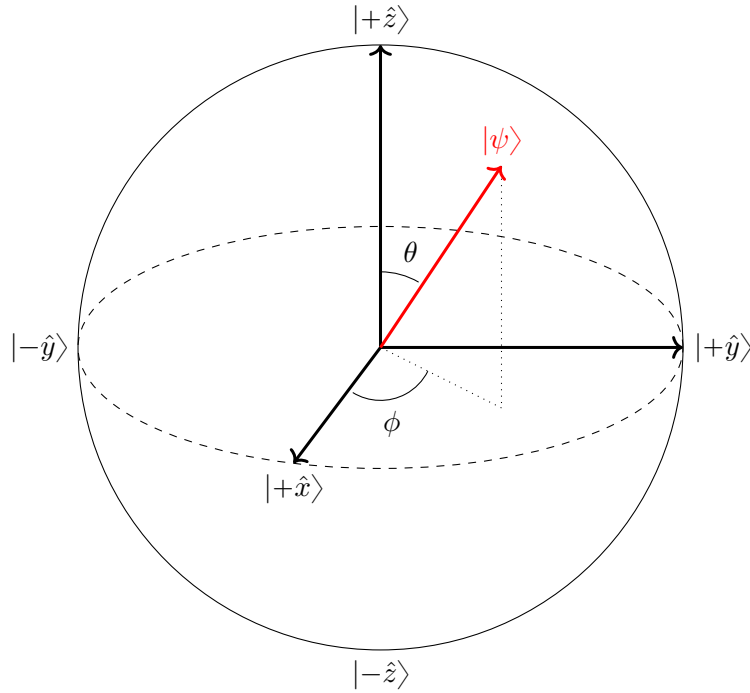


Figure 2.4: Visual representation of the Bloch sphere.

## 2.5 Quantum Measurement

Quantum measurements are described by a collection  $\{M_n\}$  of *measurement operators*. The index  $n$  refers to possible outcomes of the measurement and is therefore dependent on the quantum system we perform measurements on. As established in Section 2.4.1 we can obtain the probability for the measurement outcome  $n$ , measuring on a state  $|\psi\rangle$  with

$$p(n) = \langle \psi | M_n^\dagger M_n | \psi \rangle . \quad (2.59)$$

The post-measurement state is then given by

$$\frac{M_n |\psi\rangle}{\sqrt{\langle \psi | M_n^\dagger M_n | \psi \rangle}} . \quad (2.60)$$

### 2.5.1 Projective Measurements

*Projective measurements* are a special case of the general measurement. A projective measurement is described by an observable  $M$  which can be written down as a spectral decomposition

$$M = \sum_n n P_n , \quad (2.61)$$

with  $P_n$  being the projectors on the eigenvectors of  $M$  with the corresponding eigenvalues  $n$ . The operator  $M$  is hermitian.

The probability for an outcome  $n$  therefore reduces to

$$p(n) = \langle \psi | P_n | \psi \rangle , \quad (2.62)$$

and the post-measurement state is defined as

$$\frac{P_n |\psi\rangle}{\sqrt{p(n)}} . \quad (2.63)$$

**Example 2.5.** *Projective Measurement of the Spin of a Neutron:*

With the help of Eq. (2.62) we can define the observable  $M$  for a measurement in the basis  $\{|\uparrow\rangle, |\downarrow\rangle\}$  of the spin:

$$M = \sum_n n P_n = P_\uparrow - P_\downarrow = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \hat{\sigma}_3 . \quad (2.64)$$

We have now indirectly shown that the Pauli matrix  $\hat{\sigma}_3$  fulfills all demands of a projective measurement. With the same argument it is trivial to show that this holds for all Pauli matrices  $\hat{\sigma}_i$  with respect to their eigenbases.

We can now generalize a projective measurement in *any* direction  $\vec{n}$  by finding the eigensystem of  $\hat{S}$  (see Eqs. (2.57) and (2.58)). We define a new base with the eigenvectors of  $\hat{S}$ :

$$|n+\rangle = \frac{1}{\sqrt{1 + e^{2\text{Im}\{\phi\}} \cot^2(\frac{\theta}{2})}} \begin{pmatrix} e^{-i\phi} \cot(\frac{\theta}{2}) \\ 1 \end{pmatrix} \quad \text{with eigenvalue } +\frac{\hbar}{2} , \quad (2.65)$$

$$|n-\rangle = \frac{1}{\sqrt{1 + e^{2\text{Im}\{\phi\}} \tan^2(\frac{\theta}{2})}} \begin{pmatrix} -e^{-i\phi} \tan(\frac{\theta}{2}) \\ 1 \end{pmatrix} \quad \text{with eigenvalue } -\frac{\hbar}{2} , \quad (2.66)$$

and find

$$M = \hat{S} = \frac{\hbar}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} , \quad (2.67)$$

in its diagonal form with respect to its eigenbase.  $\diamond$

### 2.5.2 POVM Measurements

*POVM* stands for *Positive Operator Valued Measure*. This kind of measurement is usually performed, when the post measurement state of the system is not of specific interest, but the measurement probabilities are focused on. They can be seen as the most general formulation of a set of measurement operators that fulfill the properties of a general measurement, as introduced above.

Again we want to perform a measurement on a quantum system in a state  $|\psi\rangle$  by the help of a set of measurement operators  $\{M_n\}$ . When we now define an operator

$$E_n = M_n^\dagger M_n , \quad (2.68)$$

we receive the probabilities of measuring the outcome  $n$  with

$$p(n) = \langle \psi | E_n | \psi \rangle , \quad (2.69)$$

and the post-measurement state as

$$\frac{M_n}{\sqrt{p(n)}} |\psi\rangle . \quad (2.70)$$

The  $M_n$  are called *Kraus operators*.

We can now define any set of operators  $\{E_n\}$  to be a POVM when the following properties are fulfilled:

1. All  $E_n$  are *positive operators*.
2. The  $E_n$  fulfill the *completeness relation*  $\sum_n E_n = \mathbb{1}$ .

We now want to have a look at an example of a POVM measurement which was taken from [22, p. 92]:

**Example 2.6.** *POVM Measurement of the Spin of a Neutron:*

We set up a quantum system that can be in two different states,

$$|\psi_1\rangle = |\uparrow\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad |\psi_2\rangle = \frac{1}{\sqrt{2}}(|\uparrow\rangle + |\downarrow\rangle) = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

As the two states are not orthonormal, there is no quantum measurement that can distinguish the states with perfect reliability. With the help of a POVM measurement, it is possible to design the measurement apparatus in a way that no misidentification of states is possible.

We consider a 3-output POVM

$$\begin{aligned} E_1 &= \frac{\sqrt{2}}{1 + \sqrt{2}} P_{\downarrow} = \frac{\sqrt{2}}{1 + \sqrt{2}} \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \\ E_2 &= \frac{\sqrt{2}}{2(1 + \sqrt{2})} (P_{\uparrow} + P_{\downarrow} - |\uparrow\rangle\langle\downarrow| - |\downarrow\rangle\langle\uparrow|) = \frac{\sqrt{2}}{2(1 + \sqrt{2})} \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}, \\ E_3 &= \mathbb{1} - E_1 - E_2 = \begin{pmatrix} 1 - \frac{1}{\sqrt{2}} & 1 - \frac{1}{\sqrt{2}} \\ 1 - \frac{1}{\sqrt{2}} & -2 + \frac{3}{\sqrt{2}} \end{pmatrix}. \end{aligned}$$

We now perform measurements on both states:

$$\begin{aligned} \langle\psi_1|E_1|\psi_1\rangle &= 0 & \langle\psi_2|E_1|\psi_2\rangle &= 1 - \frac{1}{\sqrt{2}} \\ \langle\psi_1|E_2|\psi_1\rangle &= 1 - \frac{1}{\sqrt{2}} & \langle\psi_2|E_2|\psi_2\rangle &= 0 \\ \langle\psi_1|E_3|\psi_1\rangle &= \frac{1}{\sqrt{2}} & \langle\psi_2|E_3|\psi_2\rangle &= \frac{1}{\sqrt{2}} \\ \sum &= 1 \quad \checkmark & \sum &= 1 \quad \checkmark \end{aligned}$$

When a measurement is taken with the help of this POVM, we can distinguish which state was prepared *with absolute certainty* for two of the 3 possible outcomes  $E_1, E_2, E_3$ . If we measure  $E_1$ , we can conclude that the prepared state must be  $|\psi_2\rangle$ , as  $\langle\psi_1|E_1|\psi_1\rangle = 0$ .

For measuring  $E_2$ , we know that the prepared state must be  $|\psi_1\rangle$ , as  $\langle\psi_2|E_2|\psi_2\rangle = 0$ . Only for the measurement of  $E_3$  we can not say which state was prepared. By usage of the POVM formalism of a measurement we are therefore able to design a measurement apparatus with which we will never fail to identify the state of the system. On the other hand, here in the case of measuring  $E_3$ , we will not gain information about the state the system is in at all.  $\diamond$

## 2.6 Uncertainty Relations

The uncertainty principle introduced by *Werner Heisenberg* in 1927 was the first theoretical framework that described the inability of *exact* measurements of incompatible observables in quantum mechanics [1].

The original formulation proposed a *reciprocal relation* for the measurement *mean error* and the *disturbance* of the successively measured observable. He used a  *$\gamma$ -ray microscope thought experiment* with



the *Compton effect* as physical background to describe the trade-off of the accuracy of the measurement of position and momentum of an electron.

The symbolical formulation of the principle is

$$p_1 q_1 \sim h , \quad (2.71)$$

with  $p_1$  being the discontinuous change of momentum for a mean error of a position measurement  $q_1$ . The tradeoff is symbolically limited with the *Planck constant*  $h$ .

In the original formulation there is no rigid mathematical proof of the principle provided and the lower limit for the uncertainty is not defined in an exact manner. A modern reference of the same principle for two observables  $Q, P$  is given by

$$\epsilon(Q)\eta(P) \geq \frac{\hbar}{2} , \quad (2.72)$$

with  $\epsilon$  being the error of the first measurement (of  $Q$ ) and  $\eta$  being the disturbance on a second measurement (of  $P$ ). The descriptions of error and disturbance are operator based. They are treated symbolically in the Heisenberg paper and are not yet formally defined.

The Heisenberg uncertainty principle is considered the first formulation of a *measurement uncertainty principle*.

The first reformulation of the Heisenberg uncertainty principle was done by *Earle Hesse Kennard* just a few months afterwards [2]. He generalized the Heisenberg relation and used *standard deviations* instead of abstract measurement error operators which were only heuristically motivated.

The Kennard relation is given by

$$\sigma(Q)\sigma(P) \geq \frac{\hbar}{2} , \quad (2.73)$$

with the standard deviation  $\sigma$  of any variable  $X$  being

$$\sigma(X) = \sqrt{\langle X^2 \rangle - \langle X \rangle^2} . \quad (2.74)$$

The fundamental difference to the original formulation is that the uncertainty between the two non compatible observables of the quantum system has now become *intrinsic*. That means, that the uncertainty is *completely independent* of any measurement apparatus and therefore a *fundamental property* of the quantum system.

The Kennard relation is considered the first formulation of a *preparation uncertainty principle*. An example of the relation is given in Fig. 2.5.

In 1929 *Howard Percy Robertson* developed a generalized uncertainty relation that describes the relation between *any two non-commuting* observables  $A, B$  [8]. So far only the uncertainty of the measurement of position and momentum has been discussed.

The Robertson uncertainty relation uses the expectation value of the commutator to describe the tradeoff relation between the observables:

$$\sigma(A)\sigma(B) \geq \frac{1}{2} |\langle \psi | [A, B] | \psi \rangle| , \quad (2.75)$$

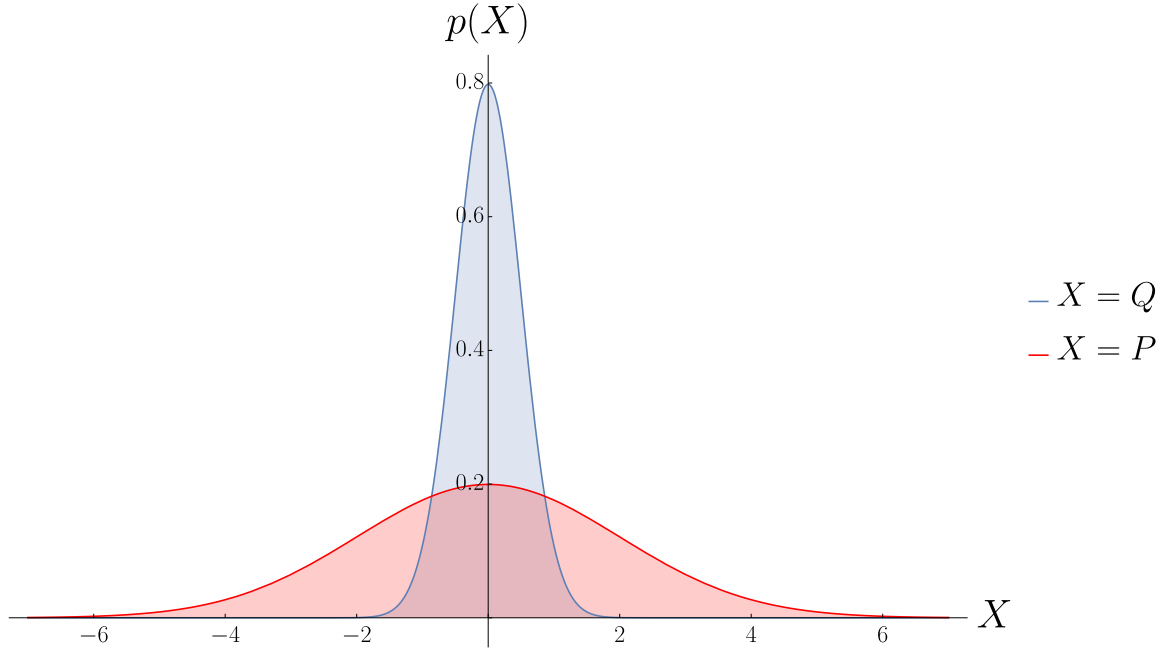


Figure 2.5: Illustration of the Kennard preparation uncertainty of two non-commuting observables  $Q, P$ . The mean value is set to  $\mu = 0$ ,  $\sigma^2(Q) = 0.5$  for the Gaussian distribution used (see Eq. (2.9)).

*Erwin Schrödinger* tightened the Robertson relation in 1930 [9]. To *tighten* an uncertainty relation means, that the lower bound that describes the minimal boundary of the uncertainty gets smaller. A *tight* uncertainty relation represents the case where the smallest possible lower bound has been found.

The result is referred to as *Robertson-Schrödinger uncertainty relation* and is denoted as

$$\sigma(A)\sigma(B) \geq |\langle \psi | \{A, B\} | \psi \rangle - \langle \psi | A | \psi \rangle \langle \psi | B | \psi \rangle|^2 + \left| \frac{1}{2i} \langle \psi | [A, B] | \psi \rangle \right|^2. \quad (2.76)$$

$\{A, B\} = AB + BA$  is called the *anti-commutator* of  $A, B$  and is defined as the sum of both permutations of the operators.

The uncertainty relations we discussed so far use the standard deviation or heuristically motivated symbolic operators as means to mathematically describe the noise and disturbance of a measurement or the intrinsic uncertainty of two observables in a quantum system. In 1983 *David Deutsch* published the article *Uncertainty in Quantum Measurements* [3] where he introduced the Shannon entropy as a more proper means to establish a *tight* uncertainty relation.

The original formulation is given by

$$S_{\hat{A}}(|\psi\rangle) + S_{\hat{B}}(|\psi\rangle) \geq 2 \ln \frac{2}{1 + \sup | \langle a | b \rangle |}, \quad (2.77)$$

with  $|\psi\rangle$  being the state of the system under observation and  $\{|a\rangle\}, \{|b\rangle\}$  being the eigenstates of the operators  $\hat{A}$  and  $\hat{B}$ , respectively.

The entropies  $S_{\hat{X}}(|\psi\rangle)$  are defined as

$$S_{\hat{X}}(|\psi\rangle) = - \sum_x |\langle x | \psi \rangle|^2 \ln |\langle x | \psi \rangle|^2. \quad (2.78)$$

As the  $|\langle x|\psi\rangle|^2$  can be interpreted as probabilities (see Eq. (2.59)) we can write Eq. (2.77) in terms of Shannon entropies (defined in Eq. (2.15)) by dividing by  $\ln 2$ :

$$H(\hat{A}) + H(\hat{B}) \geq 2 \left( 1 + \frac{\ln(1 + \sup | \langle a|b\rangle |)}{\ln 2} \right). \quad (2.79)$$

Note that all uncertainty relations we discussed so far, except the original formulation from Heisenberg, are preparation uncertainty relations. They all describe the intrinsic property of quantum systems and their observables not to be determinable in a perfectly exact manner.

### 2.6.1 Information Theoretic Noise and Disturbance in Quantum Measurements

The uncertainty relation the quantification of noise and disturbance in our experiment is based on is using information theoretic definitions, as first proposed by *Deutsch* (see above) and further used by others (e.g. [23, 24]). So far these works did not deal with measurement related noise and disturbance.

In 2014 *Buscemi et al.* introduced an *entropic measurement uncertainty relation* in their paper *Noise and Disturbance in Quantum Measurements: An Information-Theoretic Approach* [4]. They shown that for any measuring apparatus  $\mathcal{M}$  and any nondegenerate observables  $\hat{A}, \hat{B}$  the relation

$$N(\mathcal{M}, \hat{A}) + D(\mathcal{M}, \hat{B}) \geq -\log_2 c, \quad (2.80)$$

with

$$c := \max_{a,b} \left| \langle \psi^a | \phi^b \rangle \right|^2, \quad (2.81)$$

holds.

The  $|\psi^a\rangle, |\phi^b\rangle$  are the eigenstates of  $\hat{A}$  and  $\hat{B}$ , respectively.

To understand the quantification of noise we will follow the thought experiment presented in the paper. The measuring apparatus gets fed by random eigenstates of  $\hat{A}$  and provides us with two measurement outputs: the measurement outcome  $M = m$  and the disturbed output quantum system  $S'$ . For the noise we do not care about  $S'$  but only the measurement outcome  $m$ . Now we need to quantify the correlation between the eigenvalues  $\xi_a$  and the outputs of the measurement apparatus  $M$  with the help of a *guessing function*. The procedure is shown in Fig. 2.6.

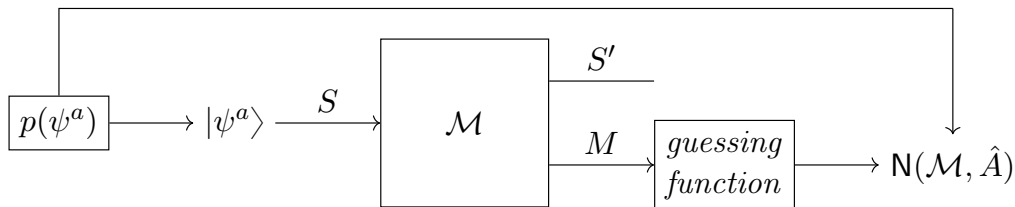


Figure 2.6: Illustration of the definition of the noise. A quantum system  $S$  gets prepared in a state  $|\psi^a\rangle$ . The measurement apparatus  $\mathcal{M}$  is changing the state of  $S$ . We receive an altered quantum system  $S'$  and a measurement outcome  $M$ . A guessing function tries to find the right eigenvalues associated with the measurement output  $M$ . The correlation of the input state and the measurement output is quantified by the noise  $N(\mathcal{M}, \hat{A})$ .

We can describe the measurement of  $m$  given  $|\psi^a\rangle$  with the *conditional probability*  $p(m|\psi^a)$ . We obtain the joint probability by using Eq. (2.12) and receive

$$p(m, \psi^a) = p(\psi^a)p(m|\psi^a) . \quad (2.82)$$

We can fix the  $p(\psi^a)$  as we usually have access to the information on how the states are prepared. For a uniform distribution with  $d$  possible outcomes we get

$$p(m, \psi^a) = \frac{1}{d}p(m|\psi^a) . \quad (2.83)$$

Now we define the noise with the help of the conditional entropy (see Eq. (2.18)) using all permutations of  $p(m, \psi^a)$  and receive

$$N(\mathcal{M}, \hat{A}) = H(\hat{A}|M) = \sum_m \sum_{\psi^a} p(m, \psi^a) \log_2 \frac{p(m)}{p(m, \psi^a)} . \quad (2.84)$$

For the disturbance we feed the measurement apparatus  $\mathcal{M}$  with random eigenstates  $|\phi^b\rangle$  of the observable  $\hat{B}$ . From the measurement apparatus  $\mathcal{M}$  we receive the measurement output  $M$  and the disturbed quantum system  $S'$ . Now, prior to a measurement  $\mathcal{M}'$  of  $\hat{B}$  we feed both outputs of the measurement apparatus  $\mathcal{M}$  into a *correction procedure*  $\epsilon$ . This correction procedure now tries to reconstruct the original quantum system  $S$  from the output system  $S'$  and the information  $m$  of the measurement output  $M$ . The output of  $\epsilon$  is the corrected quantum system  $S''$  on which the measurement  $\mathcal{M}'$  with output  $M' = m'$  gets performed. The measurement  $\mathcal{M}'$  can be of any measurement type that is suited to measure the eigenstates of  $\hat{B}$ , e.g. simply a projective measurement.

Now we once again use the joint probability to quantify the disturbance of the measurement apparatus  $\mathcal{M}$  on the quantum system  $S$  using a correction procedure  $\epsilon$  and assuming a uniform distribution for the input states:

$$p(m', \phi^b) = p(\phi^b)p(m'|\phi^b) = \frac{1}{d}p(m'|\phi^b) . \quad (2.85)$$

We now define the disturbance in terms of Shannon entropy and get

$$D(\mathcal{M}, \hat{B}) = \min_{\epsilon} H(\hat{B}|M') = \sum_{m'} \sum_{\phi^b} p(m', \phi^b) \log_2 \frac{p(m')}{p(m', \phi^b)} . \quad (2.86)$$

The term  $\min_{\epsilon} H(\hat{B}|M')$  means that we take the minimum of the entropy given all possible correction functions  $\epsilon$ . This ensures that the correction function is optimal and the disturbance therefore gets minimized.

In general a correction function  $\epsilon$  can be modeled by any *completely positive, trace preserving map* that reconstructs  $S$  from  $S'$  with the information provided by  $M$ . We can define the probability  $p_e$  of making a guessing error as

$$p_e = \sum_{\phi^b} \sum_{\phi^b \neq m'} p(m', \phi^b) , \quad (2.87)$$

and receive the probability of a correct guess as the complementary probability

$$1 - p_e = \frac{1}{d} \sum_{\phi^b} F\{S''(P_{\phi^b}), SP_{\phi^b}\} . \quad (2.88)$$

We used the *fidelity*  $F\{\rho, \sigma\} := \text{Tr}\{\sqrt{\sqrt{\rho}\sigma\sqrt{\rho}}\}^2$  between two states  $\rho, \sigma$  and the projectors  $P_{\phi^b}$  onto the eigenstates  $|\phi^b\rangle$  of the operator  $\hat{B}$  for the description. The fidelity is a measure that expresses the probability that one quantum state can be rightfully identified as the other one.

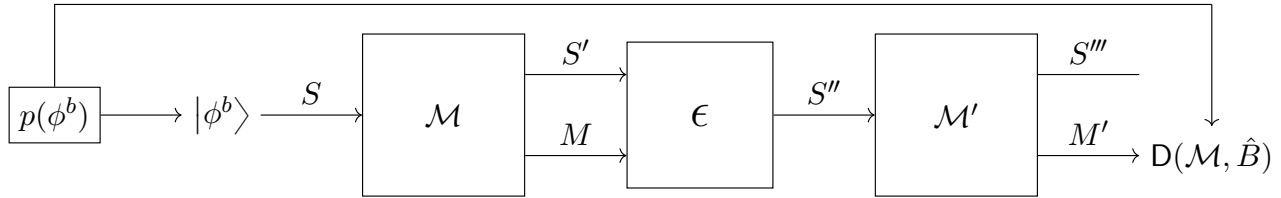


Figure 2.7: Illustration of the definition of the disturbance. The first part is analogue to the noise definition. A quantum system  $S$  is prepared in an eigenstate  $|\phi^b\rangle$ . The system  $S$  gets fed into the measurement apparatus  $\mathcal{M}$ . The output is a quantum system  $S'$ , altered by the measurement, and the measurement output  $M$ . They both get now fed in a correction procedure  $\epsilon$ , which tries to restore the original state  $S$  with the help of the information of  $M$ . The result is the corrected quantum system  $S''$ . A successive measurement apparatus  $\mathcal{M}'$  performs a measurement on  $S''$ . The correlation between the input state of  $S$  and the measurement output  $M'$  is quantified by the disturbance  $D(\mathcal{M}, \hat{B})$ .



Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.  
The approved original version of this thesis is available in print at TU Wien Bibliothek.

---

## Experimental Setup

---

I am afraid neutrons will not be of any use to any one.

---

*Sir James Chadwick*

In this chapter the experimental setup used will be introduced. At the beginning the properties and means of the production of free neutrons will be discussed, as the neutron is the quantum system we are experimenting with. The energy distribution of the neutron beam port we are working on at the TRIGA Mark II fission reactor situated at the Atominstitut will be shown. Then key elements of neutron polarimetry used in the setup will be discussed. Typical steps in the calibration of the components used will be explained and shown exemplarily. Then a short explanation of the experiment by *Sulyok et. al* will be given, as it can be seen as a predecessor of our setup [6]. They used projective measurements for an entropic measurement uncertainty relation. Our measurements, using a POVM, yield in a tighter result. Concluding, the setup specifications used for our noise and disturbance measurements are specified.

### 3.1 The Neutron

In 1920 *Ernest Rutherford* discovered that atoms consist of an electrically positive charged heavy core and an equally negatively charged hull [25]. The proton mass was already known by that time, but an about equal mass could not be associated with any particle already known. Rutherford proposed an electrically neutral nucleon concluding from his research. The issue was resolved as *James Chadwick* experimentally proved the existence of the *neutron* in 1932 [26].

The neutron is a *baryon* and consists of one up-quark and two down-quarks, often denoted as  $|udd\rangle$ . That state is not stable but can be stabilized by binding the neutron to a proton with the *strong force* in the nucleus of an atom. An unbound neutron has a mean lifetime of  $\tau_n = (880.2 \pm 1.0)$  s [27, p. 1653]. This is a reasonable amount of time to perform experiments on and with the help of free neutrons.

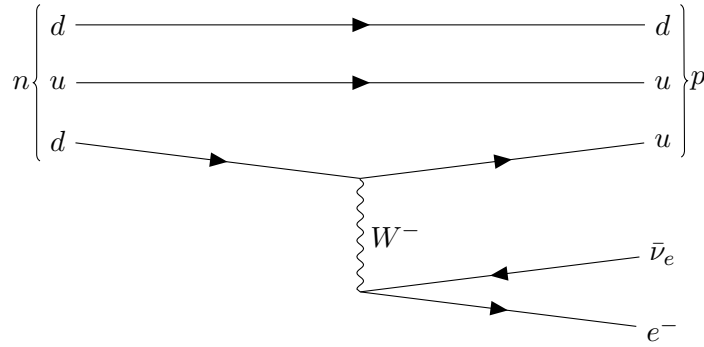


Figure 3.1: Feynman diagram illustrating the decay mode (1) given in Eq. (3.1). A down-quark of the neutron decays into an up-quark, emitting an intermediary  $W^-$  gauge boson. The electron and the electron-antineutrino are generated from that boson.

The free neutron  $n^0$  decays as a  $\beta^-$ -emitter and has three differently probable known decay modes, which can be denoted as

$${}^0_1n \rightarrow \begin{cases} \frac{1}{1}p + e^- + \bar{\nu}_e + Q, & p \cong 1 & (1) \\ \frac{1}{1}p + e^- + \bar{\nu}_e + \gamma, & p = (9.2 \pm 0.7) \times 10^{-3} & (2) \\ \frac{1}{1}H + \bar{\nu}_e, & p < 3 \times 10^{-2} \dagger & (3) \end{cases}, \quad (3.1)$$

with  $p^+$  being a proton,  $e^-$  an electron,  $\bar{\nu}_e$  an electron antineutrino,  $Q$  the decay energy,  $\gamma$  a photon and  $H^1_1$  a hydrogen atom. The probabilities  $p$  corresponding to the decay modes are taken from [27, p. 1655].

Almost every neutron decays through mode (1), with a decay energy of  $Q = 0.782343 \text{ MeV}$  <sup>‡</sup>. A small fraction is through mode (2), where a photon is emitted as an additional particle. It can be thought of as the result of an *internal bremsstrahlung* based on the interaction of the electron and the proton. Mode (3) describes the case where the electron does not gain enough kinetic energy to *escape* the proton and therefore gets bound to it, forming a hydrogen atom. A *Feynman diagram* illustrating the most common decay mode (1) is given in Fig. 3.1.

The name neutron comes from its electrical neutrality as it does not carry any charge. That fact is commonly accepted. The upper limit for a hypothetical charge  $q_n$  of the neutron is currently defined as  $q_n = (-0.2 \pm 0.8) \times 10^{-21} e$ ,  $e$  being the *elementary charge* [27, p. 1654]. This value is consistent with 0, given the experimental uncertainties.

The neutron is considered a spin  $\frac{1}{2}$  particle, hence it belongs to the particle class of *fermions*. According to the standard model of particle physics it carries a very small *electric dipole moment*. The predicted values are a lot smaller than the current precision of experiments investigating it. It also carries a *magnetic dipole moment*. This was controversial for a long time. Usually, to have a magnetic dipole moment, a particle has to have both a spin and an electric charge. In the case of the neutron the spin is a combination of the magnetic moments of the constituent quarks, which could be explained with the introduction of the quark model in the 1960s.

The magnetic dipole moment for the neutron is given as  $\mu_n = g \mu_N = -1.91304273(45) \mu_N$  [27, p. 1653].  $g$  is the so called *g-factor*,  $\mu_N$  is the *nuclear magneton* which is defined as

$$\mu_N = \frac{e\hbar}{2m_p}, \quad (3.2)$$

<sup>†</sup>Please note that there is no sufficient experimental data that can give a probability with proper confidence. It is only possible to define this lower bound, as stated in [27, p. 1655]. In general, this mode is considered the least probable.

<sup>‡</sup>The decay energy is the difference in kinetic energy before and after the decay. It can be computed via the rest masses of the decay which is not explicitly done here.



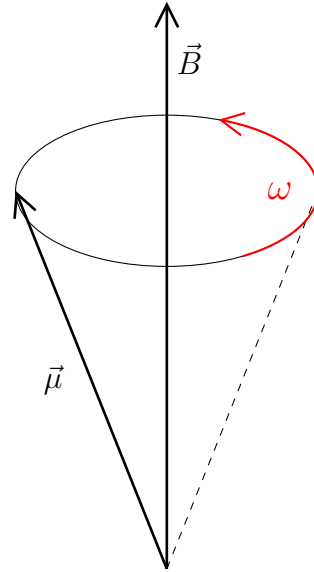


Figure 3.2: Larmor precession of a magnetic moment  $\vec{\mu}$  around an external magnetic field  $\vec{B}$  with angular frequency  $\omega$ .

with  $e$  being the elementary charge,  $\hbar$  the reduced Planck constant and  $m_p$  the proton mass. The magnetic dipole moment is antiparallel to the spin.

As the neutron does not carry any charge, it is not exposed to the *Coulomb force*. Only a very small part of the kinetic energy of the neutron is dependent on a magnetic field interaction. It is neglectable for the neutron energies we experiment with and gets only relevant in very small energy scales. The spin nevertheless orients itself parallel or antiparallel to the magnetic field in an adiabatic case, meaning a slowly changing external magnetic field. It is also possible to influence the direction of the spin with the help of external magnetic fields, which are nonadiabatic according to the following principle.

The change in orientation of the magnetic moment  $\mu_n$  in an external magnetic field  $\vec{B}$  is described via the so-called *Larmor precession*, as shown in Fig. 3.2. A *torque*  $\vec{\tau}$  of the magnetic moment is exerted by the field, mathematically

$$\vec{\tau} = \vec{\mu} \times \vec{B} . \quad (3.3)$$

The magnetic moment precesses about the magnetic field with the *Larmor frequency*

$$\omega = -\gamma |\vec{B}| , \quad (3.4)$$

where

$$\gamma = g \frac{\mu_N}{\hbar} = g \frac{e}{2m_p} , \quad (3.5)$$

is the *gyromagnetic ratio*.

### 3.1.1 Neutron Sources for Research

Neutrons for research are typically produced in research fission reactors, as for example the *Institut Laue-Langevin* in Grenoble, France or the *TRIGA Mark II* research reactor situated at the *Atominstut* of the Technical University Vienna, Austria, where our experiment is located. An alternative for fission reactors are so called *spallation sources*, where highly energetic incident particles cause the target material to dissolve into its constituents, releasing free neutrons. An example for a spallation

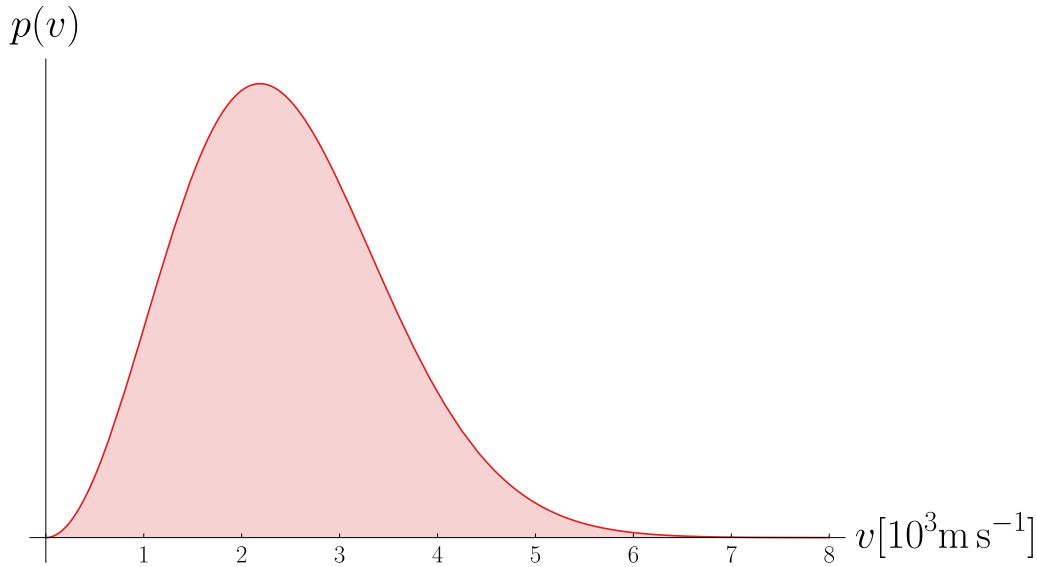


Figure 3.3: Probability density function of a Maxwell-Boltzmann velocity distribution for thermal neutrons at  $T = 290$  K (see Eq. (3.6)).

source is the *Paul Scherrer Institut* near Zurich in Switzerland.

The neutrons that get produced by nuclear fission have a mean kinetic energy in the area of  $\bar{E}_{kin} = 2$  MeV and are called *fission neutrons*. They belong to the energy range of *fast neutrons*. The fast neutrons lose energy by scattering at a *moderator*, as for example hydrogen. In general neutrons with lower energies have a larger *cross section* with matter. That means that interactions with matter, both scattering and absorption, get more probable with lower energies, which is desirable as we get access to more means of manipulation of the neutron. The velocity of the neutrons is described with the *Maxwell-Boltzmann-distribution*.

After a certain number of scattering events, the neutrons arrive at a mean energy of around  $\bar{E}_{kin} = 25$  meV, which is called the *thermal energy range*. The velocity distribution for thermal neutrons is shown in Fig. 3.3.

The probability density function for a Maxwell-Boltzmann velocity distribution  $p(v)$  is given as

$$p(v) = \sqrt{\frac{2}{\pi}} \frac{v^2}{a^3} e^{-\frac{v^2}{2a^2}}, \quad (3.6)$$

with

$$a = \sqrt{\frac{kT}{m}},$$

$m$  being the *particle mass*,  $k$  the *Boltzmann constant* and  $T$  the *temperature*.

Typically, as in our experiment, we want to have so called *monochromatic neutrons*. This means, that the wavelength  $\lambda$  is the same for all neutrons used for our experiment. The connection between the velocity  $v$  of the neutron and its corresponding wavelength is given via the *de Broglie relation*

$$\lambda_{deBroglie} = \frac{h}{m_n v}, \quad (3.7)$$

with the Planck constant  $h$ . We used the rest mass of the neutron  $m_n$ , as  $v \ll c$  for thermal neutrons. Therefore no relativistic considerations are necessary. The natural constant  $c$  is the speed of light.

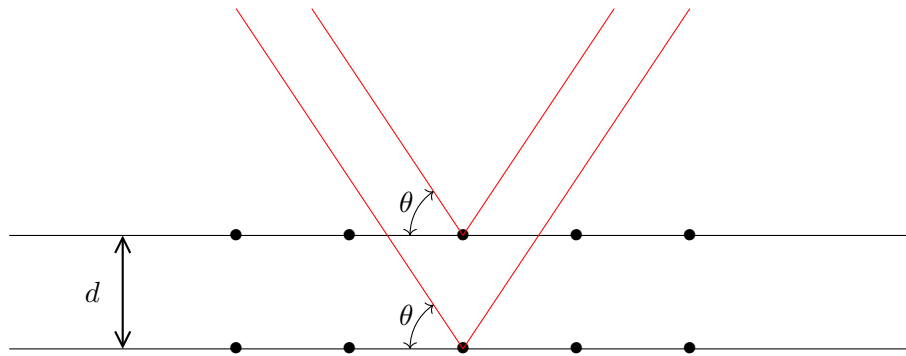


Figure 3.4: Illustration of the Bragg principle. Incoming beams (in red) interfere constructively if the Bragg condition (Eq. (3.8)) holds. Therefore one can monochromatize, as all wavelengths other than the desired one do not interfere constructively and get suppressed. The intensity of the outgoing beam decreases with higher diffraction orders.

The monochromatization is performed by a *pyrolytic graphite crystal*. It is a mosaic crystal, meaning that it consists of numerous perfect crystals that are randomly misoriented to some extent. This ensures a higher *spectral width* of  $\Delta\lambda/\lambda = 0.015$ . If, e.g. we would use a perfect Si-crystal, we would have much sharper peaks in our energy spectrum but significantly less neutron flux. In our case the sharpness of the peak is not as important as a high monochromatic neutron flux, therefore a mosaic crystal is the monochromator of choice.

Even though the theoretical description of a mosaic crystal is given by the so called *Darwin-Hamilton equations*, a simplified description of the principle of monochromatization is given by the *Bragg law*

$$n\lambda = 2d \sin \theta , \quad (3.8)$$

where  $n \in \mathbb{N}$  is the diffraction order,  $d$  is the distance between two lattice planes and  $\theta$  is the angle between the incident beam and the lattice plane. Constructive interference of the incident beam, reflected at different lattice planes, occurs when the Bragg law is fulfilled. Therefore one can tune the desired wavelength by choosing a crystal with the right lattice plane structure and finetune by adjusting the incident angle. The effect is illustrated in Fig. 3.4.

The resulting velocity distribution for the incident beam of our experiment is given in Fig. 3.5.

### Why we use neutrons for our research

We can now summarize the main reasons why neutrons are a good fit for us to be used as a quantum system to experiment with:

- The detection efficiency of neutrons is close to unity ( $> 0.99$ ).
- The efficiency of the spin manipulation is very high. For a well adjusted spin flip coil it is close to 100 %.
- The lifetime of the free neutron is high enough to perform measurements on them without time being an issue.
- The neutron has less interaction with matter compared to e.g. photons or electrons. Therefore more sophisticated experimenting environments, like a vacuum or special radiation shielding are not necessary.

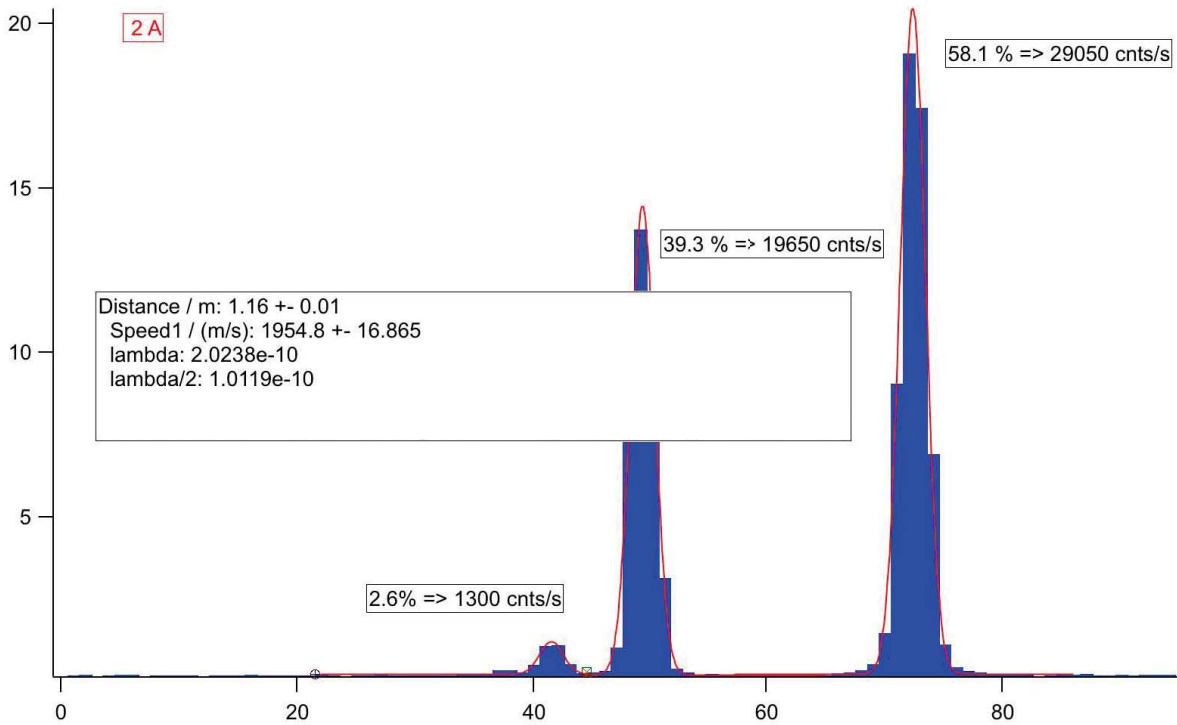


Figure 3.5: Old measurement of the incident beam energy spectrum with the Time-Of-Flight method. The units are arbitrary. We see a major peak at  $\sim 2 \text{ \AA}$ . The other peaks are higher orders from the monochromator crystal.

- As the neutron does not carry a electric dipole moment but is a spin  $\frac{1}{2}$  fermion we can alter the orientation of the spin with the help of magnetic fields without *significantly* changing the neutron flight path (The kinetical energy is marginally dependent on the orientation of the magnetic dipole moment. For thermal neutrons and the magnetic field strengths used in our setup this effect is neglectable).
- The neutron represents a relatively easy to handle qubit system which is necessary to test our Noise-Disturbance relations.

## 3.2 Elements of Neutron Polarimetry

In neutron polarimetry the modification of the spin of the neutron as well as the polarization and detection of neutrons are key requirements to form a proper experimental setup. In the following the main elements that help us achieve those operations are introduced and the configuration and calibration of these elements are shown exemplarily for better understanding.

### 3.2.1 Neutron Supermirrors

Neutron supermirrors polarize an incoming neutron beam. The neutrons coming directly from the monochromator have an arbitrary spin orientation. If we would measure the spin orientation of this beam with respect to one of the cartesian axes we would expect an uniform distribution (compare Section 2.4.2). When performing a measurement, the arbitrary spin state jumps into a well defined post measurement state.

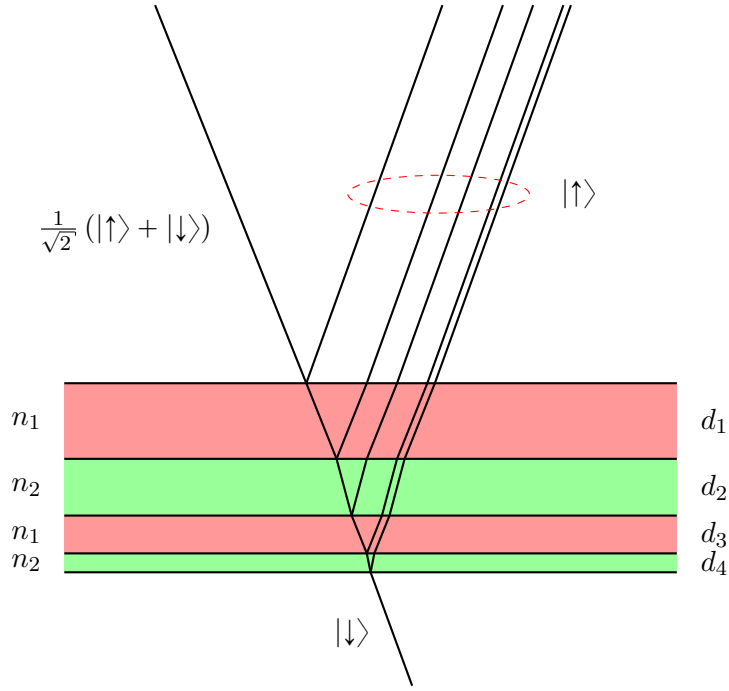


Figure 3.6: Principle of a neutron supermirror. Alternating layers of magnetic and nonmagnetic materials are used. The diameter of the layers is monotonously decreasing ( $d_1 > d_2 > d_3 > \dots$ ). An incoming beam with arbitrary spin orientations gets split into two beams in a pure spin state (ideal case). The undesired spin state (here:  $|\downarrow\rangle$ ) gets absorbed later on by a substrate, the other polarized beam leaves the supermirror.

This effect is called *collapse of the wavefunction* and is described in detail in Section 2.4.1.

The neutron passing a supermirror can be interpreted as an alteration of the incoming wavefunction, selecting only one of the two possible spin states,  $|\uparrow\rangle$ , suppressing  $|\downarrow\rangle$ . As the wavefunction which describes the premeasurement state collapses, the postmeasurement state is then defined as being purely in  $|\uparrow\rangle$ . We lose intensity but receive a polarized beam of neutrons.

Of course a real supermirror is not ideal, the quality of it is mostly expressed in terms of *polarization*  $P$ . It is defined analogously to the *Michelson contrast* (see Eq. (4.5)) and therefore as

$$P = \frac{I_{\uparrow} - I_{\downarrow}}{I_{\uparrow} + I_{\downarrow}}, \quad (3.9)$$

where  $I_{\uparrow}$  respectively  $I_{\downarrow}$  are the intensities of the neutron beams with the two possible spin orientations  $|\uparrow\rangle$  and  $|\downarrow\rangle$ . For an ideal supermirror  $I_{\downarrow}$  vanishes and we receive a polarization of  $P = 1$ . A well aligned supermirror which selects  $|\uparrow\rangle$  has a polarization in the area of  $P \gtrsim 0.98$ . Note, that  $-1 \leq P \leq +1$ , as a beam can be polarized both in  $|\downarrow\rangle$ - and  $|\uparrow\rangle$ -direction and therefore we get  $P < 1$  for  $I_{\uparrow} < I_{\downarrow}$ . This differs from the contrast  $C$ , which is fixed as  $0 \leq C \leq 1$ .

The description of the interaction of (nonrelativistic) neutrons and matter is given by the time independent *Schrödinger equation* with a potential  $V$ .

The Schrödinger equation in its time independent form is given by

$$\left( \vec{\nabla}^2 + \frac{2m(E - V(\vec{r}))}{\hbar^2} \right) \Psi(\vec{r}) = \left( \vec{\nabla}^2 + k^2 \right) \Psi(\vec{r}), \quad (3.10)$$

where  $E$  is the kinetic energy and  $k$  is the wave vector of the wavefunction  $\Psi(\vec{r})$ .

The potential consists of a term for the nuclei the neutrons interact with and a magnetic term. The term for the nuclei can be described via a *Fermi pseudopotential*. We receive for the potential

$$V = V_{nuc} + V_{mag} = \underbrace{\frac{2\pi\hbar^2 N}{m} b_c}_{\text{fermi potential}} \underbrace{\pm \mu B_{eff}}_{\text{spin dependent}} . \quad (3.11)$$

The Fermi potential is dependent on the mass  $m$  of the nuclei, the nuclei density  $N$  and the scattering length  $b_c$ .

From the Schrödinger equation we can directly derive that the wave vector length  $k$  is given by

$$k = \sqrt{\frac{2m(E - V(\vec{r}))}{\hbar^2}} . \quad (3.12)$$

The refractive index  $n$  of a material is given as the ratio of the wave vector  $k_0$  in absence of a potential ( $V = 0$ ) and the wave vector with an additional potential and therefore

$$n = \frac{k}{k_0} = \sqrt{1 - V/E} . \quad (3.13)$$

We see, that we receive two different refractive indices for different spin orientations because of the magnetic term in the potential. This effect is used in supermirrors, we can differentiate the incident beam in two subbeams. By choosing a material with a specific Fermi pseudopotential  $V > E$  we can also get  $n \in \mathbb{C}$ , which yields in absorption or incoherent scattering.

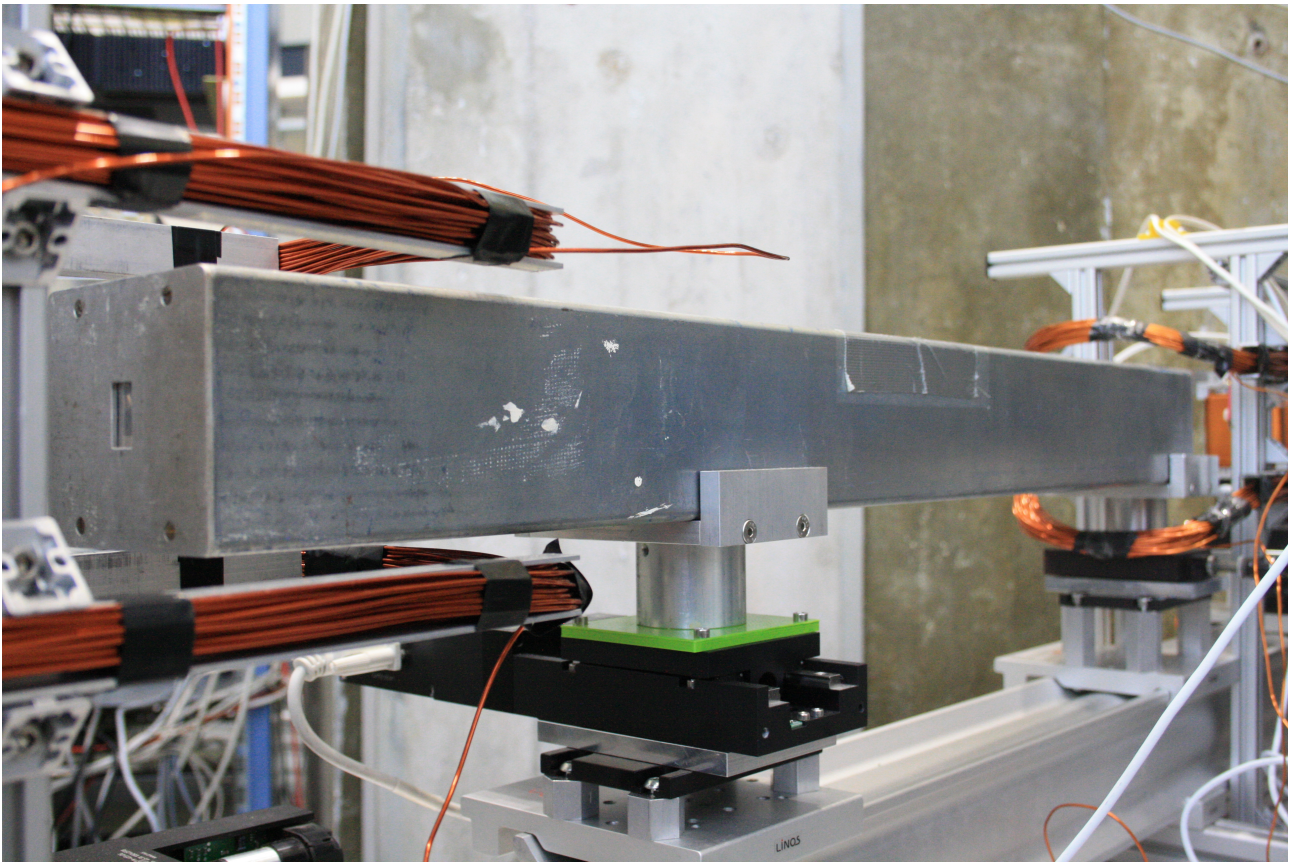


Figure 3.7: Supermirror used in our setup. The exit window for the beam is visible on the left side of the mirror.

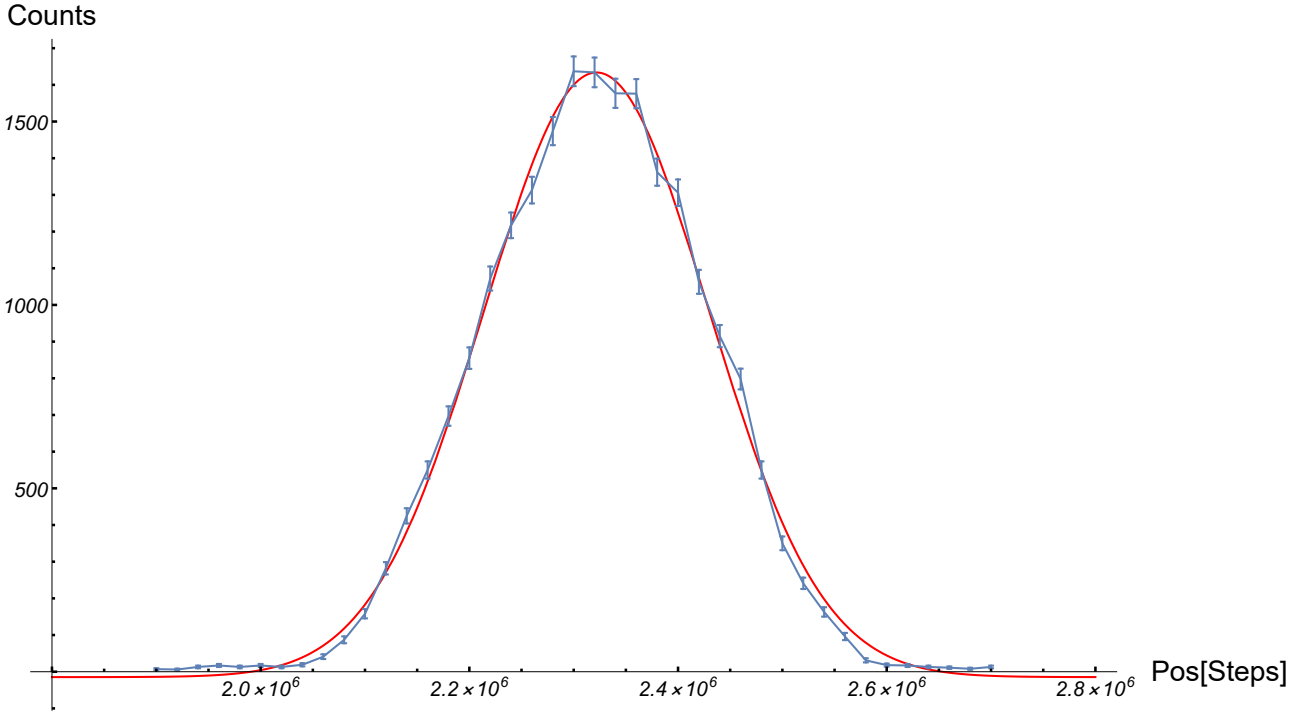


Figure 3.8: Calibration measurement of the supermirror tilt and the resulting intensities dependent on the tilt angle  $\theta$ . The actual values on the  $x$ -axis are positions of a stepper motor performing the tilt and are therefore not relevant. The full range of the plot covers an angle of only about  $1.3^\circ$ .

A neutron supermirror consists of alternating layers of magnetic and nonmagnetic materials. The layer thickness decreases with the layers. The basic structure is shown in Fig. 3.6. There are material combinations that have quite different refractive indices for different spin alignments. Therefore it is technically possible to create supermirrors with high reflectivity for one spin component but high transmittivity for the other one. For the  $|\uparrow\rangle$  component a refractive index of  $1 - n \sim 10^{-6}$  is usual. The material combination used for our supermirrors is *Co* and *Ti*.

In theory, we then would get an intensity at the output of the supermirror that is half of the incoming beam, as we expect half of the neutrons of the incoming beam to be selected by the supermirror. Practically due to material losses, and as the incident neutron beam is not a plain wave and therefore the neutrons impact the supermirror not only from top but also sideways the total intensity further decreases. We end up with an output intensity in the range of  $I_{out} \cong \frac{1}{4}I_{in}$ .

We can also use the supermirror as an analyzer of the neutron spin state. When an incoming neutron beam is in a state  $|\psi(\theta, \phi)\rangle$  on the Bloch sphere (see Eq. (2.56)), we expect an intensity that is proportional to the overlap of the incoming state and the state  $|\uparrow\rangle$ , which gets transmitted, mathematically

$$I(\phi, \theta) \propto |\langle \uparrow | \psi(\theta, \phi) \rangle|^2. \quad (3.14)$$

A typical supermirror looks basically like a lengthy cuboid that has entrance respectively exit windows for the neutron beam on the smaller faces. A picture of one of the supermirrors used in our experiment is shown in Fig. 3.7. It should be aligned in a way that the center of the entrance window matches the maximum of the neutron beam, as we want the highest intensity possible to enter the supermirror in order to lose as few neutrons as practicable. This is mostly done by hand monitors or neutron detectors fixed in front of the entrance window.

After the alignment of the entrance window we need to properly adjust the tilt of the supermirror relative to the neutron beam. When tilting the supermirror we receive a Gaussian function alike shape for the output intensities measured. An example for a calibration measurement is given in Fig. 3.8. It is important to note that this measurement only gives us information about the intensity of the outgoing neutron beam but not at all about the direction of the magnetic moment of the neutrons, as neutron detectors are not spin sensitive.

In fact, choosing a tilt that configures the supermirror position in the area of the maximum of the distribution is a bad choice, as the degree of polarization in the maximum is rather low. Good values for the degree of polarization of the neutron beam are in the flanks of the distribution. Therefore there is always a tradeoff between intensity and degree of polarization of the beam exiting a supermirror. Dependent on the actual demands of the experiment one has to find the best setup.

### 3.2.2 Guidefield

The degree of polarization of a neutron beam decreases if we do not ensure a static magnetic field that has the same direction as the magnetic dipole of the beam. This is the case, as the spin is always prone to stray fields from the direct environment or the magnetic field of the earth. This means, that in the areas between supermirrors it is necessary to constitute a strong magnetic guide field that has the same direction as the magnetic field of the supermirror.

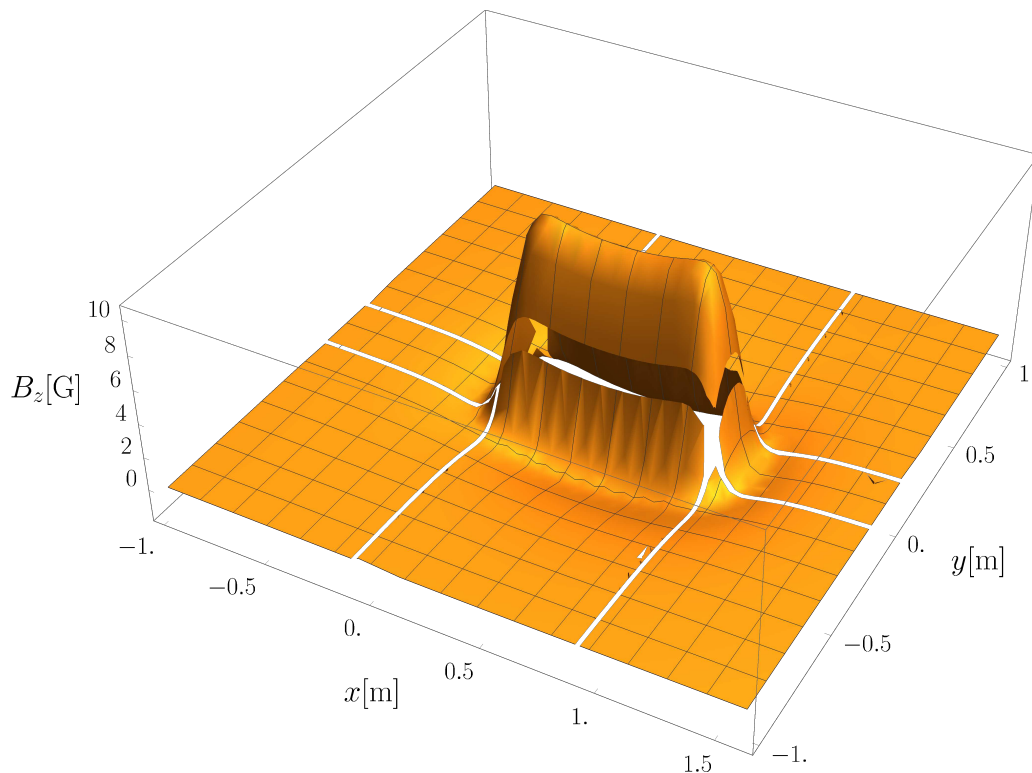


Figure 3.9: Simulation of the field caused by the rectangular Helmholtz coils used in the setup. The plot shows the field distribution in  $z$ -direction for different  $x$ - and  $y$ -values. The 2D plot (see Fig. 3.10) is the cut of this 3D plot at  $y = 0$ . The magnetic field in the area where the neutron beam is located is very homogeneous. The white areas are due to a divergence of the integral at the source of the magnetic field. The field inside the coils is not computed.



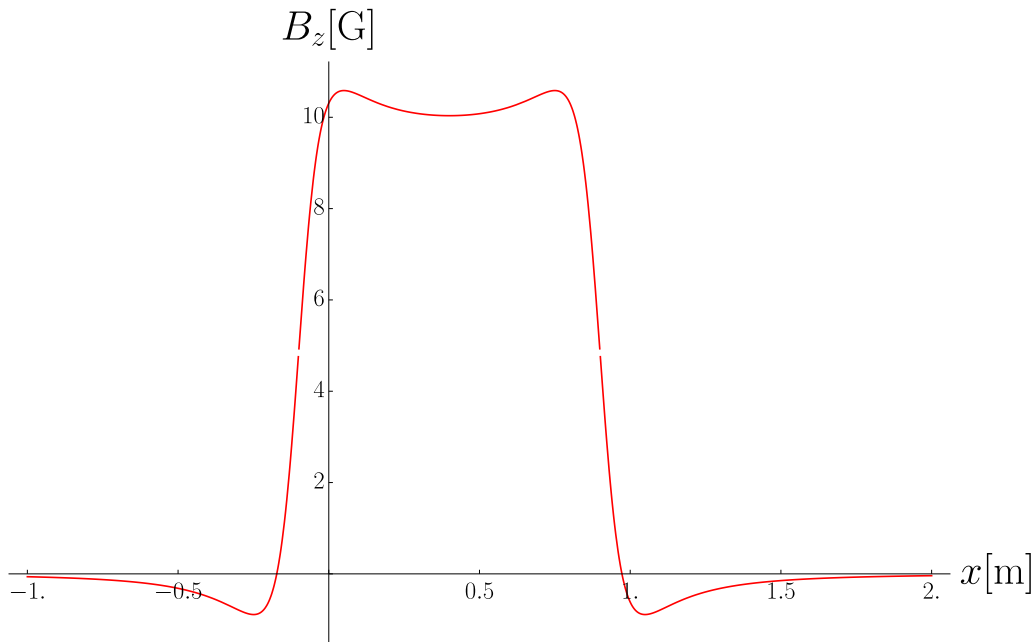


Figure 3.10: Simulation of the magnetic field of the guidefield. The plot shows the magnetic field in the middle of the coil setup over the full  $x$ -range. The neutron beam is most likely to travel in the center in a well aligned setup. Therefore this simulation shows the magnetic field the beam is most likely to be exposed to. The field strength in the middle is in the area of  $B_z \sim 10$  G.

In this case the spin follows the laws of the Larmor precession, which was introduced in Eq. (3.3). For an ideally polarized beam the magnetic moment of the neutrons and the guide field are parallel and the torque of the spin vanishes. Therefore the degree of polarization is conserved.

When they are not parallel the magnetic moment begins to precess about the guidefield direction. The projection relative to the axis of the guidefield is then also conserved (e.g., when a magnetic moment precesses about a magnetic field in  $z$ -direction, the expectation value of the spin in  $z$ -direction is conserved. This does not hold for the other directions.)

To achieve a relatively homogeneous magnetic field in  $z$ -direction *Helmholtz coils* are used. Helmholtz coils consist of a pair of two equal round coils which are aligned in parallel. The radius of the single coils is equal to the distance of the coils. The derivatives of first, second and third order of the magnetic field between the coils vanish. This guarantees a perfect homogeneous magnetic field between the coil pair in the ideal case.

Due to spacing reasons we did not use round Helmholtz coils but a rectangular setup. Simulations of the magnetic field in  $z$ -direction (parallel to the magnetic field of the supermirrors) have been done to assess the quality of the field before manufacturing and installing the guidefield coils. The results of the simulation are shown in Figs. 3.9 and 3.10. Our guidefields operate at  $B_z \sim 14$  G in the center of the field. The values received are consistent with the simulations. The higher field strength compared to the simulated values is given due to higher currents used in the setup compared to the simulation. The fields were simulated with a current of  $I = 5$  A, the actual current used was  $I \sim 6.5$  A for both guidefield coils.

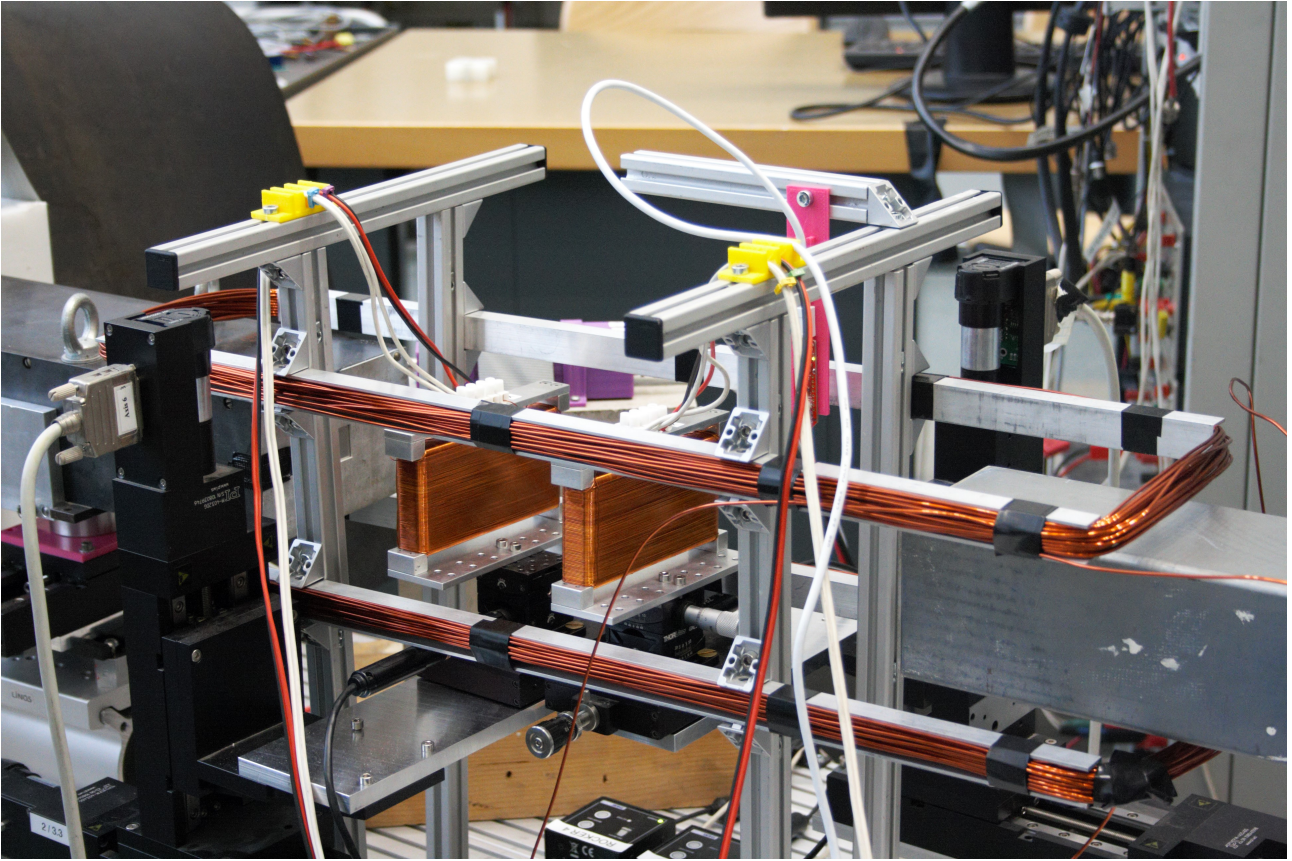


Figure 3.11: Rectangular Helmholtz coils forming a guidefield. The adjacent supermirrors and two spinrotator coils in the area of the guidefield are also visible. The spinrotators are fixed on a pedestal that can be moved by stepper motors to adjust the distance for a full Larmor precession.

We apply two regions with a guidefield in our setup. They are placed between the supermirrors in use. They are of identical construction. A picture of one of the guidefields and two spinflipper coils is given in Fig. 3.11.

### 3.2.3 Spinrotator Coils

The spin of the neutrons can be tuned by external magnetic fields. After passing a supermirror the neutron beam is in a defined spin state. In our case that is  $|\uparrow\rangle$ . The degree of polarization of a neutron beam is then conserved as the neutron beam travels through a guidefield that is oriented in  $z$ -direction and therefore parallel to the spin orientation. A spinrotator coil is placed inside the guidefield. We want to change the spin orientation in a way that it precesses about the  $x$ - $z$ -plane. Again we can describe the time evolution of the spin orientation with the help of the Larmor precession, given in Eq. (3.3). We see that for a magnetic field purely in  $y$ -direction we achieve the desired movement of the magnetic moment.

A spinrotator coil actually consists of two rectangular coil windings that share a common frame. A picture of a typical spinrotator coil is given in Fig. 3.12. One winding generates a magnetic field in  $z$ -direction, antiparallel to the guidefield. The coil needs to be designed and the operating current needs to be calibrated in a way, that the guidefield can be compensated locally inside the coil. The resulting field if we would just use this single winding would in the ideal case therefore be  $\vec{B} = 0$ .

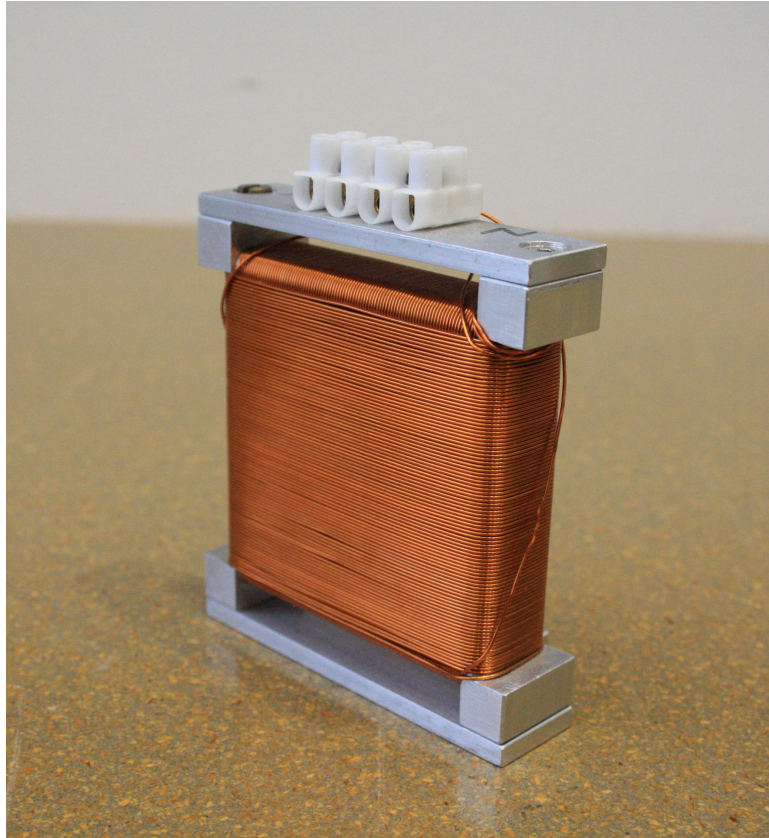


Figure 3.12: Picture of a typical spinrotator coil. The current is applied via the screw joints on top. Coils of the same dimensions are used between the polarizer and first analyzer of our setup.

To calibrate the  $z$ -field of the coil, we fix a current for the  $y$ -winding, typically one that represents a  $\pi$ -flip. Then we sweep the current applied to the  $z$ -coil and receive a measurement curve that can be described by a 2nd order polynomial. The ideal setting lies in the minimum of the measurement curve. In this case we receive the minimal countrate, as our resulting state is purely  $|\downarrow\rangle$ . Only for an ideally compensated  $z$ -field we receive a pure  $y$ -field and are able to perform a full  $\pi$ -flip. Nevertheless we get a small number of counts due to imperfections of both the coil itself, the adjustment and the supermirror. A typical calibration measurement for the  $y$ -field is shown in Fig. 3.13.

The second winding is physically rotated by  $\frac{\pi}{2}$  about the first winding. This guarantees that the magnetic fields generated by the spinrotator coil are perpendicular to each other. The coil generates a field in  $y$ -direction that can be tuned by the current applied. For that reason the alignment of the spin can be changed by applying different currents. We finally receive for the magnetic field  $\vec{B} = B_y \propto I_y$ .

We expect a sinusoidal shape when we measure the counts for different  $y$ -currents. The coil needs to be aligned respectively to the guidefield in a way that the  $z$ -fields are perfectly compensated. The  $y$ -winding is then automatically adjusted perpendicular to the guidefield due to the design of the spinrotator. This alignment is very sensitive and performed via stepper motor driven goniometer stages. For a very well adjusted spinrotator coil we can reach contrasts of up to  $C \cong 0.98$  and a symmetry of  $(90 \pm 1)^\circ$  for the sine fit (For  $90^\circ$  we would have an ideal mirror symmetry about  $I_y = 0$ ). A measurement of a well adjusted spinrotator coil is given in Fig. 3.14.

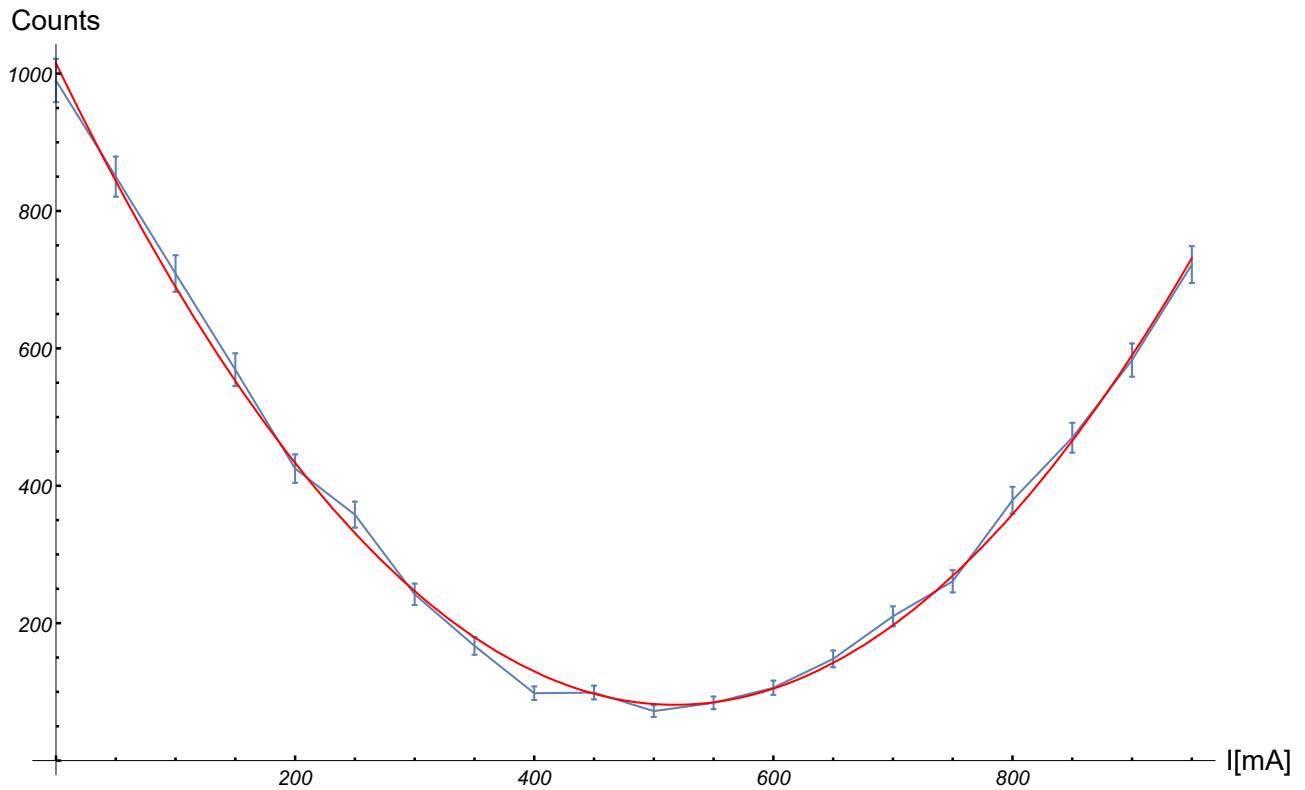


Figure 3.13: Current sweep for the z-winding of a spinrotator coil. The ideal setting is found in the minimum.

In our setup we use pairs of two spinrotator coils. With the first one a desired spin state can be prepared, the second one turns the Bloch vector of a measurement instruction, e.g. a POVM or a projective measurement, in  $|+z\rangle$  direction. With a supermirror as spin state analyzer positioned after coil two we can therefore perform a quantum mechanical measurement on any Bloch vector on the  $x$ - $z$ -plane.

We must not forget that after performing a rotation operation with the first coil the neutrons enter the guidefield again and begin to precess about its direction. Therefore the distance between the spinrotator coil pair must be fixed to a value that represents an integral multiple of full rotations about the magnetic field. This is done by performing a  $\frac{\pi}{2}$  flip with the first and a  $-\frac{\pi}{2}$  flip with the second coil while varying the distance. For the maximal countrate of the resulting measurement plot an ideal setting is found. An exemplary measurement is given in Fig. 3.15. An alternative approach would be performing two consecutive  $\frac{\pi}{2}$  flips and looking for a minimum.

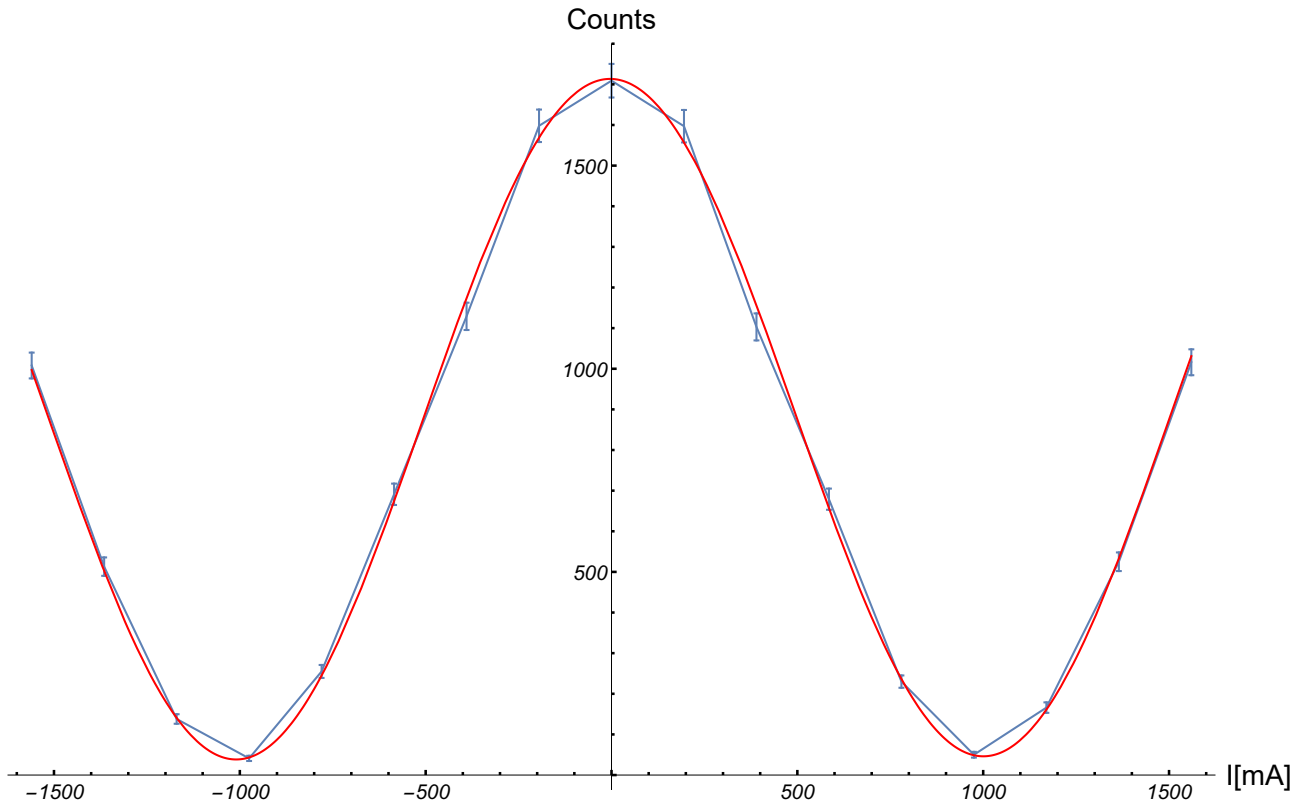


Figure 3.14: Current sweep for the y-winding of a spinrotator coil. With different currents we can alter the spin orientation by performing different spin rotations.

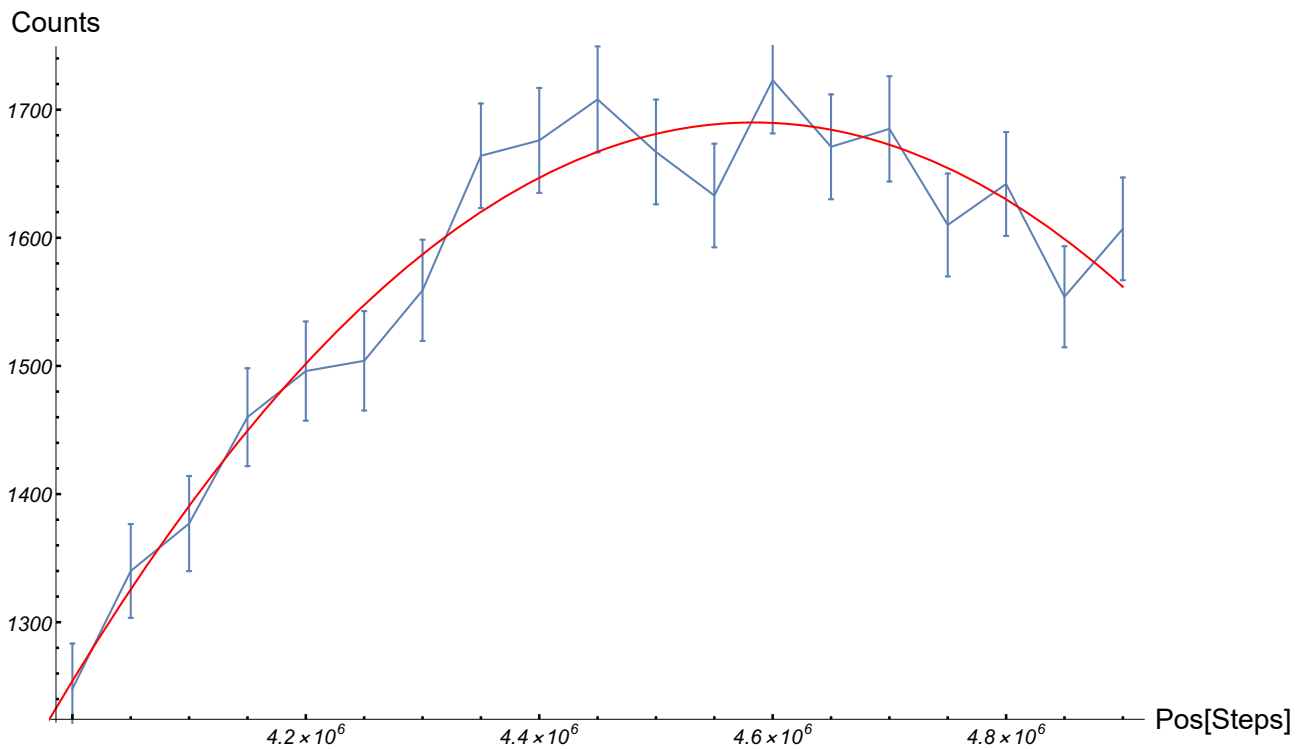
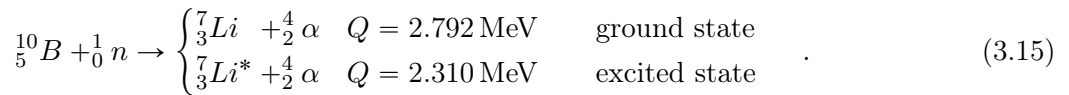


Figure 3.15: Distance measurement of a spinrotator pair. The distance representing a full Larmor precession is given for the maximum of the countrate The distance between the coils is varied with a stepper motor. The positions of the spinrotator coils DC3 and DC4 and the stepper motor that varies the distance is shown in Fig. 3.11

### 3.2.4 Neutron Detectors

To detect neutrons, it is necessary to find neutron-nucleus reactions that have a high cross section and are well distinguishable from a  $\gamma$ -background. The emitted particles must also be easily detectable. When those two properties are fulfilled, it is possible to manufacture compact and highly effective neutron detectors. It is also desirable that the isotopes used have a high natural occurrence within the element. In that case it is cost efficient to enrich the needed isotope. Nuclear reactions of other isotopes do not disturb our detector signal significantly for highly enriched detection elements.

Two elements which have a high cross section for neutron capture in the thermal range are  ${}^3\text{He}$  and  ${}^{10}\text{B}$ . The  ${}^{10}\text{B}(n, \alpha)$  reaction is proper to convert an incident neutron into an easily detectable  $\alpha$ -particle. The reaction is denoted as



In both cases a charged  $\alpha$ -particle is emitted. In most cases  $\text{Li}$  is excited and then decays relatively fast, emitting a  $\gamma$ -ray. Boron trifluoride  $\text{BF}_3$  is used as a detection gas. The boron is highly enriched with  ${}^{10}\text{B}$ , which results in a higher detection efficiency.

The other nuclear reaction that is relevant for us is  ${}^3\text{He}(n, p)$ , which is given as



The cross section for this reaction is even higher than the one for  ${}^{10}\text{B}(n, \alpha)$  and therefore is very attractive in terms of neutron detection efficiency. As  $\text{He}$  is a noble gas we can only use gas detectors with it. A downside is the very small natural occurrence of  ${}^3\text{He}$ , therefore the costs are relatively large, as a high purity is necessary.

We use so called proportional counter tubes filled with detector gas in our setup. The main measurements have been done with a  $\text{BF}_3$  detector, calibration measurements with different sizes of  $\text{He}$  detectors. The principle of a proportional counter tube is simple. A tube is filled with detector gas. In the center of the tube a conducting wire is positioned. A high voltage is applied between the detector casing and the conducting wire. A neutron passes the detector casing and hits a nucleus of the detector gas. As described above, a charged particle is emitted when the neutron gets absorbed by the nucleus. This particle is referred to as *primary ion*, as it is directly caused by neutron capture. Due to its charge, the emitted particle is drawn to the wire or the casing. On its way it interacts with the other atoms or molecules in the detector gas, *secondary ions* get emitted (mostly electrons), that cause a charge avalanche. When the ions hit the electrodes of the detector, a pulse is generated via a load resistor.

The relative increase in ions per distance is described by the *Townsend equation*

$$\frac{dn}{n} = \alpha dx , \quad (3.17)$$

where  $n$  is the number of ions,  $x$  is the distance and  $\alpha$  is the so called *first Townsend coefficient* which is dependent on the electric field strength and therefore the tube diameter, the applied voltage and the position where the ionization incident occurs.

For the right setting of the detector voltage it is possible to receive a pulse amplitude spectrum that is not dependent on the position of the nuclear reaction caused by neutron absorption in the tube. We operate in the *proportional* area of the detector. To find the right voltage for the detector, it is placed in front of a neutron beam with constant intensity. For a sweep of the voltage we get different counts

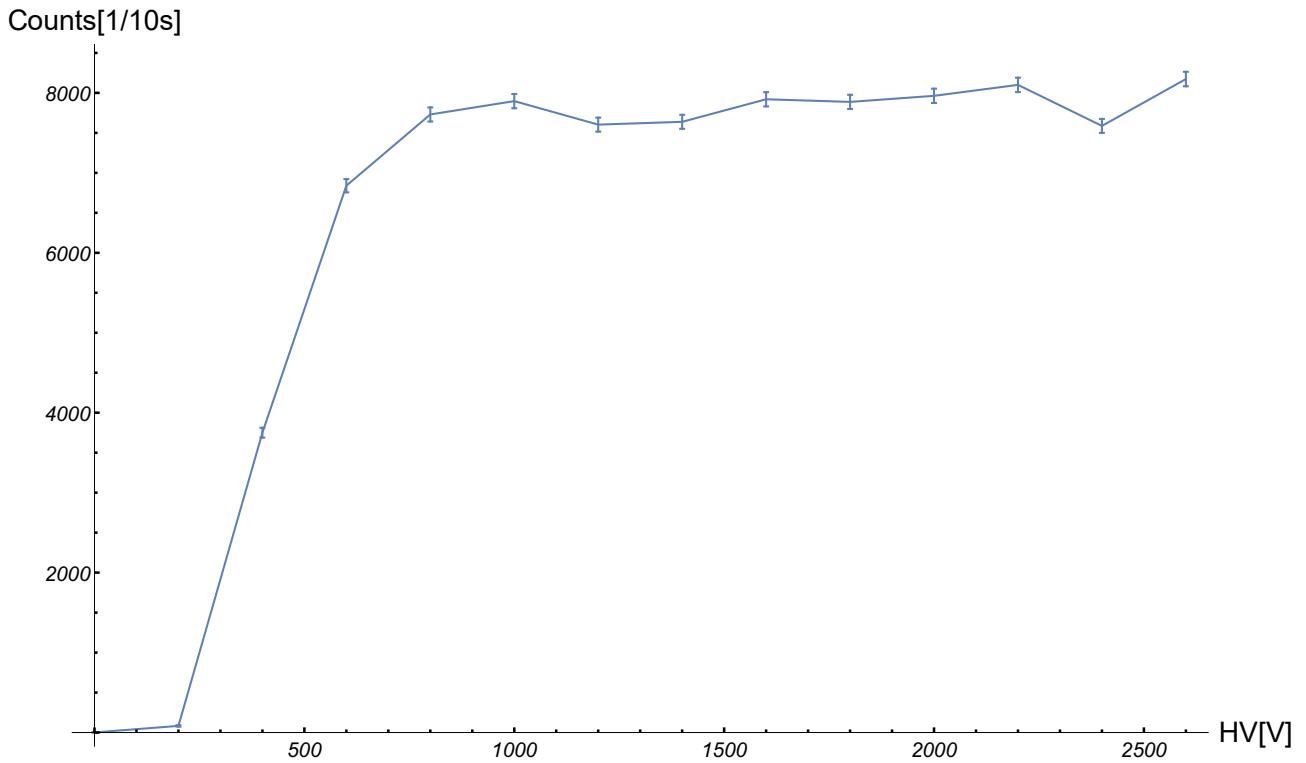


Figure 3.16: Calibration measurement of the neutron detector used as monitor for the setup. A high voltage (HV) is applied, the corresponding countrates are measured. The final setting for the HV was chosen at  $U = 1$  kV. For this setting we are in saturation and a further increase would not improve the detection quality.

for the same time intervals. When the curve flattens and reaches a limit, a further increase of the detector voltage is not expedient. The curve for the monitor detector used in the setup is shown in Fig. 3.16.

The raw pulses from the detector get amplified and forwarded to a *discriminator*. When we look at the spectrum of an (amplified) detector signal, we also see a  $\gamma$ -peak and some detector specific effects that cause pulses that do not belong to an incident neutron. The discriminator only selects a specific range of pulse amplitudes that can be associated with a detected neutron and discriminates other pulses. The spectrum of the neutron detector used as monitor is shown in Fig. 3.17.

Neutron detectors are available in different sizes. For example for the adjustment of the supermirror entrance windows neutron detectors with a smaller diameter were used. They were placed directly in front of center the entrance window with a mount. With this setup we have a high sensitivity for the position of the maximum of the neutron beam and are able to align more sensitively. For the detector used for the disturbance measurement on the other hand it was crucial to measure neutrons that left the last supermirror. Due to the application of POVM weights via the supermirror tilt the position of the neutron beam varied over several cm. A bigger detector was chosen to guarantee a high detection rate over the full range of the horizontal motion of the supermirror. A picture with two detectors in use for different alignment measurements is given in Fig. 3.18.

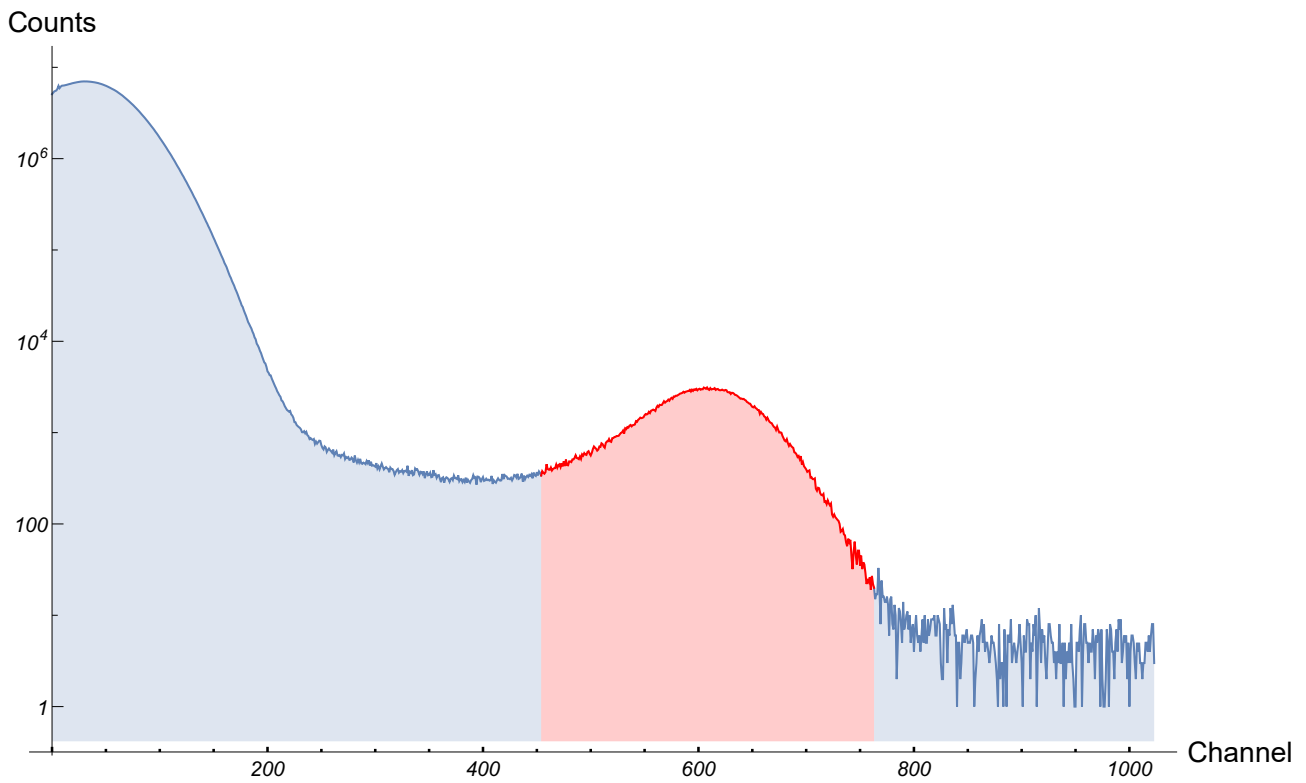


Figure 3.17: Spectrum of the  $He$  detector used as a monitor in our setup, measured with a multichannel analyzer. The domain that is associated with an incident neutron is highlighted in red. Only in that energy range occurring pulses are considered neutrons and get counted. The  $\gamma$ -peak is well recognizable on the left side of the plot.



Figure 3.18: Two differently sized  $He$  neutron detectors. Both of them were used during the calibration of the final setup for our noise-disturbance measurements.



### 3.3 Setup for Noise and Disturbance Measurements

#### 3.3.1 POVM and Correction Procedure

The POVM we use for our measurements was proposed by *Abbott and Branciard* in 2016 [5]. They introduced a 3-outcome POVM

$$M^\theta = \{M_{-1}^\theta, M_0^\theta, M_{+1}^\theta\} \quad \text{for } \theta \in \left[0, \frac{\pi}{2}\right], \quad (3.18)$$

with the elements of the POVM being denoted as

$$M_m^\theta = p_m(\mathbb{1} + \vec{n}_m \vec{\sigma}). \quad (3.19)$$

$\vec{\sigma}$  is the vector holding all Pauli matrices, mathematically

$$\vec{\sigma} = \begin{pmatrix} \sigma_x \\ \sigma_y \\ \sigma_z \end{pmatrix}. \quad (3.20)$$

The orientation of the measurement vectors within the Bloch sphere is given by the Bloch vector

$$\vec{n}_m = \begin{pmatrix} (-1)^m \cos(m\theta) \\ 0 \\ \sin(m\theta) \end{pmatrix}. \quad (3.21)$$

The  $p_m$  can be interpreted as probabilities that a certain outcome of the POVM is measured. They will further be referred to as *POVM weights*. They are defined as

$$p_0 = \frac{\cos \theta}{1 + \cos \theta}, \quad (3.22)$$

$$p_{-1} = p_{+1} = \frac{1}{2(1 + \cos \theta)}. \quad (3.23)$$

An illustration of  $M_m^\theta$  on the x-z-plane of the Bloch sphere is given in Fig. 3.19.

For the disturbance measurement a correction procedure is necessary in addition, as stated in Section 2.6.1. *Abbott and Branciard* proposed a correction procedure  $\epsilon = \{\epsilon_m\}_m$  with

$$\epsilon_0(\rho) = \rho, \quad (3.24)$$

$$\epsilon_{-1}(\rho) = \epsilon_{+1}(\rho) = \frac{1}{2}(\mathbb{1} - \sigma_x), \quad (3.25)$$

where  $\rho$  is the density matrix of the post POVM measurement state. This resembles leaving the state unchanged for  $m = 0$  and mapping  $\rho$  onto the negative  $x$ -axis for  $m = -1, +1$ .

The correction actually performed in our experiment uses a different correction operation for  $m = 0$ . It was proposed by *Stephan Sponar* [28]. Simulations have shown that the correction procedure is equivalent to the one proposed by *Abbott and Branciard* and can therefore be considered as an optimal correction. For  $m = 0$  we map the post POVM measurement state onto the positive  $x$ -axis, mathematically

$$\epsilon_0(\rho) = \frac{1}{2}(\mathbb{1} + \sigma_x). \quad (3.26)$$

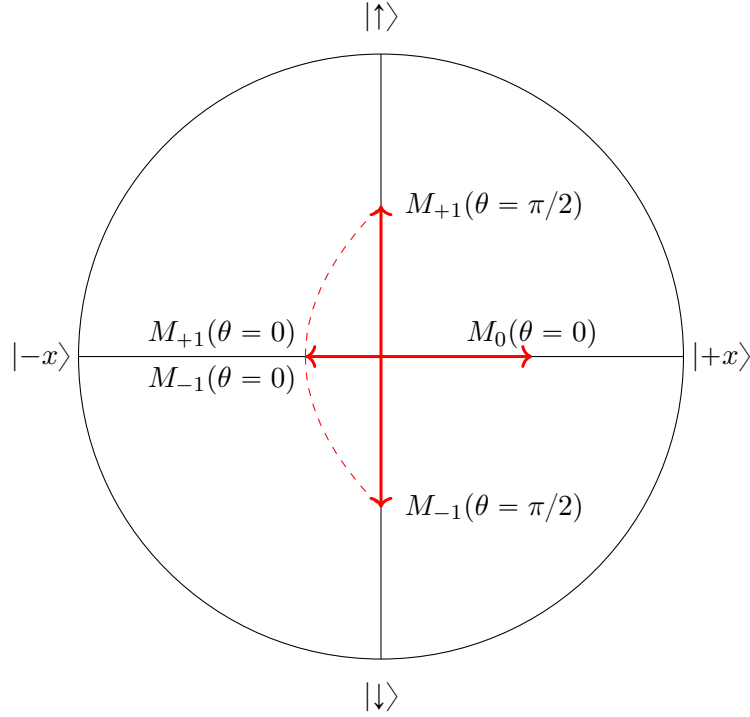


Figure 3.19: Illustration of the POVM under investigation in the experiment.  $M_0$  always points in  $|+x\rangle$ -direction and is shortened with increasing  $\theta$ , vanishing for  $\theta = \pi/2$ .  $M_{+1}$  and  $M_{-1}$  are equal for  $\theta = 0$  and point in  $|+x\rangle$ -direction. For increasing  $\theta$  the vectors begin to move towards  $|\uparrow\rangle$  respectively  $|\downarrow\rangle$  and are elongated. Their movement is along the dashed red lines.

Using the definition for entropic measurement uncertainties from *Buscemi et al.* (see Section 2.6.1, [4]), the noise and disturbance expected for the POVM was now calculated. We receive for the noise, preparing the state in  $\rho = \sigma_z$ ,

$$\mathbf{N}(\mathcal{M}^\theta, \sigma_z) = \frac{\cos \theta + h(\sin \theta)}{1 + \cos \theta}, \quad (3.27)$$

and for the disturbance with a prepared input state  $\rho = \sigma_x$ ,

$$\mathbf{D}_\epsilon(\mathcal{M}^\theta, \sigma_x) = \frac{h(\cos \theta)}{1 + \cos \theta}. \quad (3.28)$$

The function  $h(x)$  is given as

$$h(x) = -\frac{1+x}{2} \log_2 \left( \frac{1+x}{2} \right) - \frac{1-x}{2} \log_2 \left( \frac{1-x}{2} \right). \quad (3.29)$$

The realization of mapping procedures with the help of the spinrotator will not be discussed in detail, as the principle was already introduced in Section 3.2.3. Furthermore the manipulation of the spin with the help of spinrotator coils is a common procedure for many decades.

The application of the POVM weights results in measurement vectors which are not on the surface of the Bloch sphere but inside the volume of the sphere. They are shortened. It is therefore necessary to reduce the intensity of the neutron beam according to the applied POVM weights. In our case that was done with a tilt of the supermirror that is representing the POVM weights. This is a novelty in neutron polarimetry.

When a supermirror is tilted, the resulting intensity is of a Gaussian shape. That was already discussed in Section 3.2.1. An exemplary measurement is given in Fig. 3.8. We now defined the position  $-500\text{ k}$  as the setting where our POVM weight is  $p_m := 1$ . This is in the range of  $\sim 0.7$  of the maximum of the Gaussian shape. All POVM weights are applied *relative* to this point. They are therefore all situated at the flank of the Gaussian. High flip ratios for all weights are expectable in this region.

A detailed measurement with the relative intensities and the corresponding positions of the stepper motor axis is given in Fig. 3.20. Another advantage of using the Gaussian flank for the POVM weights is that the intensity is almost linear in that area.

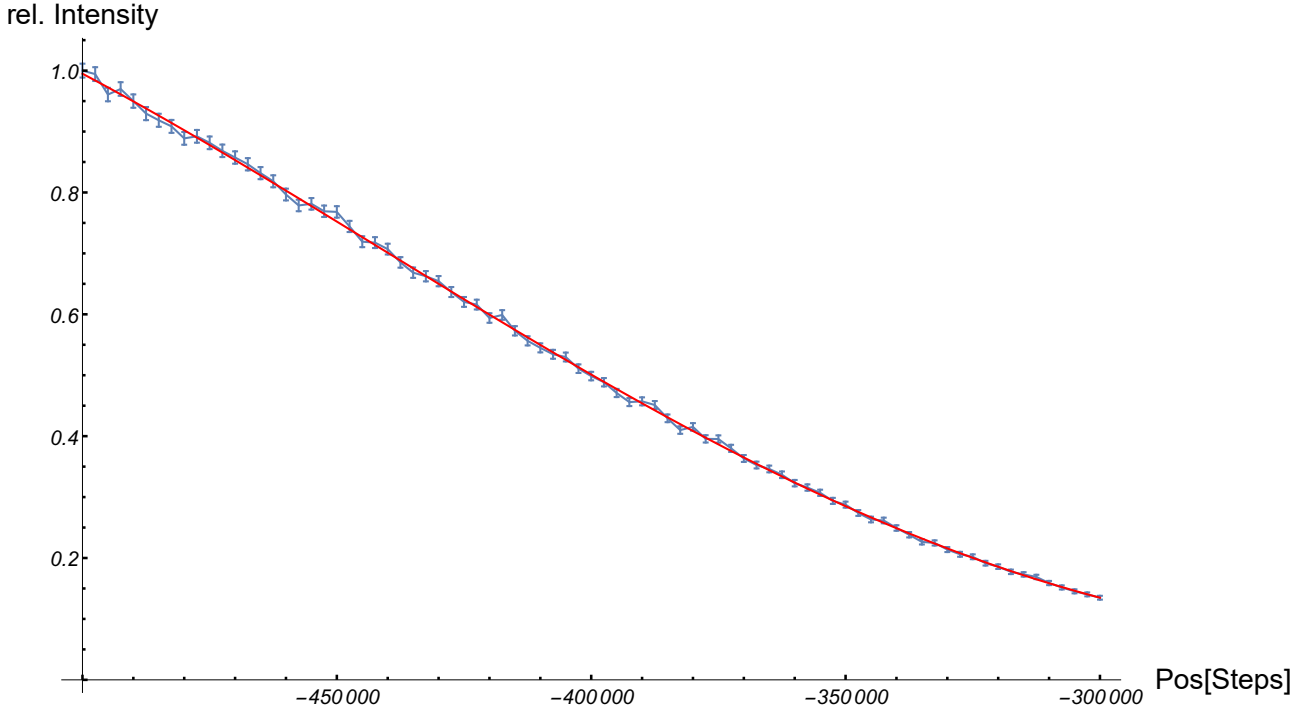


Figure 3.20: Detailed measurement of the relative POVM weights and the corresponding stepper motor positions. The intensities are normalized to the defined value of  $p_m := 1$  at axis position  $-500\text{ k}$ .

### 3.3.2 Setup for Noise Measurement

For the noise measurement we need to prepare input states with unitary probability. A supermirror polarizes the incident neutron beam and prepares it in state  $|\uparrow\rangle$ . The first spinrotator coil either performs a  $\pi$ -flip, resulting in the state  $|\downarrow\rangle$  or no current is applied at all, leaving the beam in state  $|\uparrow\rangle$ . The probability of a performed  $\pi$ -flip is  $p(\pi) = \frac{1}{2}$ . We receive a unitary probability for the states  $|\uparrow\rangle$  and  $|\downarrow\rangle$  for that behaviour of the DC1.

The measurement of the three POVM elements is done consecutively. The Bloch vectors corresponding to the POVM elements are mapped on the  $|\uparrow\rangle$ -axis by the second spinrotator coil DC2. A second supermirror acts as analyzer and also applies the respective POVM weights by being tilted as described in Section 3.3.1. This results in a measurement as defined in the POVM. A neutron detector is fixed right after the exit window of the analyzer with a mount. We measure the joint probabilities  $p(m, a)$  with  $|a\rangle$  being the input state and  $m$  the index of the POVM  $M_m^\theta$ . A scheme of the setup is given in Fig. 3.21. A picture of the setup for the noise measurement is given in Fig. 3.22.

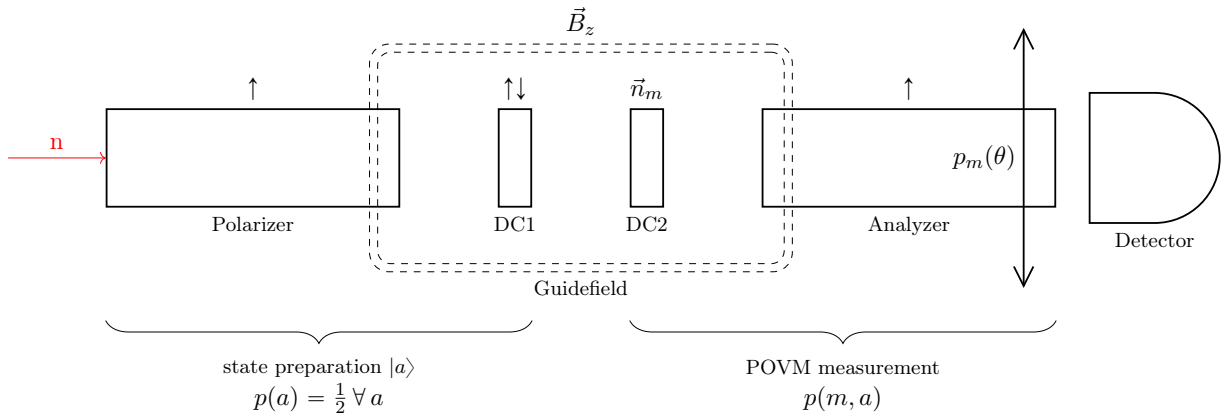


Figure 3.21: Scheme of the setup for the noise measurement. The first supermirror and spinrotator prepare the state with unitary probability for  $|\uparrow\rangle$  and  $|\downarrow\rangle$ . The neutron beam leaves the polarizer in the pure state  $|\uparrow\rangle$ . The spinrotator coil DC1 performs a  $\pi$ -flip, resulting in the state  $|\downarrow\rangle$  or no current is applied at all, leaving the beam in state  $|\uparrow\rangle$ . The probability of a performed  $\pi$ -flip is  $p(\pi) = \frac{1}{2}$ . The second coil DC2 maps the Bloch vector of the POVM to the  $|\uparrow\rangle$  axis and the second supermirror acts as analyzer and applies the POVM weights. We measure the joint probabilities  $p(m, a)$  with a neutron detector that is directly mounted on the analyzer and placed right after the exit window of the supermirror.

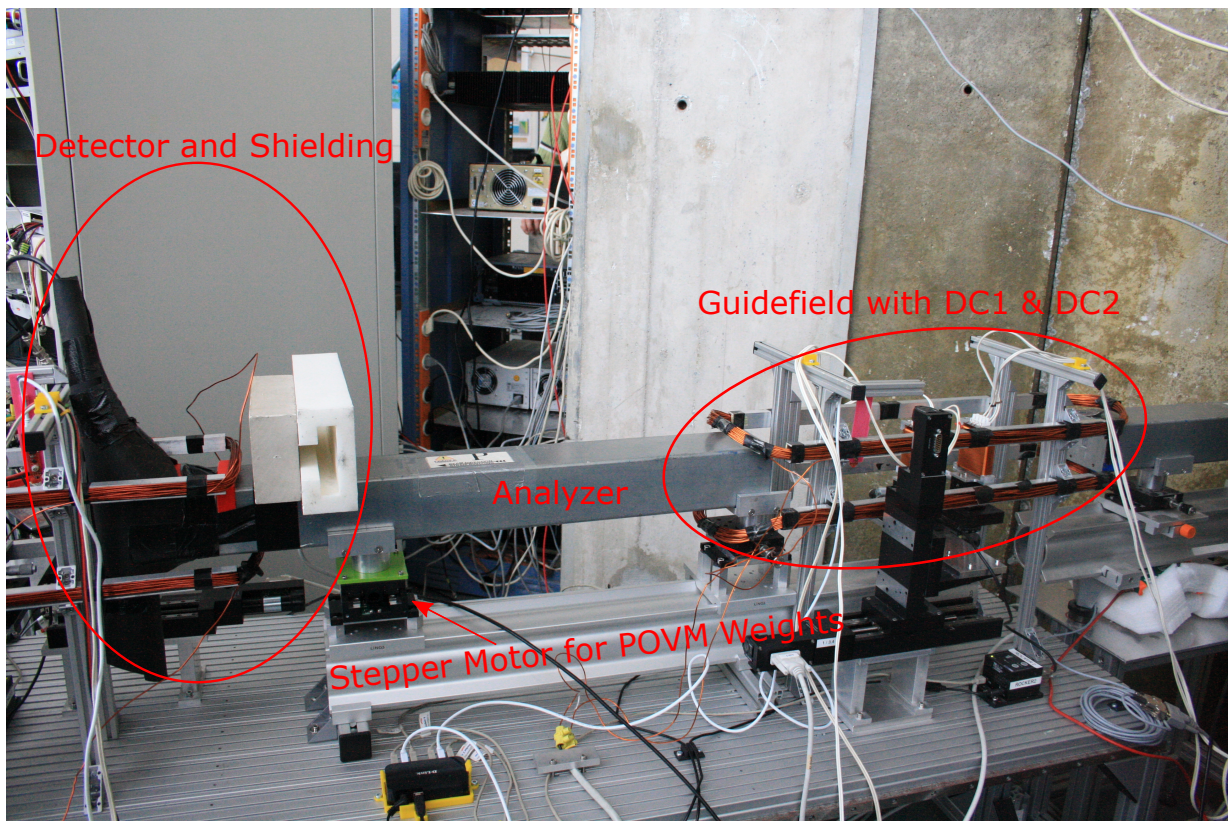


Figure 3.22: Picture of the setup for the noise measurement. The detector is directly mounted on the analyzer supermirror. The POVM weight tilt is performed by a stepper motor. The spin orientation is manipulated by a coil pair inside a guidefield, that is situated between polarizer and analyzer supermirror. The first coil prepares the desired input state  $|a\rangle$ , the second coil maps the POVM Bloch vectors on the  $|\uparrow\rangle$ -axis. The second supermirror acts as an analyzer of the spin state. It also applies the reduction of the intensity according to the POVM weight function  $p_m$  by being tilted.

### 3.3.3 Setup for Disturbance Measurement

For the disturbance measurement we need to prepare the input states  $|+\rangle$  and  $|-\rangle$ . The neutron beam leaves the first supermirror in the state  $|\uparrow\rangle$ . The first spinrotator coil DC1 performs either a  $+\frac{\pi}{2}$  flip, resulting in the state  $|+\rangle$  or a  $-\frac{\pi}{2}$  flip, resulting in  $|-\rangle$ . The probabilities for the flips are  $p(+\frac{\pi}{2}) = p(-\frac{\pi}{2}) = \frac{1}{2}$ . This results in a unitary probability for the input states  $|b\rangle$ .

The POVM measurement is performed exactly the same way as for the noise measurement. The measurement of the three POVM elements is done consecutively. The Bloch vectors corresponding to the POVM elements are mapped on the  $|\uparrow\rangle$ -axis by the second spinrotator coil DC2. A second supermirror acts as analyzer and also applies the respective POVM weights by being tilted as described in Section 3.3.1. This results in a measurement as defined in the POVM. After the first analyzer another spinrotator coil pair is situated. The first one, DC3, is performing the correction procedure for the currently measured POVM element. For the correction of  $M_0^\theta$  this means a  $+\frac{\pi}{2}$  rotation, for the correction of  $M_{\pm 1}^\theta$  this means a  $-\frac{\pi}{2}$  rotation. The second one, DC4, consecutively maps  $|+\rangle$  and  $|-\rangle$  onto the  $|\uparrow\rangle$ -axis. Together with the second analyzer supermirror it forms a projective measurement for  $\sigma_x$ . The detector is placed after the last supermirror. We receive the joint probabilities  $p(b', b)$ .

For a detailed look of the region with DC1 and DC2 please refer to the picture of the noise setup, given in Fig. 3.22. The detector that can be seen in the picture is removed for the disturbance measurement. The other elements stay exactly the same for the disturbance measurement. The region with DC3 and DC4 is presented in detail in Fig. 3.23. A picture as an overview of the full disturbance setup is given in Fig. 3.24. A scheme of the disturbance setup is presented in Fig. 3.25.

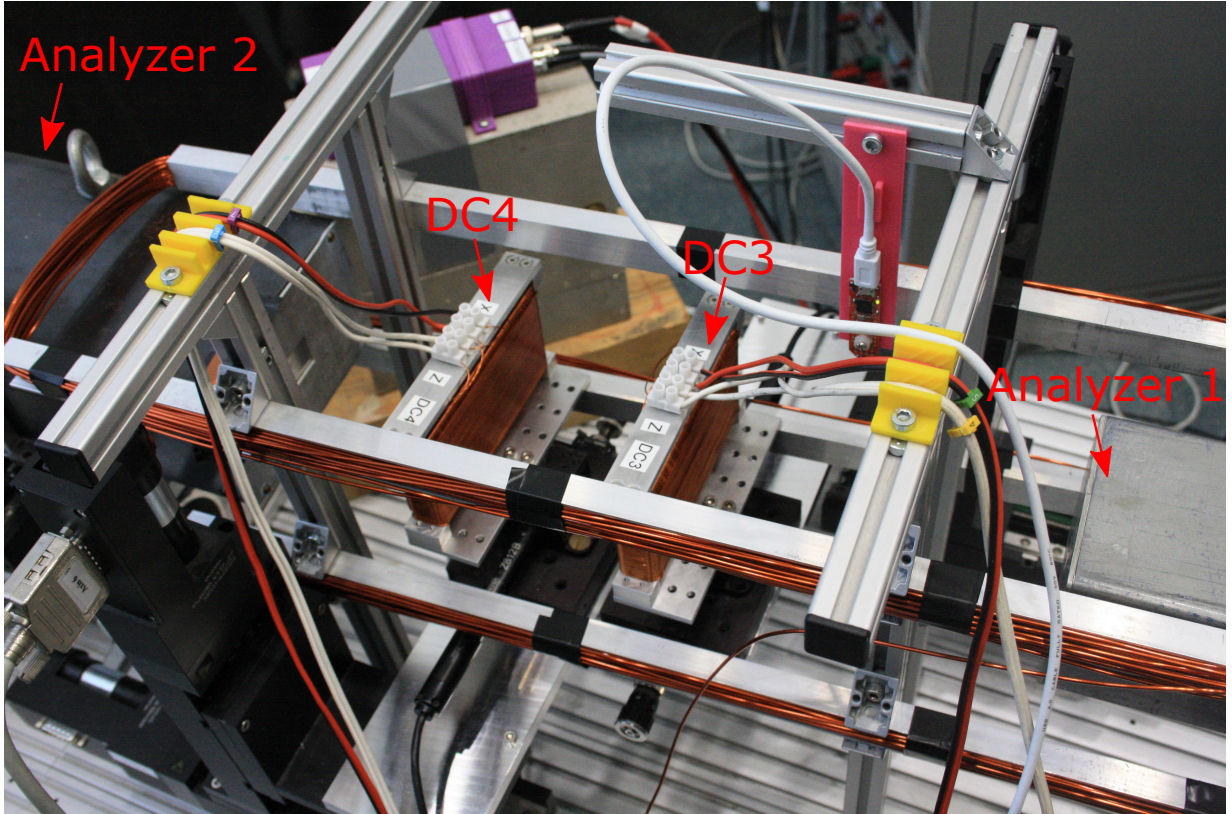


Figure 3.23: Detailed view of the region with DC3 and DC4. For the disturbance measurement the exit window of the first analyzer supermirror is not covered by the detector that was mounted directly on the supermirror for the noise measurement. The coil pair DC3 and DC4 is placed with a distance that represents a full Larmor precession. The DC3 performs the correction operation, the DC4 turns the projectors of the  $B$  measurement in  $|\uparrow\rangle$ -direction.

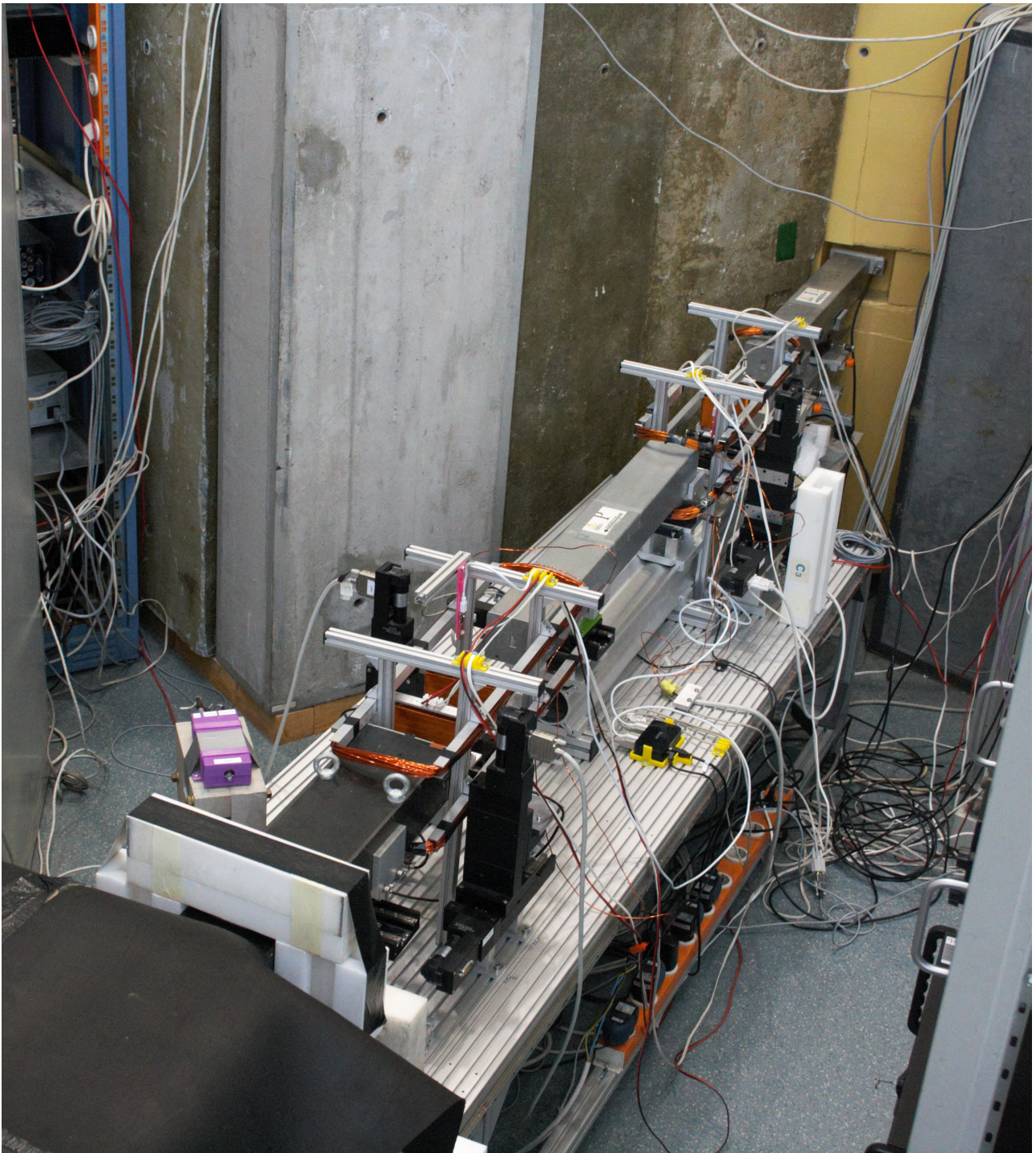


Figure 3.24: Overview of the full setup. All 3 supermirrors and both guidefield regions with each 2 spinrotator coils are visible. The last supermirror is a different type than the other ones. The detector for the disturbance measurement is in the background shielding built from bricks of borated paraffin (in white, coated with black boron carbide mats).

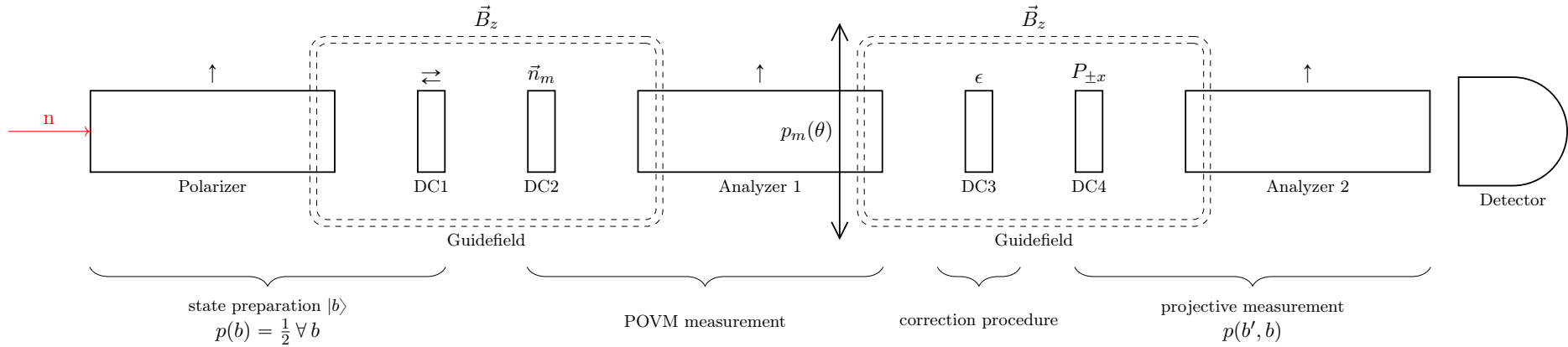


Figure 3.25: Scheme of the setup for the disturbance measurement. For the disturbance measurement we need to prepare the input states  $|+\rangle$  and  $|-\rangle$ . The neutron beam leaves the first supermirror in the state  $|\uparrow\rangle$ . The first spinrotator coil DC1 performs either a  $+\frac{\pi}{2}$  flip, resulting in the state  $|+\rangle$  or a  $-\frac{\pi}{2}$  flip, resulting in  $|-\rangle$ . The probabilities for the flips are  $p(+\frac{\pi}{2}) = p(-\frac{\pi}{2}) = \frac{1}{2}$ . This results in a unitary probability for the input states  $|b\rangle$ . The POVM measurement is performed exactly the same way as for the noise measurement. The measurement of the three POVM elements is done consecutively. The Bloch vectors corresponding to the POVM elements are mapped on the  $|\uparrow\rangle$ -axis by the second spinrotator coil DC2. A second supermirror acts as analyzer and also applies the respective POVM weights by being tilted as described in Section 3.3.1. This results in a measurement as defined in the POVM. After the first analyzer another spinrotator coil pair is situated. The first one, DC3, is performing the correction procedure for the currently measured POVM element. For the correction of  $M_0^\theta$  this means a  $+\frac{\pi}{2}$  rotation, for the correction of  $M_{\pm 1}^\theta$  this means a  $-\frac{\pi}{2}$  rotation. The second one, DC4, consecutively maps  $|+\rangle$  and  $|-\rangle$  onto the  $|\uparrow\rangle$ -axis. Together with the second analyzer supermirror it forms a projective measurement for  $\sigma_x$ . The detector is placed after the last supermirror. We receive the joint probabilities  $p(b', b)$ .



Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.  
The approved original version of this thesis is available in print at TU Wien Bibliothek.



Someone is sitting in the shade today  
because someone planted a tree a long time  
ago.

Warren Buffett

## 4.1 Noise Measurement

### 4.1.1 Raw Data and Data Treatment

For our noise measurements we measured 18 equidistant settings for the parameter  $\theta \in [0, \frac{\pi}{2}]$ , including the values  $\theta = 0$  and  $\theta = \frac{\pi}{2}$ . For every point on the noise plot all three POVM elements must be measured. This is done consecutively. For a full measurement set we therefore have 54 different measurement settings. The measurement time was 400 s per measurement setting. The raw data which we received directly from the experiment is given in Fig. 4.1.

The raw data then is decomposed with respect to the three possible outcome channels  $M_{-1}^\theta, M_0^\theta$  and  $M_{+1}^\theta$ . Two data correction procedures are applied. First, the neutron background is subtracted. This correction procedure can mathematically be written as

$$M_{m,bgCorr}^\theta = M_{m,raw}^\theta - bgps * t_{meas} , \quad (4.1)$$

where  $t_{meas} = 400$  s is the measurement time and,  $bgps$  is the average background per second for the experiment, which was measured as

$$bgps = \frac{2460}{1800 \text{ s}} \simeq 1.37 \text{ s}^{-1} . \quad (4.2)$$

A second data correction was performed incorporating the POVM weights  $p_m$  and the contrast of the system  $C$  as well as the parameters  $curr_m(\theta)$ ,  $flipcurr$  and  $a(\theta)$  which all three are associated with the spin rotation performed and are explained in detail below. The correction used was of the form

$$M_{m,corr}^\theta = M_{m,bgCorr}^\theta - I_{min}(C) \cdot p_m(\theta) \cdot \sin \left( \left[ \frac{curr_m(\theta)}{flipcurr} + a(\theta) \right] \frac{\pi}{2} \right) \quad (4.3)$$

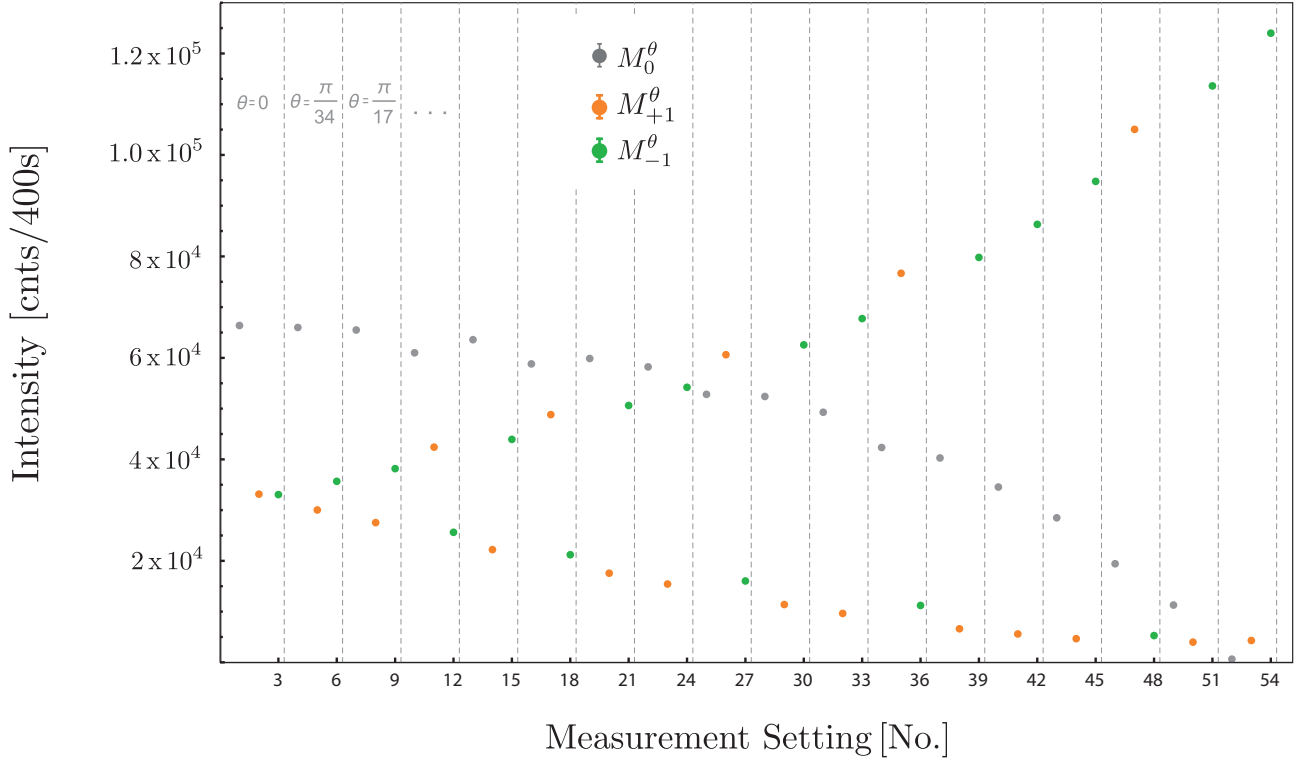


Figure 4.1: Raw measurement data directly received from the experiment. Three consecutive measurement settings form a point for the parameter  $\theta$ .

with

$$a(\theta) = \begin{cases} 0 & \text{for } |a\rangle = |\uparrow\rangle \\ 1 & \text{for } |a\rangle = |\downarrow\rangle \end{cases} . \quad (4.4)$$

The parameter  $I_{min}$  can be computed from the contrast  $C$  of the system and the knowledge of the maximum expected intensity  $I_{max}$  which can directly be obtained from the background corrected data set  $M_{m,bgCorr}^\theta$ . For the setting  $\theta = \frac{\pi}{2}$  we expect  $M_{-1}$  to have maximum intensity, if the input state  $a = |\downarrow\rangle$  was selected randomly. On the other hand, if  $a = |\uparrow\rangle$  was selected, we expect  $M_{+1}$  to have maximum intensity.

In our case, looking at the last two measurement settings in Fig. 4.1, we can see that  $a = |\downarrow\rangle$  was selected. With the definition of the contrast

$$C = \frac{I_{max} - I_{min}}{I_{max} + I_{min}} , \quad (4.5)$$

we can now calculate the expected count rate in the minimum. This represents the count rate we still expect to measure when the system is purely in state  $|\downarrow\rangle$ , as our system is not ideal. We receive

$$I_{min}(C) = I_{max} \frac{1 - C}{1 + C} . \quad (4.6)$$

The sine term in the correction scales the correction dependent on the total rotation applied to the spin by the spinflipper pair. The term  $a(\theta)$  corrects the state preparation coil DC1. In the case where  $|\downarrow\rangle$  is selected as input state a  $\pi$ -flip is performed and we therefore need to correct maximally. In the other case we do not need to correct at all. The term  $\frac{curr_m(\theta)}{flipcurr}$  corrects the DC2, which turns the POVM Bloch vectors in  $|\uparrow\rangle$ -direction. Here we scale with the momentarily applied current  $curr_m(\theta)$  relative to the current applied for a full  $\pi$ -flip.

An additional scaling factor is given via the POVM weights  $p_m(\theta)$ . As the measurement vectors of the POVM are shortended, the intensities are therefore also reduced by that factor and we need to apply the same scaling to the correction procedure. Typical contrasts for our system are in the range of  $C = 94\%$  to  $98\%$ . It can be obtained taking the average contrast of both spinflippers calculated from the minimum and maximum of the fitting curve of a DC sweep as shown exemplarily in Fig. 3.14.

After the decomposition and data correction procedure we receive three corrected channels representing the POVM elements for different  $\theta$ . They are presented in Fig. 4.2.

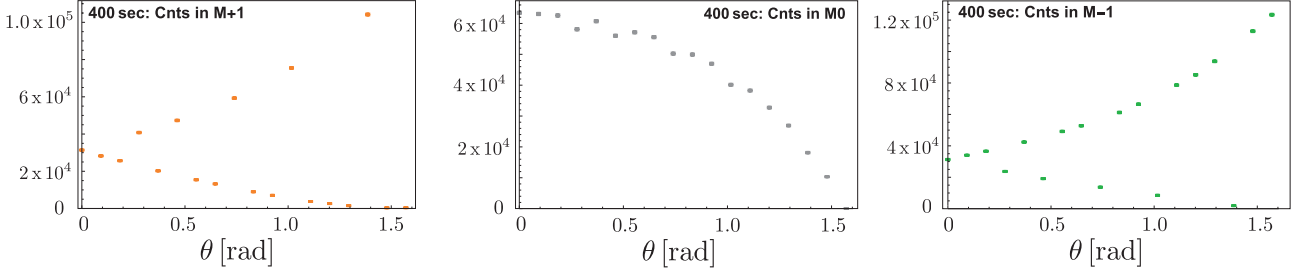


Figure 4.2: Corrected and detangled measurement data for all three POVM channels.

#### 4.1.2 Computation of Probabilities

In a next step we obtain the joint probabilities  $p(m, a)$  by normalizing the measured intensities. The probability for a full set of POVM channels is given with  $p = 1$ . We can now obtain the relative probabilities by normalizing over the sum of all counts, mathematically

$$p(m, a) = \frac{M_m^\theta}{M_{-1}^\theta + M_0^\theta + M_{+1}^\theta} . \quad (4.7)$$

There are two types of errors to be considered. First, the purely statistical error, being  $\sqrt{N}$  for  $N$  counts, as we deal with a *Poissonian distribution*. The second type of error is a systematic error originating in the imperfections of the spinflippers. It is given via an average angle deviation for a manipulation of the spin. For our system it was measured as  $\Delta_{sys} \simeq 0.7^\circ$ . The value is obtained from the deviation of the ideal phase of the sinusoidal fit for our spinflipper system, as exemplarily shown in Fig. 3.14. A comparison of the joint probabilities with and without systematic error is given in Fig. 4.3.

When we compare the theoretical predictions for the joint probabilities with the values received by measurement we can see that for the channels  $M_{-1}^\theta$  and  $M_{+1}^\theta$  the initial state  $|a\rangle$  can be inferred. The only value for  $\theta$  where a deduction is not possible is  $\theta = 0$ , where we expect maximum noise. With increasing values of  $\theta$  the distinction of the two possible input states becomes more significant, the noise declines.

For the computation of the conditional probabilities we need full sets of  $p(a = -1, m)$ ,  $p(a = +1, m)$ , meaning, that the joint probability for both input states is necessary for every  $M_m^\theta$ . In our measurements we only obtain a joint probability for one of the input states. It is coincidence which one is selected by the state preparation coil. We use the theoretical prediction for the sum  $p(a = \pm 1, m) = p(a = -1, m) + p(a = +1, m)$  of the probabilities to calculate the missing value for every  $m = \pm 1$  and  $\theta$  and to receive all permutations of  $p(a, m = \pm 1)$ . The handling of  $M_0^\theta$  is somehow different and described below.

For the channel  $M_{+1}^\theta$  the calculation of  $p(a = -1, m = +1)$  and  $p(a = +1, m = +1)$  is shown in Fig. 4.4.

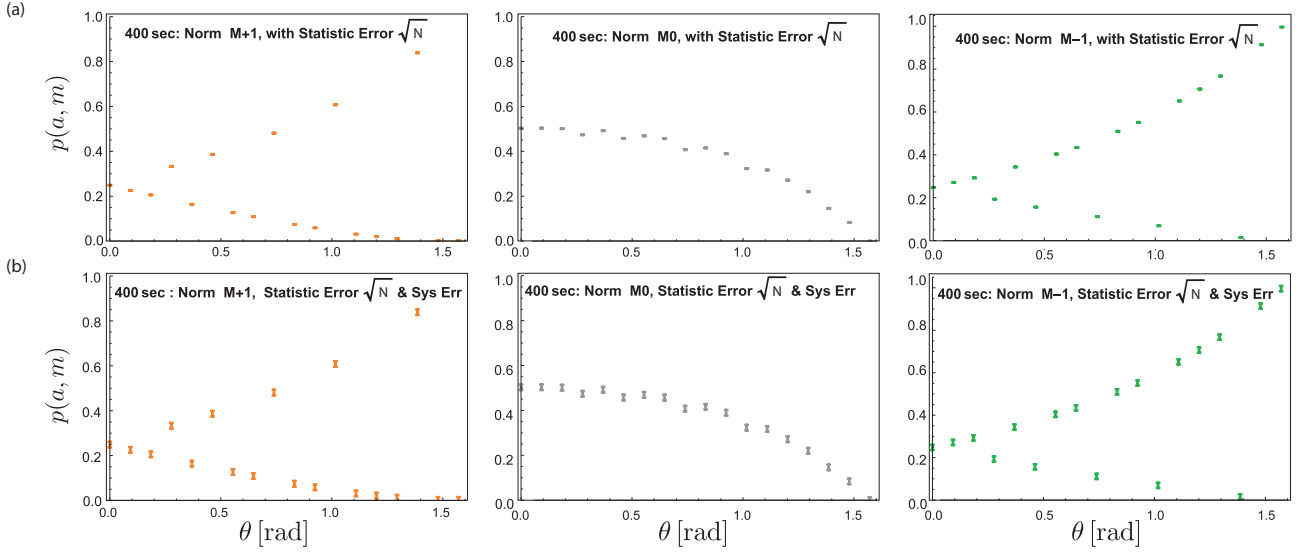


Figure 4.3: Joint probabilities acquired by normalization of the intensities.

(a) Including a statistical error  $\sqrt{N}$ .

(b) Additional consideration of a systematic error  $\Delta_{sys} \simeq 0.7^\circ$ .

The conditional probabilities  $p(a = +1|m = +1)$  and  $p(a = -1|m = +1)$  can be acquired using

$$p(a|m) = \frac{p(a, m)}{p(m)} = \frac{p(a, m)}{\sum_a p(a, m)}. \quad (4.8)$$

They are depicted together with the theoretical predictions in Fig. 4.5.

For  $M_0^\theta$  it is impossible to infer which input state was sent to the measurement apparatus. The theoretical predictions are exactly the same for both possible inputs  $|a\rangle$ . The *identical* data set is taken for the joint probabilities. The measured values are used for both input states. The principle is plotted in Fig. 4.6. The conditional probabilities are computed with Eq. (4.8) and give us exact equal probability for both input states, as the joint probabilities are equal. The results are shown in Fig. 4.7.

Finally we will have a look at the last POVM channel  $M_{-1}^\theta$ . The predicted theoretical curves for  $a = +1$  and  $a = -1$  are swapped with those for  $m = +1$ . The computation of the joint probabilities for both input states follows the same principle as the one describes above for  $M_{+1}^\theta$ . The results are shown in Fig. 4.8, the computed conditional probabilities in Fig. 4.9.

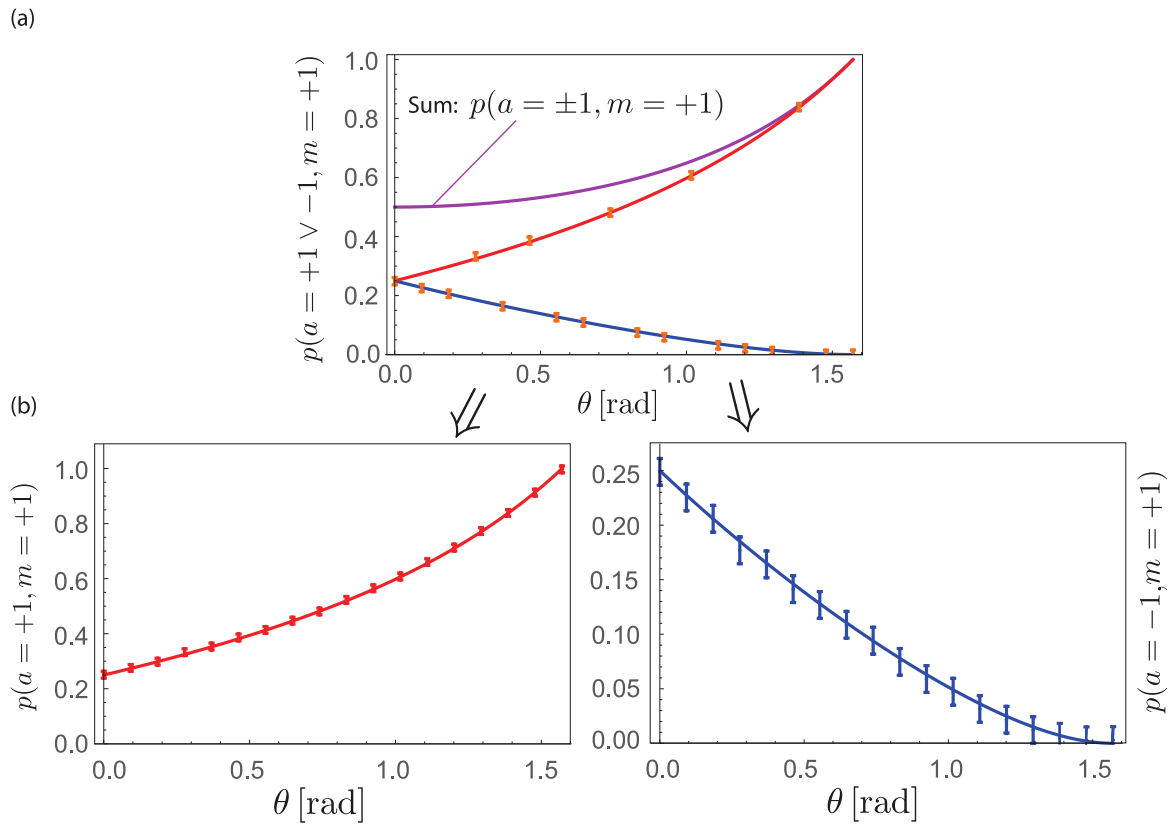


Figure 4.4: Joint probabilities for  $M_{+1}^\theta$ .

(a) Joint probability received by the measurement data plotted over the theoretical curves for  $M_{+1}^\theta$ . The input state is still "indefinite".

(b) Joint probabilities  $p(a = +1, m = +1)$  (red) and  $p(a = -1, m = +1)$  (blue) consisting of a combination of measured and computed values.

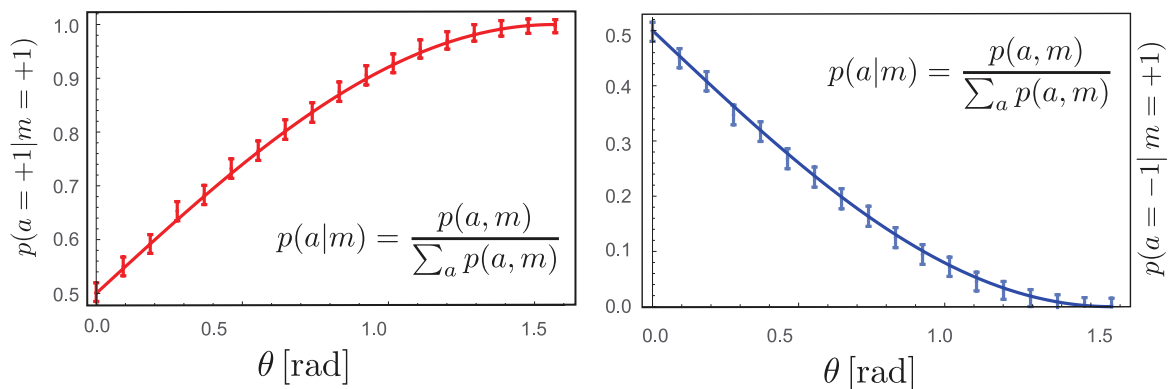


Figure 4.5: Conditional probabilities for the  $M_{+1}^\theta$  channel. On the left it is given for the input  $|\uparrow\rangle$  (red), on the right for  $|\downarrow\rangle$  (blue).

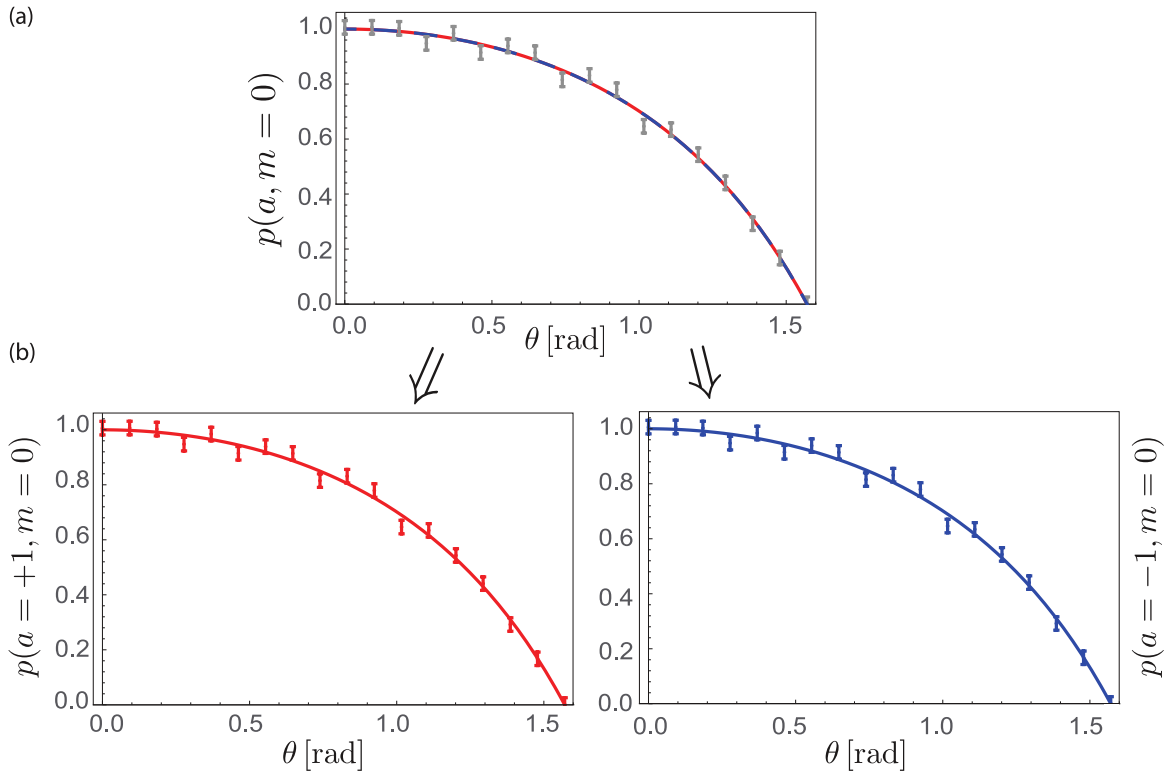


Figure 4.6: Joint probabilities for  $M_0^\theta$ .  
 (a) Joint probability received by the measurement data plotted over the theoretical curves for  $M_0^\theta$ . The input state is "indefinite".  
 (b) Joint probabilities  $p(a = +1, m = 0)$  (red) and  $p(a = -1, m = 0)$  (blue) which are equal for all  $\theta$ .

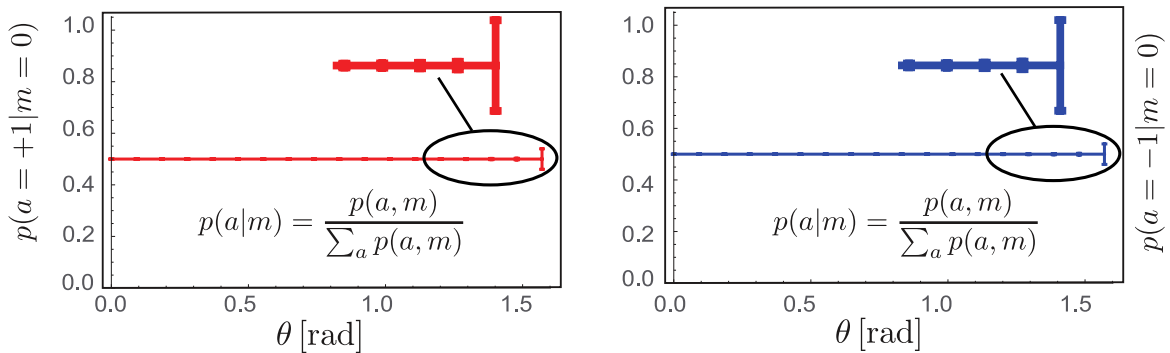


Figure 4.7: Conditional probabilities for the  $M_0^\theta$  channel. On the left it is given for the input  $|\uparrow\rangle$  (red), on the right for  $|\downarrow\rangle$  (blue). They are  $p = \frac{1}{2}$  for all  $\theta$ . The values have a very small error, please note the zoomed area in the plots.

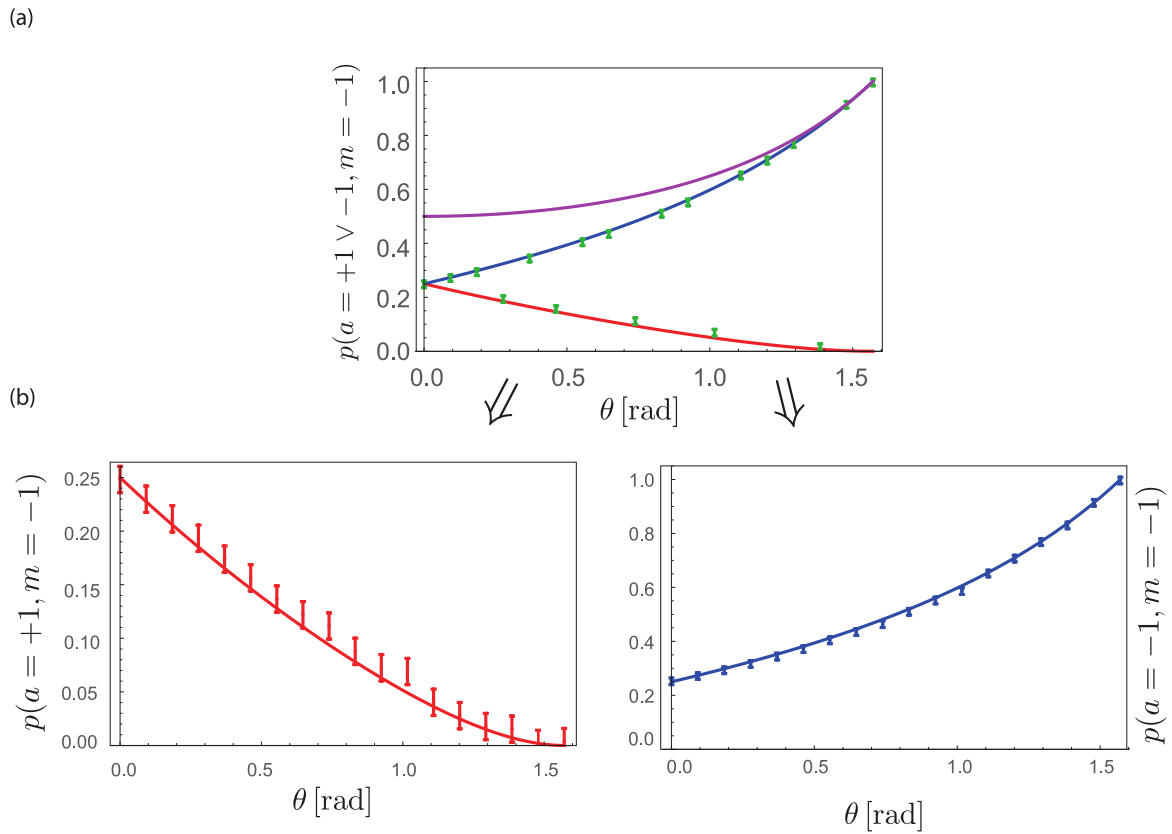


Figure 4.8: Joint probabilities for  $M_{-1}^\theta$ .

(a) Joint probability received by the measurement data plotted over the theoretical curves for  $M_{-1}^\theta$ . The input state is "indefinite".

(b) Joint probabilities  $p(a = +1, m = -1)$  (red) and  $p(a = -1, m = -1)$  (blue) consisting of a combination of measured and computed values.

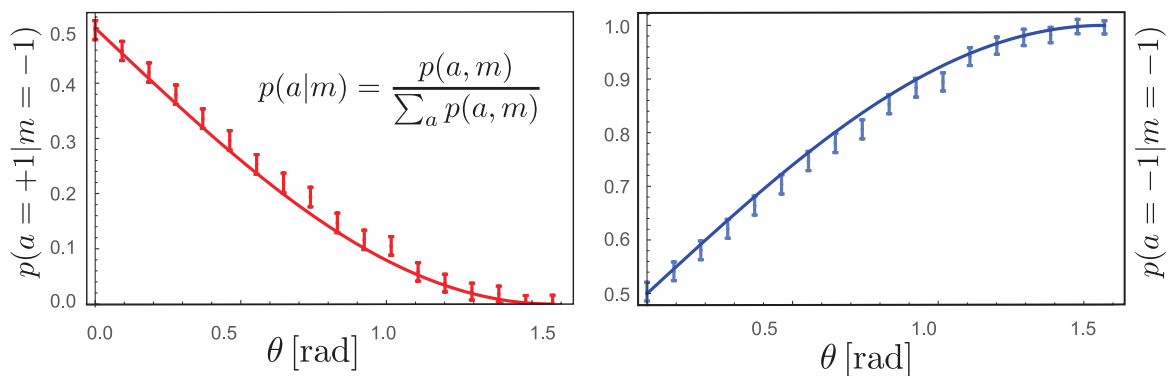


Figure 4.9: Conditional probabilities for the  $M_{-1}^\theta$  channel. On the left it is given for the input  $|\uparrow\rangle$  (red), on the right for  $|\downarrow\rangle$  (blue).

### 4.1.3 Final Noise Plot

With the six conditioned probabilities  $p(a|m)$  we can calculate the noise  $N(\mathcal{M}^\theta, \sigma_z)$  via its definition

$$N(\mathcal{M}^\theta, \sigma_z) = H(\mathbb{A}|\mathbb{M}) = - \sum_m p(m) \sum_a p(a|m) \log_2 p(a|m) . \quad (4.9)$$

The marginal probabilities for the POVM elements are given as

$$p(m) = \frac{1}{d} \text{Tr}[M_m] = \frac{1}{2} \text{Tr}[M_m] . \quad (4.10)$$

The final results are presented in Fig. 4.10. The theoretical predictions for the noise  $N(\mathcal{M}^\theta, \sigma_z)$  are reproduced over the entire range of the POVM parameter  $\theta$ .

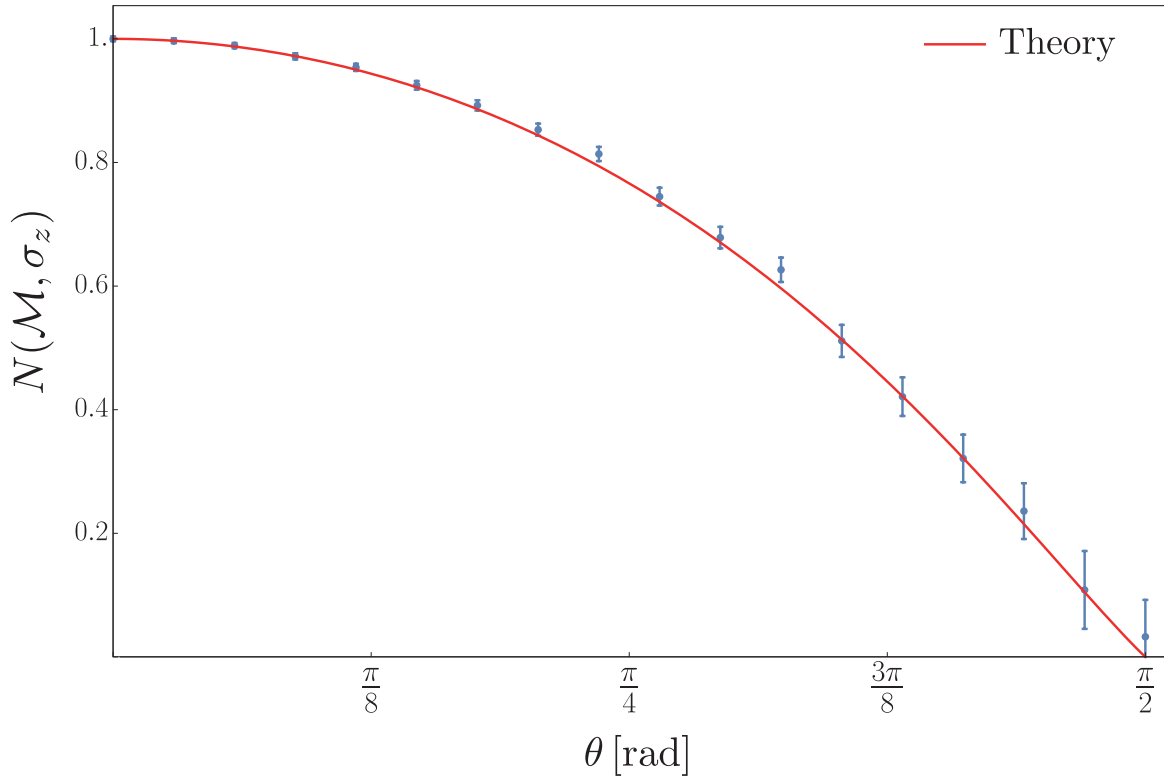


Figure 4.10: Final plot of the noise  $N(\mathcal{M}^\theta, \sigma_z)$  of the POVM  $M_m^\theta$  with input states as eigenstates of  $\sigma_z$  as a function of the POVM parameter  $\theta$ . The red line is the theoretical prediction.

## 4.2 Disturbance Measurement

### 4.2.1 Raw Data and Data Treatment

For our disturbance measurement we measured the same 18 equidistant POVM parameter settings for  $\theta \in [0, \frac{\pi}{2}]$ . The eigenstates  $|b\rangle$  are now uniformly distributed eigenstates of  $\sigma_x$ . The POVM measurement causes a disturbance on a subsequent projective measurement. To quantify the disturbance, for every setting of  $\theta$  all permutations of the POVM measurement channels  $\{M_{-1}^\theta, M_0^\theta, M_{+1}^\theta\}$  and the projective measurement channels  $\{|-x\rangle\langle -x|, |+x\rangle\langle +x|\}$  must be measured. This leaves us with six measurement settings per  $\theta$ . The raw data acquired through measurement is depicted in Fig. 4.11.



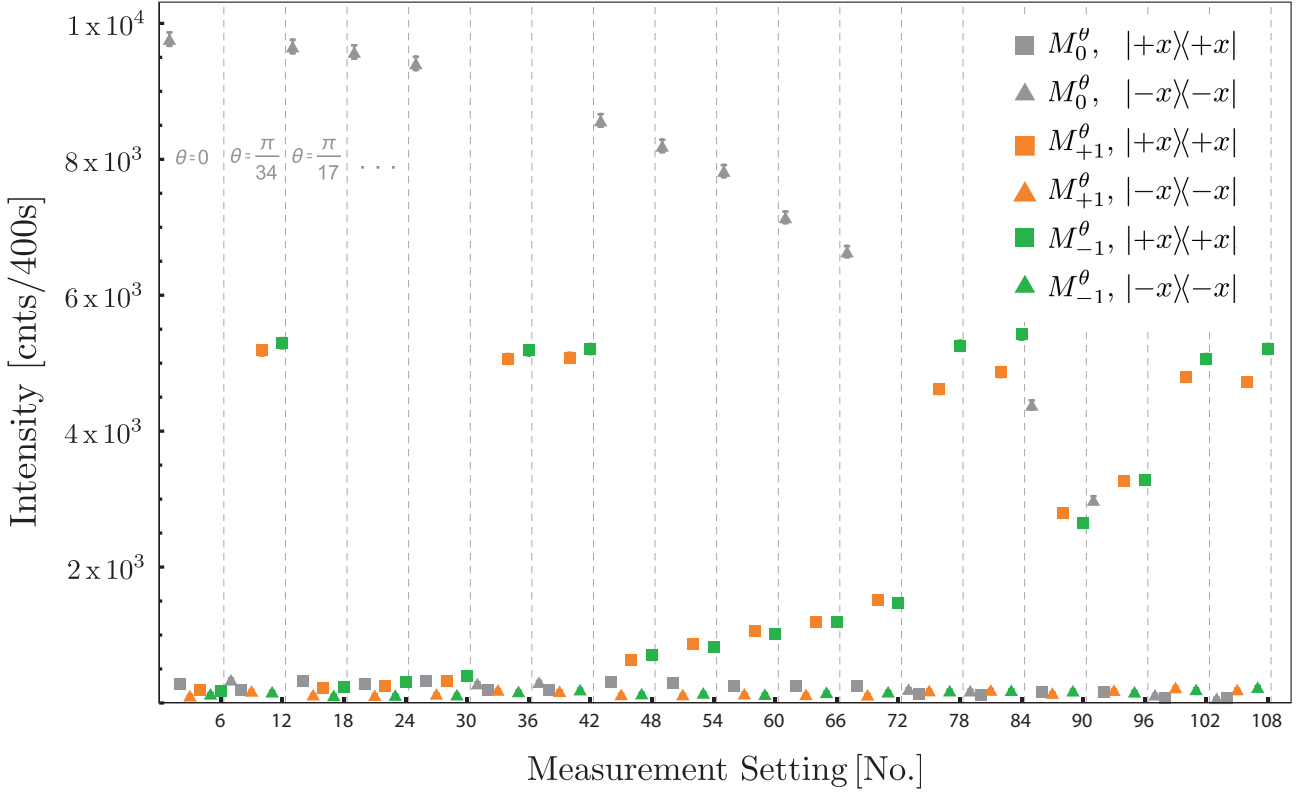


Figure 4.11: Raw data of the *disturbance* measurement  $D_\epsilon(M^\theta, \sigma_x)$  of the three output POVM  $M^\theta$  and projective  $B = \sigma_x$  measurement for 400 seconds.

The raw data is then decomposed with respect to the possible outcome channels  $I_{m,b'}^\theta$ , where  $m$  is the index of the POVM channel and  $b'$  the index of the projective measurement. We use  $b' = +1$  for the measurement  $|+x\rangle\langle+x|$  respectively  $b' = -1$  for  $|-x\rangle\langle-x|$ . The data is corrected similar to the noise measurement. First, the background of the measurement is subtracted, mathematically written as

$$I_{m,b',bgCorr}^\theta = I_{m,b',raw}^\theta - bgps * t_{meas} , \quad (4.11)$$

where  $t_{meas}$  is the measurement time and,  $bgps$  is the average background per second for the experiment, which was measured as

$$bgps = \frac{317}{1800 \text{ s}} \simeq 0.176 \text{ s}^{-1} . \quad (4.12)$$

The background is a lot lower than the one measured for the noise measurement (see Eq. (4.2)). This is due to a more sophisticated background shielding for the detector used for the disturbance measurements. For the noise measurements the detector was directly mounted on the analyzer supermirror. The space usable to add shielding for the noise detector was very limited. It was compensated, as the count rates for the noise measurements were higher than the ones for the disturbance measurements. The measurements presented are measured with a measurement time of  $t_{meas} = 400 \text{ s}$ .

The second correction is an extension of the procedure for the noise and can mathematically be written as

$$I_{m,b',corr}^{\theta} = I_{m,b',bgCorr}^{\theta} - I_{min}(C) \cdot p_m(\theta) \cdot \sin \left( \left[ \frac{curr_m(\theta)}{flipcurr} - (-1)^m \frac{b(\theta)}{2} + c_{b'} \right] \frac{\pi}{2} \right) \quad (4.13)$$

with

$$b(\theta) = \begin{cases} -1 & \text{for } |b\rangle = |-x\rangle \\ +1 & \text{for } |b\rangle = |+x\rangle \end{cases}, \quad (4.14)$$

and

$$c_{b'} = \begin{cases} 0 & \text{for } I_{0,+1}, I_{\pm 1,-1} \\ 1 & \text{for } I_{0,-1}, I_{\pm 1,+1} \end{cases}. \quad (4.15)$$

The usage of the minimum count rate  $I_{min}(C)$  and the POVM weights  $p_m(\theta)$  is the same as for the correction of the noise measurement. The rest of the correction is also just an extension of the correction already introduced. Every term in the sine function refers to a certain coil or coil segment.

The term

$$\frac{curr_m(\theta)}{flipcurr}, \quad (4.16)$$

is referring to the DC2 in the setup. The correctional factor is scaled with the momentarily applied current representing the mapping of the POVM element Bloch vector onto the  $|\uparrow\rangle$ -axis. It is the only term in the sinus unchanged compared to the noise correction.

The term

$$(-1)^m \frac{b(\theta)}{2}, \quad (4.17)$$

is correcting the state preparation operation of the DC1. The absolute value of the term is always fixed to  $\frac{1}{2}$ . This represents the rotation of  $\pm \frac{\pi}{2}$  which maps the initial state  $|\uparrow\rangle$  on the  $|-x\rangle$  respectively  $|+x\rangle$  axis. The  $(-1)^m$  subterm guarantees that the correction direction is correct relative to the spin rotation performed by the DC2.

The variable  $c_{b'}$  is referring to the DC3 and DC4 coil pair. The coil pair combination either performs a full  $\pi$ -flip, leading to a maximal correction represented by  $c_{b'} = 1$  or two consecutive  $\pm \frac{\pi}{2}$  rotations with alternating signs, leading to no correction, represented by  $c_{b'} = 0$ .

## 4.2.2 Computation of Probabilities

In a next step we obtain the joint probabilities  $p(b, m, b')$  by normalizing the measured intensities. The probability for a full set of measurement channels is given with  $p = 1$ . We can now obtain the relative probabilities by normalizing over the sum of all counts, given by

$$p(b, m, b') = \frac{(M_m^{\theta}, |b'\rangle\langle b'|)}{\sum_{m,b} (M_m^{\theta}, |b'\rangle\langle b'|)}. \quad (4.18)$$

Analogous to the noise measurement we only measure the probabilities  $p(b, m, b')$  for one input state  $|b\rangle$  and compute the probabilities for the other input state which was not measured. The normalized data is presented in Fig. 4.12.

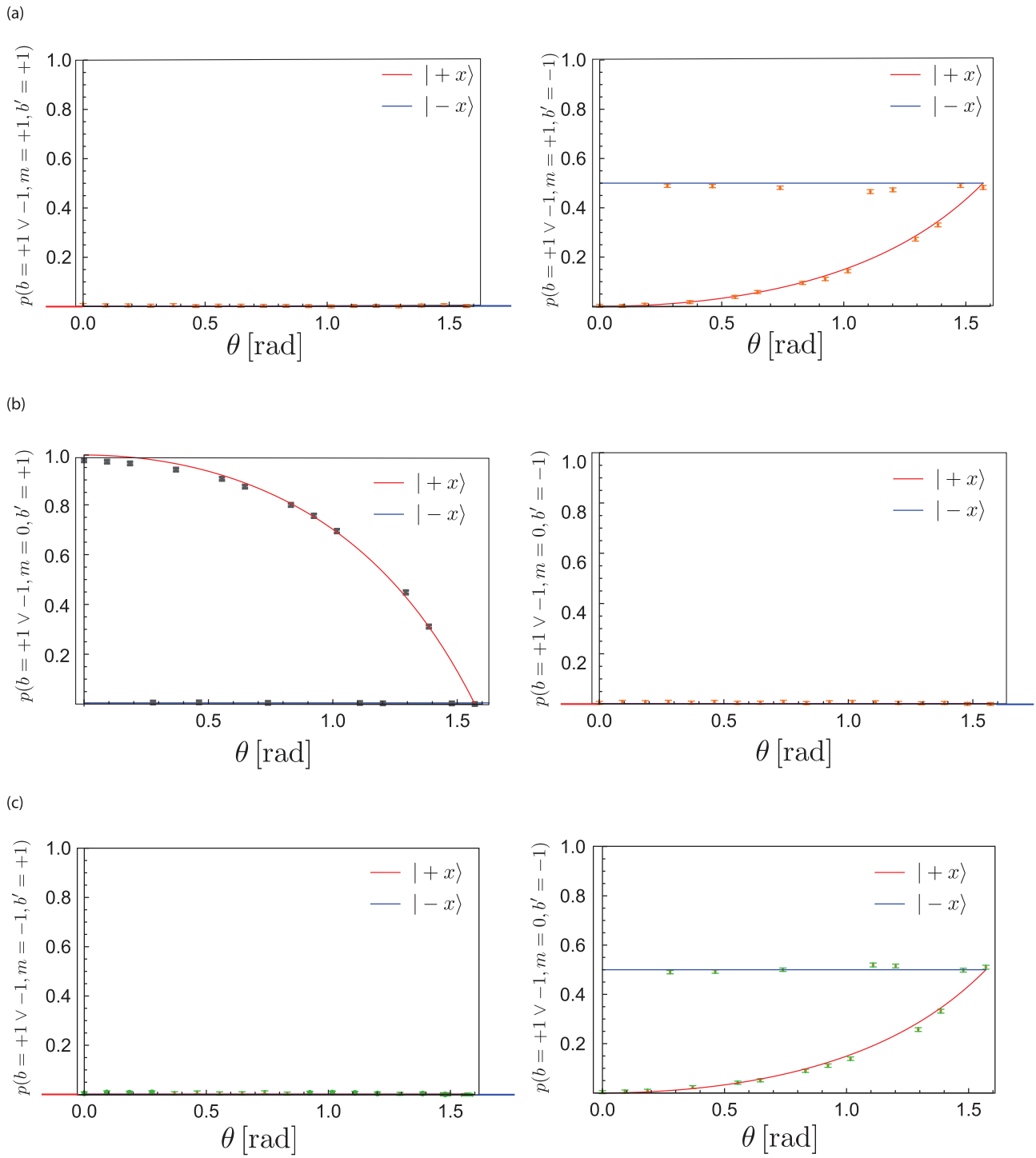


Figure 4.12: Normalized data with "indefinite" input state  $|b\rangle = |+x\rangle \vee |-x\rangle$  split up in the two output channels of the subsequent projective measurement for  $\sigma_x$ .

- (a) Measurement channel  $M_{+1}^\theta$ .
- (b) Measurement channel  $M_0^\theta$ .
- (c) Measurement channel  $M_{-1}^\theta$ .

The theoretical prediction for the joint probability is given with

$$\begin{aligned}
p(b, b') &= \frac{1}{2} \text{Tr} \left( \sum_{m=-1}^1 \epsilon_m (\mathcal{M}_m(|b\rangle\langle b|)) |b'\rangle\langle b'| \right) \\
&= \frac{1}{2} \text{Tr} \left( \sum_{m=-1}^1 p_m \epsilon_m (P(\vec{n}_m) |b\rangle\langle b| P(\vec{n}_m) |b'\rangle\langle b'|) \right) \\
&= \sum_{m=-1}^1 p_m \left( \frac{1 + b \vec{e}_x \cdot \vec{n}_m}{2} \right) \langle b' | \epsilon_m (P(\vec{n}_m)) | b' \rangle .
\end{aligned} \tag{4.19}$$

Using the *optimal correction*

$$\epsilon_m (P(\vec{n}_m)) = \frac{1}{2} (\mathbb{1} + (-1)^m \sigma_x) , \tag{4.20}$$

we finally get

$$p(b, b') = \sum_{m=-1}^1 p_m \frac{1}{2} \left( \frac{1 + b \vec{e}_x \cdot \vec{n}_m}{2} \right) (\mathbb{1} + (-1)^m b') = \frac{1 - b' + (1 + b' + 2bb') \cos \theta}{4(1 + \cos \theta)} . \tag{4.21}$$

The *marginal probability*  $p(b')$  used is acquired by summation

$$p(b') = \sum_b p(b, b') = \frac{1 - b' + \cos \theta + b' \cos \theta}{2 + 2 \cos \theta} . \tag{4.22}$$

From the normalized joint probabilities we can now calculate the conditional probabilities  $p(b, b')$  by summation, denoted as

$$p(b, b') = \sum_m p(b, m, b') , \tag{4.23}$$

and further the marginal probabilities with

$$p(b') = \sum_b p(b, b') . \tag{4.24}$$

The calculated joint probabilities  $p(b, b')$  plotted together with the theoretical predictions are shown in Fig. 4.13. The marginal probabilities  $p(b')$  and their theoretical predictions are plotted in Fig. 4.14.

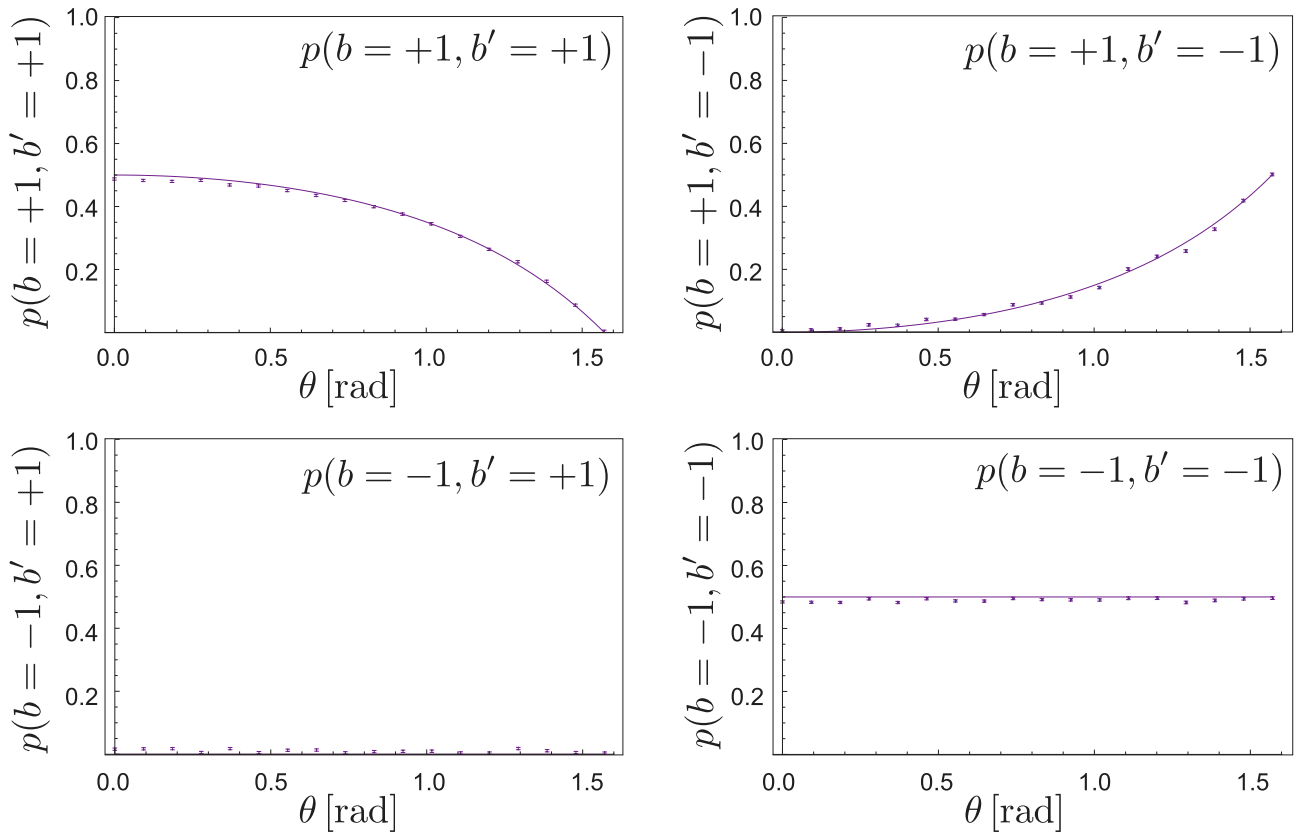


Figure 4.13: The joint probabilities  $p(b, b')$  for all four permutations of  $b$  and  $b'$  plotted together with the theoretical predictions.

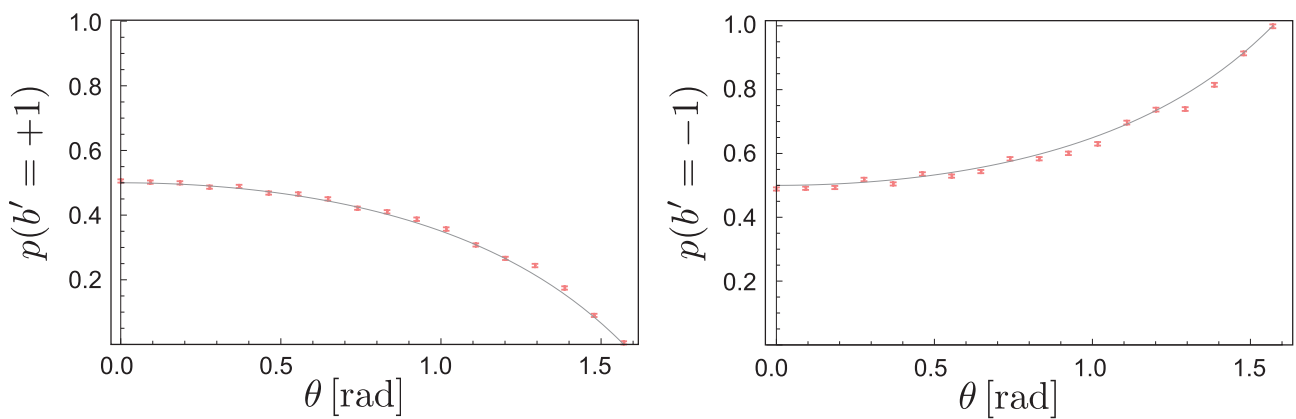


Figure 4.14: The marginal probabilities  $p(b')$  together with the theoretical predictions. On the left side we see  $p(b' = +1)$ , on the right  $p(b' = -1)$ .

## 4.2.3 Final Disturbance Plot

Finally the disturbance operator can be calculated from the four joint probabilities  $p(b, b')$  and the two marginal probabilities  $p(b')$ . The operator is given as

$$D_\epsilon(\mathcal{M}^\theta, \sigma_x) = H(\mathbb{B}|\mathbb{B}') = - \sum_{b, b'} p(b, b') \log_2 \frac{p(b, b')}{p(b')} . \quad (4.25)$$

The final disturbance plot is given in Fig. 4.15. We were able to reproduce the trend of the theoretical predictions over almost the full range of the POVM parameter  $\theta$ . The values observed are slightly higher as predicted by theory. Reasons for that deviation from theory are discussed in Section 4.4.

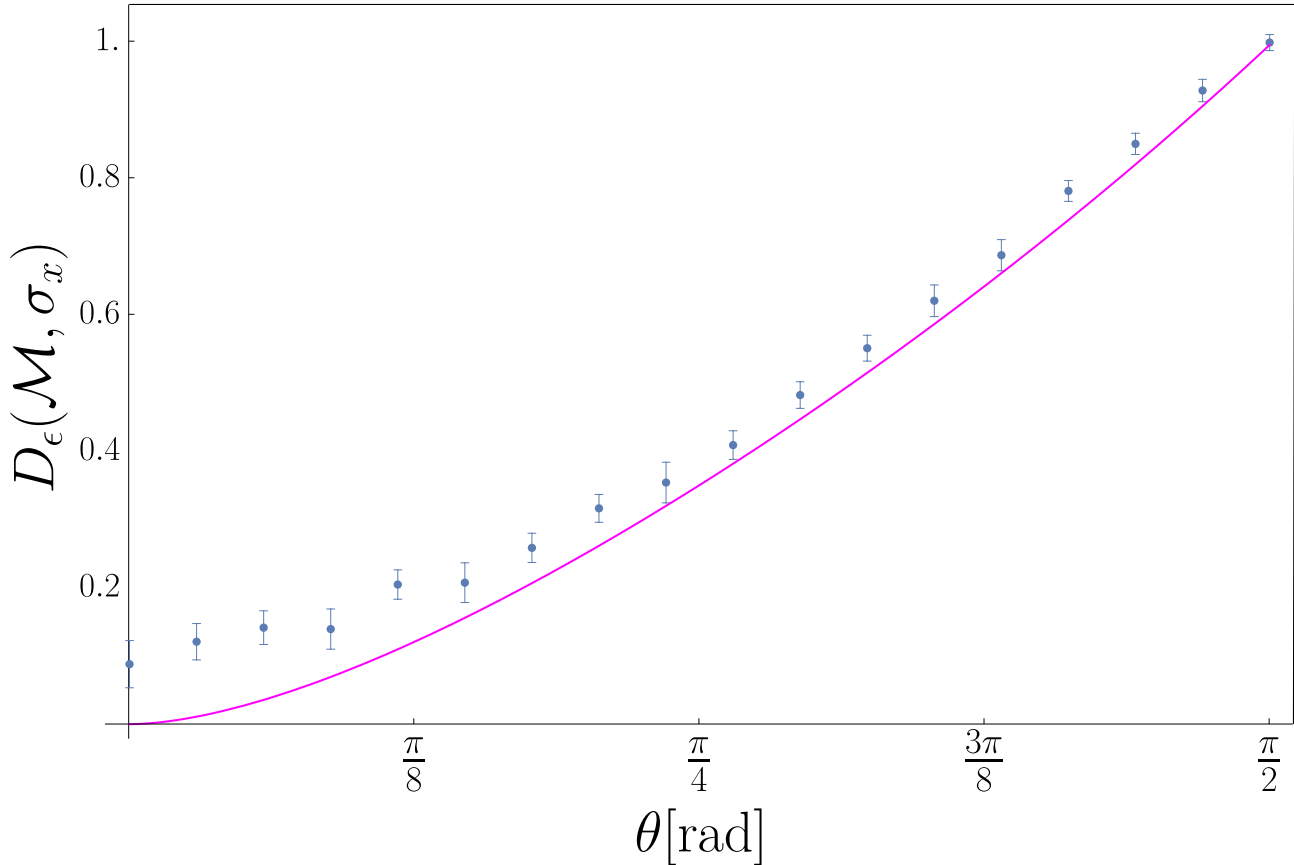


Figure 4.15: The final disturbance plot. The first ten points (for low  $\theta$ ) were measured with  $t_{meas} = 800$  s to reduce the statistical error. The trend of the theoretical prediction is reproduced over almost the entire range of the POVM parameter  $\theta$ . The disturbance values measured are slightly higher as the predictions.

### 4.3 Final Noise Disturbance Plot

The final noise disturbance plot shows equidistant pairs of the noise (see Fig. 4.10) and the disturbance (see Fig. 4.15) with matching POVM parameter  $\theta$ . As the main interest of the experiment was to show that the projective measurements by *Sulyok et al.* [6] can be outperformed with the POVM under investigation, both the theoretical curve as well as the measured points for the projective measurements are shown together with the current measurements.

We have shown that the POVM measurements yield in a tighter noise disturbance uncertainty relation than the projective measurements. The projective measurements were outperformed over almost the entire range of  $\theta$ . The results also showing the previous projective measurements are depicted in Fig. 4.16.

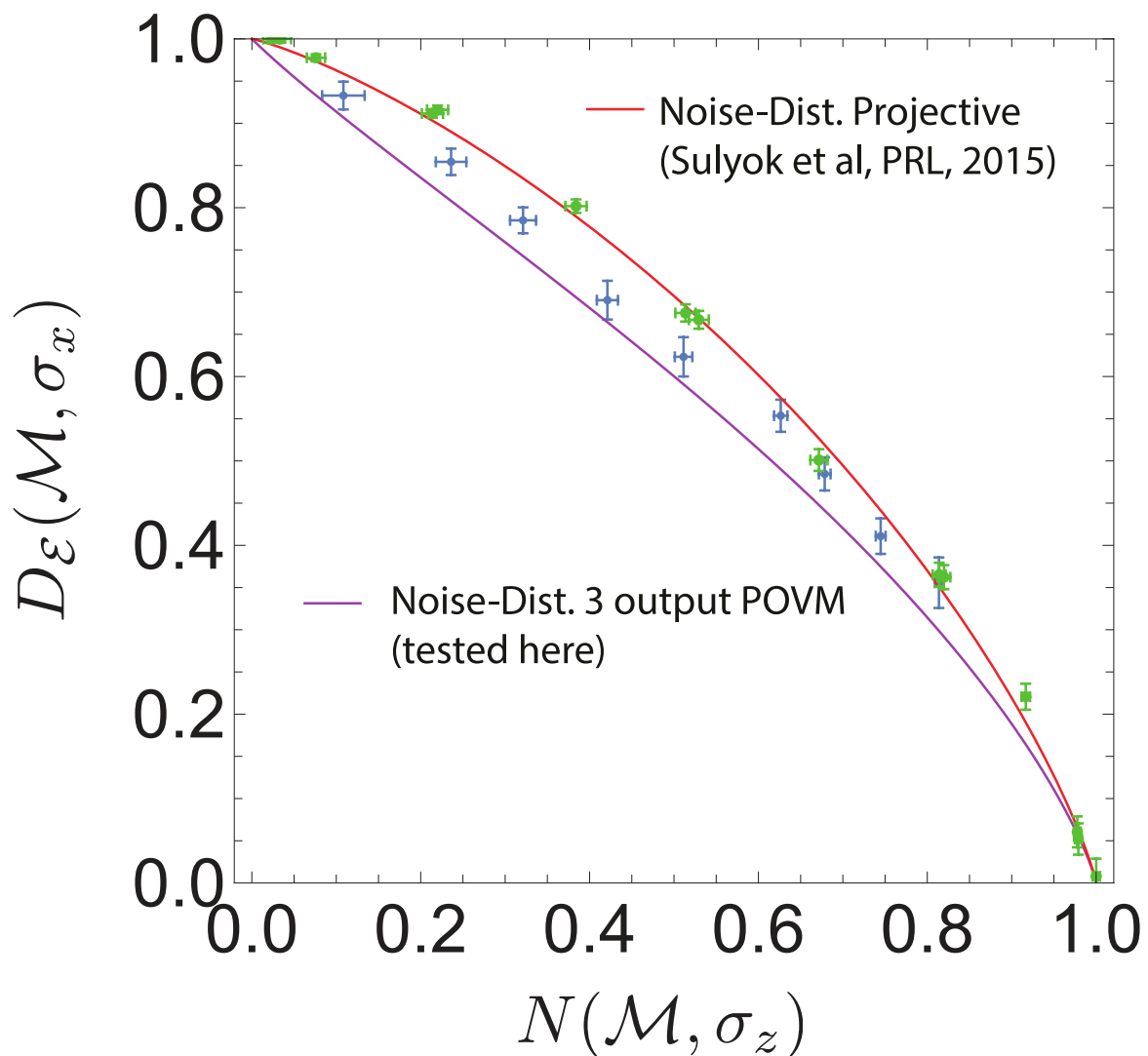


Figure 4.16: Comparison of the measurements using a three output POVM and a formerly done projective measurement (taken from [6]). The POVM measurement leads to a tighter uncertainty relation.

## 4.4 Discussion of the Results

We have shown that using a POVM proposed by *Abbott and Branciard* [5] yields in a tighter noise-disturbance uncertainty relation compared to projective measurements, as measured by *Sulyok et al.* [6].

Our measurement clearly outperforms the values obtained for the projective measurements over almost the entire range of the POVM parameter  $\theta$ . The theoretical prediction for the noise measurement could be reproduced over the full range of  $\theta$ . The values obtained for the disturbance measurement are slightly higher than the theoretical predictions. An explanation is given looking at the structure of the quantification of the disturbance.

The disturbance is computed using

$$D_\epsilon(\mathcal{M}^\theta, \sigma_x) = H(\mathbb{B}|\mathbb{B}') = - \sum_{b,b'} p(b, b') \log_2 \frac{p(b, b')}{p(b')} . \quad (4.26)$$

For simplification of the argument we set  $p(b') = 1$  and receive a sum over terms of the form

$$f(p) = -p \log_2 p . \quad (4.27)$$

When we look at the plot of  $f(p)$  (see Fig. 4.17) we can derive that for small values of  $p$  the function value  $f(p)$  is very sensitive to the input data. Only small fluctuations  $\Delta_1$  of an expected value  $p_1$  result in relatively large differences  $f(p_1 - \Delta_1) \ll f(p_1) \ll f(p_1 + \Delta_1)$ . The smaller the value  $p$  is, the more sensitive the function is to input data, as

$$p \rightarrow 0 \Rightarrow f'(p) \rightarrow \infty . \quad (4.28)$$

For probabilities with expected values  $p \gtrsim 0.1$   $f(p)$  is a lot less sensitive to fluctuations. Therefore especially the measurement of probabilities where we expect  $p = 0$  is critical. When measuring expected zero count rates, we typically measure values *close to zero, but never exactly zero*. Due to the structure of the quantification of entropic uncertainty relations discussed above, this results in a major change of the outcome.

In the case of the noise measurement the measured probabilities  $p(a|m)$  were in the range of  $p > 0.1$  over almost the entire range of the POVM parameter  $\theta$  for all input states  $a$  and measurement channels  $m$ . The results for the noise therefore resemble the theoretical prediction with a high statistical significance.

For of the disturbance measurement we used a summation over the joint probabilities  $p(b, b')$ . Unfortunately for  $p(b = -1, b' = +1)$  we expect a value of  $p = 0$ . The measurement results are plotted in Fig. 4.18.

When we now use the same measurement values as introduced in Section 4.2 for  $p(b, b')$  to quantify the disturbance by using Eq. (4.26) but replace the measured values of  $p(b = -1, b' = +1)$  with  $p(b = -1, b' = +1) = 0$  for all  $\theta$  we receive a significantly better disturbance plot which is given in Fig. 4.19. With these values for the disturbance we would not only *outperform* the projective measurements, but would also *match the theoretical predictions* for the 3-output POVM over almost the full range of the POVM parameter  $\theta$ . The noise-disturbance plot using the disturbance results of Fig. 4.19 is shown in Fig. 4.20.



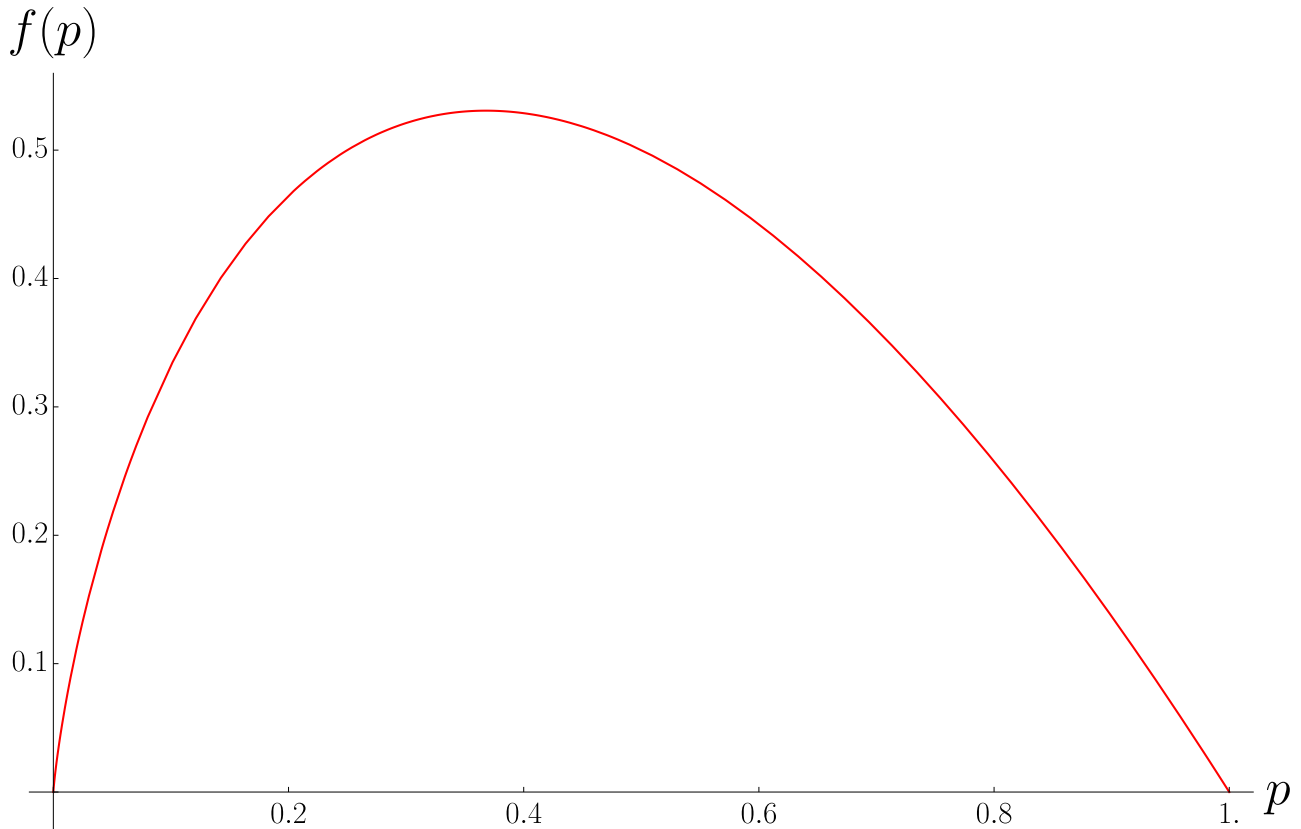


Figure 4.17: Plot of the function  $f(p) = -p \log_2 p$ . There is a steep increase of  $f(p)$  for small values of  $p$ .

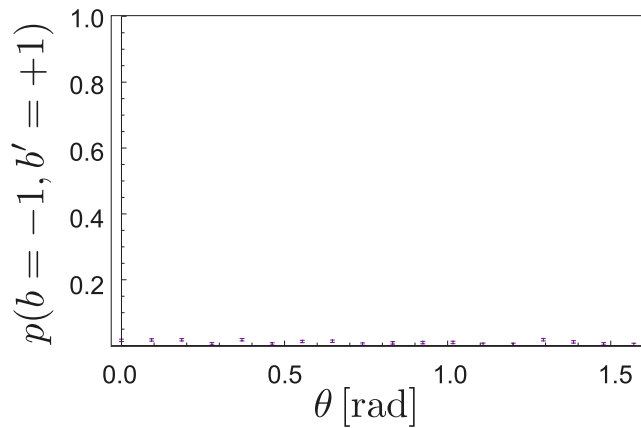


Figure 4.18: Plot of the measurement results of the joint probability  $p(b = -1, b' = +1)$  for the disturbance measurement.

The problem discussed here is not a topic that may be solved by further improvement of the setup, as it is an intrinsic property of the data handling and not an issue of the setup calibration. The setup was optimized with a lot of effort to increase the performance of the elements used. The data presented was acquired with a setup which can already be considered close to optimal. It is not realistic to further improve the disturbance measurement with the used 3-output POVM.

In general any measurements that aim to measure probabilities in the area of  $p \sim 0$  are problematic when it comes to the interpretation of the input data. A further investigation of possible data correction procedures that represent the physical reality is necessary to find a proper handling.

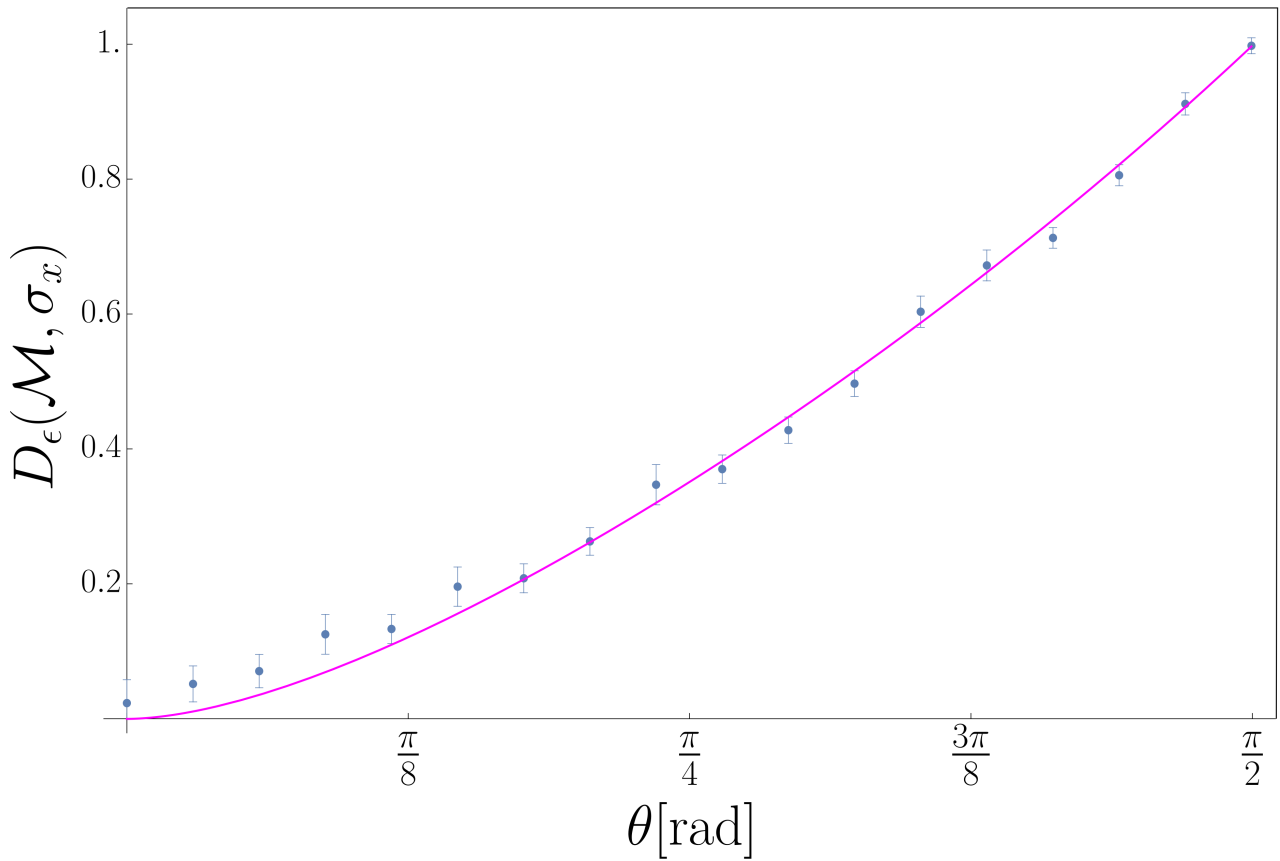


Figure 4.19: Disturbance with  $p(b = -1, b' = +1) = 0$  for all  $\theta$ . The results are significantly better than the ones obtained by using the experimental data for  $p(b = -1, b' = +1)$ . The disturbance using only experimental data is given in Fig. 4.15.

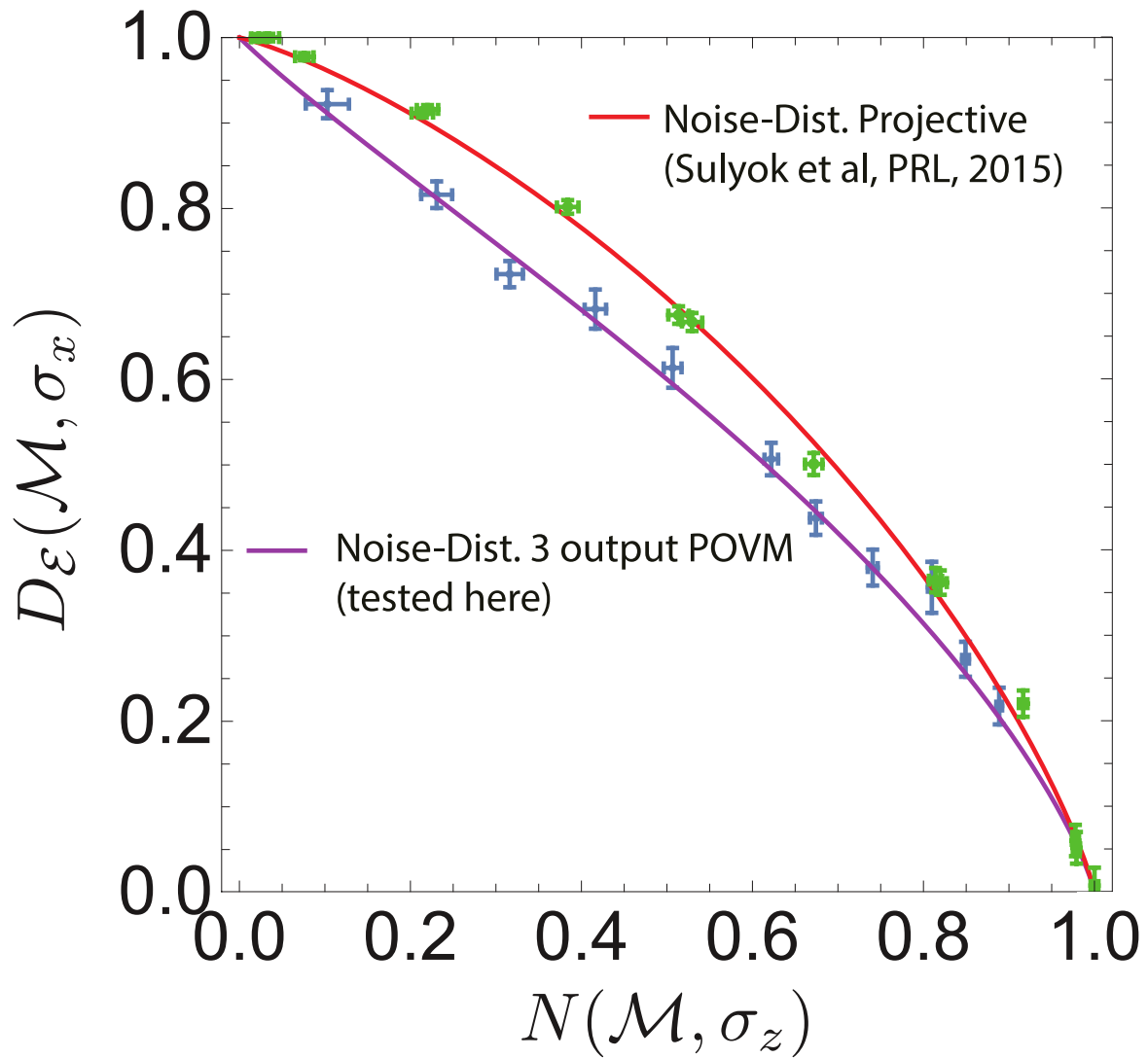


Figure 4.20: Noise-disturbance plot using the theoretical predictions for  $p(b = -1, b' = +1)$  instead of the measurement results for the disturbance measurement to show the problem of measuring probabilities  $p \sim 0$  when quantifying entropic uncertainty relations.



Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.  
The approved original version of this thesis is available in print at TU Wien Bibliothek.

---

Conclusion and Outlook

---



---

Why do all good things come to an end?

---

*Nelly Furtado*

In this thesis we experimentally investigated entropic noise and disturbance relations following the definitions from Buscemi et al. [4]. For the noise  $N(\mathcal{M}^\theta, \sigma_z)$  we used a 3-outcome POVM as proposed by Abbott and Branciard [5]. The input state was prepared as an eigenstate of  $\sigma_z$  with an uniform distribution for both possible states. The results obtained follow the theoretical predictions over the full range of the POVM parameter  $\theta$ .

For the disturbance  $D_\epsilon(\mathcal{M}^\theta, \sigma_x)$  we used the same POVM followed by a consecutive projective measurement. The input state was now prepared as an eigenstate of  $\sigma_x$  with an uniform distribution for both possible states. We were able to reproduce the trend of the theoretical predictions over the full range of the POVM parameter  $\theta$ . The results obtained from the measurement show a slightly higher disturbance than theoretically expected. This is due to probabilities that are theoretically expected to be  $p = 0$  which contribute to the disturbance quantification. For only small deviations of  $p \gtrsim 0$  a large offset of the computed disturbance values is generated. A detailed discussion of that effect is given in Section 4.4. How to properly deal with probabilities  $p \sim 0$  for entropic measurements is still an unresolved issue and needs further discussion and investigation. The problem is not only given for the measurements at hand, but for *any* entropic means of quantification which operates in the range of  $p \sim 0$ . This is a general problem of the data handling and not of the experimental setup or calibration.

The parametric noise-disturbance plot shows that the results using only projective measurements (Sulyok et al. [6]) are clearly outperformed with high statistical significance over almost the entire range of the POVM parameter  $\theta$ . The theoretical predictions could not be matched due to the non-ideal disturbance values.

For future research two main questions arise from the results of the measurements presented in this thesis. An unresolved issue is the handling of probabilities  $p \sim 0$  for entropic formulations of uncertainty relations. It has shown to be critical to find alternative approaches for the data handling that are both justifiable from a physical point of view and sufficiently resolve the issues associated with the current approach.

The second point of interest is the investigation of using two consecutive generalized measurements for the disturbance. So far only the first measurement apparatus used a generalized POVM measurement followed by a successive projective measurement. It would be of interest to replace the projective

measurement apparatus with a second generalized POVM measurement. So far, to the knowledge of the author, no theoretical work dealing with that kind of further generalization has been published. Therefore both a further theoretical examination of the usage of generalized measurements for noise-disturbance uncertainty relations as well as an experimental realization are subject to future research.

- 
- [1] W. Heisenberg, Über den anschaulichen Inhalt der quantentheoretischen Kinematik und Mechanik, *Z. Phys.* **43**, 172 (1927).
- [2] E. H. Kennard, Zur Quantenmechanik einfacher Bewegungstypen, *Z. Phys.* **44**, 326 (1927).
- [3] D. Deutsch, Uncertainty in Quantum Measurements, *Phys. Rev. Lett.* **50**, 631 (1983).
- [4] F. Buscemi, M. J. Hall, M. Ozawa and M. M. Wilde, Noise and Disturbance in Quantum Measurements: An Information-Theoretic Approach, *Phys. Rev. Lett.* **112**, 050401 (2014).
- [5] A. A. Abbott and C. Branciard, Noise and disturbance of qubit measurements: An information-theoretic characterization, *Phys. Rev. A* **94**, 062110 (2016).
- [6] G. Sulyok, S. Sponar, B. Demirel, F. Buscemi, M. J. W. Hall, M. Ozawa and Y. Hasegawa, Experimental Test of Entropic Noise-Disturbance Uncertainty Relations for Spin-1/2 Measurements, *Phys. Rev. Lett.* **115**, 030401 (2015).
- [7] M. Born, W. Heisenberg and P. Jordan, Zur Quantenmechanik. II., *Zeitschrift für Physik* **35**, 557 (1926).
- [8] H. P. Robertson, The Uncertainty Principle, *Phys. Rev.* **34**, 163 (1929).
- [9] E. Schrödinger, Zum Heisenbergschen Unschärfepinzip, *Sitzungsberichte der Preussischen Akademie der Wissenschaften, Physikalisch-mathematische Klasse* **14**, 296 (1930). Engl. translation at <http://arxiv.org/abs/quant-ph/9903100>.
- [10] P. J. Coles, M. Berta, M. Tomamichel and S. Wehner, Entropic uncertainty relations and their applications, *Reviews of Modern Physics* **89** (2017).
- [11] B. Demirel, S. Sponar and Y. Hasegawa, Measurements of Entropic Uncertainty Relations in Neutron Optics, *Applied Sciences* **10**, 1087 (2020).
- [12] A. Einstein, B. Podolsky and N. Rosen, Can Quantum-Mechanical Description of Physical Reality Be Considered Complete?, *Phys. Rev.* **47**, 777 (1935).
- [13] J. S. Bell, On the Einstein-Podolsky-Rosen paradox, *Physics (Long Island City, N.Y.)* **1**, 195 (1964).
- [14] S. J. Freedman and J. F. Clauser, Experimental Test of Local Hidden-Variable Theories, *Phys. Rev. Lett.* **28**, 938 (1972).
- [15] A. Aspect, J. Dalibard and G. Roger, Experimental Test of Bell's Inequalities Using Time-Varying Analyzers, *Phys. Rev. Lett.* **49**, 1804 (1982).
-

- [16] A. Aspect, P. Grangier and G. Roger, Experimental Realization of Einstein-Podolsky-Rosen-Bohm Gedankenexperiment: A New Violation of Bell's Inequalities, *Phys. Rev. Lett.* **49**, 91 (1982).
- [17] M. Giustina, M. A. M. Versteegh, S. Wengerowsky, J. Handsteiner, A. Hochrainer, K. Phelan, F. Steinlechner, J. Kofler, J.-A. Larsson, C. Abellán, W. Amaya, V. Pruneri, M. W. Mitchell, J. Beyer, T. Gerrits, A. E. Lita, L. K. Shalm, S. W. Nam, T. Scheidl, R. Ursin, B. Wittmann and A. Zeilinger, Significant-Loophole-Free Test of Bell's Theorem with Entangled Photons, *Phys. Rev. Lett.* **115**, 250401 (2015).
- [18] B. Hensen, H. Bernien, A. E. Dreau, A. Reiserer, N. Kalb, M. S. Blok, J. Ruitenber, R. F. L. Vermeulen, R. N. Schouten, C. Abellan, W. Amaya, V. Pruneri, M. W. Mitchell, M. Markham, D. J. Twitchen, D. Elkouss, S. Wehner and R. H. T. H. Tamini, Loophole-free Bell inequality violation using electron spins separated by 1.3 kilometres, *Nature (London)* **526**, 682 (2015).
- [19] L. K. Shalm, E. Meyer-Scott, B. G. Christensen, P. Bierhorst, M. A. Wayne, M. J. Stevens, T. Gerrits, S. Glancy, D. R. Hamel, M. S. Allman, K. J. Coakley, S. D. Dyer, C. Hodge, A. E. Lita, V. B. Verma, C. Lambrocco, E. Tortorici, A. L. Migdall, Y. Zhang, D. R. Kumor, W. H. Farr, F. Marsili, M. D. Shaw, J. A. Stern, C. Abellán, W. Amaya, V. Pruneri, T. Jennewein, M. W. Mitchell, P. G. Kwiat, J. C. Bienfang, R. P. Mirin, E. Knill and S. W. Nam, Strong Loophole-Free Test of Local Realism, *Physical Review Letters* **115** (2015).
- [20] C. E. Shannon, A Mathematical Theory of Communication, *The Bell System Technical Journal* **27**, 379 (1948).
- [21] P. A. M. Dirac, *The Principles of Quantum Mechanics* (Oxford University Press, 1981), 4 Auflage.
- [22] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Pr., 2010).
- [23] K. Kraus, Complementary observables and uncertainty relations, *Phys. Rev. D* **35**, 3070 (1987).
- [24] H. Maassen and J. B. M. Uffink, Generalized entropic uncertainty relations, *Phys. Rev. Lett.* **60**, 1103 (1988).
- [25] E. Rutherford, Bakerian Lecture. Nuclear Constitution of Atoms, *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences* **97**, 374 (1920).
- [26] J. Chadwick, Possible Existence of a Neutron, *Nature* **129**, 312 (1932).
- [27] M. Tanabashi, K. Hagiwara, K. Hikasa, K. Nakamura, Y. Sumino, F. Takahashi, J. Tanaka, K. Agashe, G. Aielli, C. Amsler, M. Antonelli, D. M. Asner, H. Baer, S. Banerjee, R. M. Barnett, T. Basaglia, C. W. Bauer, J. J. Beatty, V. I. Belousov, J. Beringer, S. Bethke, A. Bettini, H. Bichsel, O. Biebel, K. M. Black, E. Blucher, O. Buchmuller, V. Burkert, M. A. Bychkov, R. N. Cahn, M. Carena, A. Ceccucci, A. Cerri, D. Chakraborty, M.-C. Chen, R. S. Chivukula, G. Cowan, O. Dahl, G. D'Ambrosio, T. Damour, D. de Florian, A. de Gouvêa, T. DeGrand, P. de Jong, G. Dissertori et al., Review of Particle Physics, *Phys. Rev. D* **98**, 030001 (2018).
- [28] S. Sponar, *private communication* (TU Wien, Vienna, 2019).



---

## List of Figures

---

2.1	Plot of the probability mass function introduced in Example 2.3 . . . . .	9
2.2	The information content $I(x)$ of an event $x$ given by its probability $p(x)$ as defined in Eq. (2.14). One can see the $2^{-n}$ -dependency of $p(x)$ from $I(x)$ that can be derived from its definition. . . . .	12
2.3	The binary entropy function $H_b(p)$ , defined in Eq. (2.16). The maximal entropy is given at $p = 1/2 = 1/n$ . It is symmetric around $p = 1/2$ . . . . .	13
2.4	Visual representation of the Bloch sphere. . . . .	22
2.5	Illustration of the Kennard preparation uncertainty of two non-commuting observables $Q, P$ . The mean value is set to $\mu = 0$ , $\sigma^2(Q) = 0.5$ for the Gaussian distribution used (see Eq. (2.9)). . . . .	26
2.6	Illustration of the definition of the noise. A quantum system $S$ gets prepared in a state $ \psi^a\rangle$ . The measurement apparatus $\mathcal{M}$ is changing the state of $S$ . We receive an altered quantum system $S'$ and a measurement outcome $M$ . A guessing function tries to find the right eigenvalues associated with the measurement output $M$ . The correlation of the input state and the measurement output is quantified by the noise $N(\mathcal{M}, \hat{A})$ . . . . .	27
2.7	Illustration of the definition of the disturbance. The first part is analogue to the noise definition. A quantum system $S$ is prepared in an eigenstate $ \phi^b\rangle$ . The system $S$ gets fed into the measurement apparatus $\mathcal{M}$ . The output is a quantum system $S'$ , altered by the measurement, and the measurement output $M$ . They both get now fed in a correction procedure $\epsilon$ , which tries to restore the original state $S$ with the help of the information of $M$ . The result is the corrected quantum system $S''$ . A successive measurement apparatus $\mathcal{M}'$ performs a measurement on $S''$ . The correlation between the input state of $S$ and the measurement output $M'$ is quantified by the disturbance $D(\mathcal{M}, \hat{B})$ . . . . .	29
3.1	Feynman diagram illustrating the decay mode (1) given in Eq. (3.1). A down-quark of the neutron decays into an up-quark, emitting an intermediary $W^-$ gauge boson. The electron and the electron-antineutrino are generated from that boson. . . . .	32
3.2	Larmor precession of a magnetic moment $\vec{\mu}$ around an external magnetic field $\vec{B}$ with angular frequency $\omega$ . . . . .	33
3.3	Probability density function of a Maxwell-Boltzmann velocity distribution for thermal neutrons at $T = 290$ K (see Eq. (3.6)). . . . .	34
3.4	Illustration of the Bragg principle. Incoming beams (in red) interfere constructively if the Bragg condition (Eq. (3.8)) holds. Therefore one can monochromatize, as all wavelenghts other than the desired one do not interfere constructively and get suppressed. The intensity of the outcoming beam decreases with higher diffraction orders. . . . .	35

3.5	Old measurement of the incident beam energy spectrum with the Time-Of-Flight method. The units are arbitrary. We see a major peak at $\sim 2 \text{ \AA}$ . The other peaks are higher orders from the monochromator crystal. . . . .	36
3.6	Principle of a neutron supermirror. Alternating layers of magnetic and nonmagnetic materials are used. The diameter of the layers is monotonously decreasing ( $d_1 > d_2 > d_3 > \dots$ ). An incoming beam with arbitrary spin orientations gets split into two beams in a pure spin state (ideal case). The undesired spin state (here: $ \downarrow\rangle$ ) gets absorbed later on by a substrate, the other polarized beam leaves the supermirror. . . . .	37
3.7	Supermirror used in our setup. The exit window for the beam is visible on the left side of the mirror. . . . .	38
3.8	Calibration measurement of the supermirror tilt and the resulting intensities dependent on the tilt angle $\theta$ . The actual values on the $x$ -axis are positions of a stepper motor performing the tilt and are therefore not relevant. The full range of the plot covers an angle of only about $1.3^\circ$ . . . . .	39
3.9	Simulation of the field caused by the rectangular Helmholtz coils used in the setup. The plot shows the field distribution in $z$ -direction for different $x$ - and $y$ -values. The 2D plot (see Fig. 3.10) is the cut of this 3D plot at $y = 0$ . The magnetic field in the area where the neutron beam is located is very homogeneous. The white areas are due to a divergence of the integral at the source of the magnetic field. The field inside the coils is not computed. . . . .	40
3.10	Simulation of the magnetic field of the guidefield. The plot shows the magnetic field in the middle of the coil setup over the full $x$ -range. The neutron beam is most likely to travel in the center in a well aligned setup. Therefore this simulation shows the magnetic field the beam is most likely to be exposed to. The field strength in the middle is in the area of $B_z \sim 10 \text{ G}$ . . . . .	41
3.11	Rectangular Helmholtz coils forming a guidefield. The adjacent supermirrors and two spinrotator coils in the area of the guidefield are also visible. The spinrotators are fixed on a pedestal that can be moved by stepper motors to adjust the distance for a full Larmor precession. . . . .	42
3.12	Picture of a typical spinrotator coil. The current is applied via the screw joints on top. Coils of the same dimensions are used between the polarizer and first analyzer of our setup. . . . .	43
3.13	Current sweep for the $z$ -winding of a spinrotator coil. The ideal setting is found in the minimum. . . . .	44
3.14	Current sweep for the $y$ -winding of a spinrotator coil. With different currents we can alter the spin orientation by performing different spin rotations. . . . .	45
3.15	Distance measurement of a spinrotator pair. The distance representing a full Larmor precession is given for the maximum of the countrate. The distance between the coils is varied with a stepper motor. The positions of the spinrotator coils DC3 and DC4 and the stepper motor that varies the distance is shown in Fig. 3.11 . . . . .	45
3.16	Calibration measurement of the neutron detector used as monitor for the setup. A high voltage (HV) is applied, the corresponding countrates are measured. The final setting for the HV was chosen at $U = 1 \text{ kV}$ . For this setting we are in saturation and a further increase would not improve the detection quality. . . . .	47
3.17	Spectrum of the $He$ detector used as a monitor in our setup, measured with a multi-channel analyzer. The domain that is associated with an incident neutron is highlighted in red. Only in that energy range occurring pulses are considered neutrons and get counted. The $\gamma$ -peak is well recognizable on the left side of the plot. . . . .	48
3.18	Two differently sized $He$ neutron detectors. Both of them were used during the calibration of the final setup for our noise-disturbance measurements. . . . .	48

3.19	Illustration of the POVM under investigation in the experiment. $M_0$ always points in $ +x\rangle$ -direction and is shortened with increasing $\theta$ , vanishing for $\theta = \pi/2$ . $M_{+1}$ and $M_{-1}$ are equal for $\theta = 0$ and point in $ +x\rangle$ -direction. For increasing $\theta$ the vectors begin to move towards $ \uparrow\rangle$ respectively $ \downarrow\rangle$ and are elongated. Their movement is along the dashed red lines. . . . .	50
3.20	Detailed measurement of the relative POVM weights and the corresponding stepper motor positions. The intensities are normalized to the defined value of $p_m := 1$ at axis position $-500$ k. . . . .	51
3.21	Scheme of the setup for the noise measurement. The first supermirror and spinrotator prepare the state with unitary probability for $ \uparrow\rangle$ and $ \downarrow\rangle$ . The neutron beam leaves the polarizer in the pure state $ \uparrow\rangle$ . The spinrotator coil DC1 performs a $\pi$ -flip, resulting in the state $ \downarrow\rangle$ or no current is applied at all, leaving the beam in state $ \uparrow\rangle$ . The probability of a performed $\pi$ -flip is $p(\pi) = \frac{1}{2}$ . The second coil DC2 maps the Bloch vector of the POVM to the $ \uparrow\rangle$ axis and the second supermirror acts as analyzer and applies the POVM weights. We measure the joint probabilities $p(m, a)$ with a neutron detector that is directly mounted on the analyzer and placed right after the exit window of the supermirror. . . . .	52
3.22	Picture of the setup for the noise measurement. The detector is directly mounted on the analyzer supermirror. The POVM weight tilt is performed by a stepper motor. The spin orientation is manipulated by a coil pair inside a guidefield, that is situated between polarizer and analyzer supermirror. The first coil prepares the desired input state $ a\rangle$ , the second coil maps the POVM Bloch vectors on the $ \uparrow\rangle$ -axis. The second supermirror acts as an analyzer of the spin state. It also applies the reduction of the intensity according to the POVM weight function $p_m$ by being tilted. . . . .	52
3.23	Detailed view of the region with DC3 and DC4. For the disturbance measurement the exit window of the first analyzer supermirror is not covered by the detector that was mounted directly on the supermirror for the noise measurement. The coil pair DC3 and DC4 is placed with a distance that represents a full Larmor precession. The DC3 performs the correction operation, the DC4 turns the projectors of the $B$ measurement in $ \uparrow\rangle$ -direction. . . . .	53
3.24	Overview of the full setup. All 3 supermirrors and both guidefield regions with each 2 spinrotator coils are visible. The last supermirror is a different type than the other ones. The detector for the disturbance measurement is in the background shielding built from bricks of borated paraffin (in white, coated with black boron carbide mats). . . . .	54
3.25	Scheme of the setup for the disturbance measurement. For the disturbance measurement we need to prepare the input states $ +\rangle$ and $ -\rangle$ . The neutron beam leaves the first supermirror in the state $ \uparrow\rangle$ . The first spinrotator coil DC1 performs either a $+\frac{\pi}{2}$ flip, resulting in the state $ +\rangle$ or a $-\frac{\pi}{2}$ flip, resulting in $ -\rangle$ . The probabilities for the flips are $p(+\frac{\pi}{2}) = p(-\frac{\pi}{2}) = \frac{1}{2}$ . This results in a unitary probability for the input states $ b\rangle$ . The POVM measurement is performed exactly the same way as for the noise measurement. The measurement of the three POVM elements is done consecutively. The Bloch vectors corresponding to the POVM elements are mapped on the $ \uparrow\rangle$ -axis by the second spinrotator coil DC2. A second supermirror acts as analyzer and also applies the respective POVM weights by being tilted as described in Section 3.3.1. This results in a measurement as defined in the POVM. After the first analyzer another spinrotator coil pair is situated. The first one, DC3, is performing the correction procedure for the currently measured POVM element. For the correction of $M_0^\theta$ this means a $+\frac{\pi}{2}$ rotation, for the correction of $M_{\pm 1}^\theta$ this means a $-\frac{\pi}{2}$ rotation. The second one, DC4, consecutively maps $ +\rangle$ and $ -\rangle$ onto the $ \uparrow\rangle$ -axis. Together with the second analyzer supermirror it forms a projective measurement for $\sigma_x$ . The detector is placed after the last supermirror. We receive the joint probabilities $p(b', b)$ . . . . .	55

4.1	Raw measurement data directly received from the experiment. Three consecutive measurement settings form a point for the parameter $\theta$ . . . . .	58
4.2	Corrected and detangled measurement data for all three POVM channels. . . . .	59
4.3	Joint probabilities acquired by normalization of the intensities. (a) Including a statistical error $\sqrt{N}$ . (b) Additional consideration of a systematic error $\Delta_{sys} \simeq 0.7^\circ$ . . . . .	60
4.4	Joint probabilities for $M_{+1}^\theta$ . (a) Joint probability received by the measurement data plotted over the theoretical curves for $M_{+1}^\theta$ . The input state is still "indefinite". (b) Joint probabilities $p(a = +1, m = +1)$ (red) and $p(a = -1, m = +1)$ (blue) consisting of a combination of measured and computed values. . . . .	61
4.5	Conditional probabilities for the $M_{+1}^\theta$ channel. On the left it is given for the input $ \uparrow\rangle$ (red), on the right for $ \downarrow\rangle$ (blue). . . . .	61
4.6	Joint probabilities for $M_0^\theta$ . (a) Joint probability received by the measurement data plotted over the theoretical curves for $M_0^\theta$ . The input state is "indefinite". (b) Joint probabilities $p(a = +1, m = 0)$ (red) and $p(a = -1, m = 0)$ (blue) which are equal for all $\theta$ . . . . .	62
4.7	Conditional probabilities for the $M_0^\theta$ channel. On the left it is given for the input $ \uparrow\rangle$ (red), on the right for $ \downarrow\rangle$ (blue). They are $p = \frac{1}{2}$ for all $\theta$ . The values have a very small error, please note the zoomed area in the plots. . . . .	62
4.8	Joint probabilities for $M_{-1}^\theta$ . (a) Joint probability received by the measurement data plotted over the theoretical curves for $M_{-1}^\theta$ . The input state is "indefinite". (b) Joint probabilities $p(a = +1, m = -1)$ (red) and $p(a = -1, m = -1)$ (blue) consisting of a combination of measured and computed values. . . . .	63
4.9	Conditional probabilities for the $M_{-1}^\theta$ channel. On the left it is given for the input $ \uparrow\rangle$ (red), on the right for $ \downarrow\rangle$ (blue). . . . .	63
4.10	Final plot of the noise $N(\mathcal{M}^\theta, \sigma_z)$ of the POVM $M_m^\theta$ with input states as eigenstates of $\sigma_z$ as a function of the POVM parameter $\theta$ . The red line is the theoretical prediction. . . . .	64
4.11	Raw data of the <i>disturbance</i> measurement $D_\epsilon(M^\theta, \sigma_x)$ of the three output POVM $M^\theta$ and projective $B = \sigma_x$ measurement for 400 seconds. . . . .	65
4.12	Normalized data with "indefinite" input state $ b\rangle =  +x\rangle \vee  -x\rangle$ split up in the two output channels of the subsequent projective measurement for $\sigma_x$ . (a) Measurement channel $M_{+1}^\theta$ . (b) Measurement channel $M_0^\theta$ . (c) Measurement channel $M_{-1}^\theta$ . . . . .	67
4.13	The joint probabilities $p(b, b')$ for all four permutations of $b$ and $b'$ plotted together with the theoretical predictions. . . . .	69
4.14	The marginal probabilities $p(b')$ together with the theoretical predictions. On the left side we see $p(b' = +1)$ , on the right $p(b' = -1)$ . . . . .	69
4.15	The final disturbance plot. The first ten points (for low $\theta$ ) were measured with $t_{meas} = 800$ s to reduce the statistical error. The trend of the theoretical prediction is reproduced over almost the entire range of the POVM parameter $\theta$ . The disturbance values measured are slightly higher as the predictions. . . . .	70
4.16	Comparison of the measurements using a three output POVM and a formerly done projective measurement (taken from [6]). The POVM measurement leads to a tighter uncertainty relation. . . . .	71
4.17	Plot of the function $f(p) = -p \log_2 p$ . There is a steep increase of $f(p)$ for small values of $p$ . . . . .	73
4.18	Plot of the measurement results of the joint probability $p(b = -1, b' = +1)$ for the disturbance measurement. . . . .	73
4.19	Disturbance with $p(b = -1, b' = +1) = 0$ for all $\theta$ . The results are significantly better than the ones obtained by using the experimental data for $p(b = -1, b' = +1)$ . The disturbance using only experimental data is given in Fig. 4.15. . . . .	74

- 
- 4.20 Noise-disturbance plot using the theoretical predictions for  $p(b = -1, b' = +1)$  instead of the measurement results for the disturbance measurement to show the problem of measuring probabilities  $p \sim 0$  when quantifying entropic uncertainty relations. . . . . 75



Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.  
The approved original version of this thesis is available in print at TU Wien Bibliothek.

---

## Acknowledgements

---

Ich habe fertig!

---

*Giovanni Trapattoni*

This thesis would not have been possible without the support and contribution of many people. Some of them I want to thank decidedly, knowing that I will forget at least half of those who contributed. I also only focused on people that directly worked with me scientifically, otherwise this list of appreciations would be longer than the rest of the thesis, which would be kind of inappropriate as I actually really want to finally earn that degree.

I want to thank Richard Wagner whom I got to know during a scientific visit at ILL in Grenoble for pointing out the opportunity to work in the neutron interferometry group and performing my diploma thesis in this area. And being an easy and nice dude that is always great to have around. And for showing me an alternative bicycle route from home to the Atominstitut.

I specifically want to thank Armin Danner a.k.a. Typohunter3000, who introduced me to both the research group and neutron polarimetry. He was a very important support which was definitely necessary when being thrown in the cold reactor water at the very beginning of this endeavour. He also was the most conscientious reader of every single part of this thesis, resulting in a lot of constructive feedback and the death of billions of typos. That's where he's got his alter ego name from.

Another big thank you goes out to Stephan Sponar, for many interesting discussions and support, bringing in his experience and helping me with weird nasty Mathematica notebooks. And for fixing problems with the normalization in the calculation files, because I really really really did not want to do that all alone. And in general for all the coffee breaks, inspiration and good advice in the right situations. He always took time, even though his plate is usually quite full anyway.

A big thank goes out to Yuji Hasegawa who led me through the so far biggest scientific adventure of my life and both gave me guidance when needed but also left me enough free space to work on the project in my tempo and style. A great combination that is not always easy to maintain, especially when dealing with a stubborn goat as I am. And for never ever guiding with authority, but wisdom.

I also want to thank the whole group – not only the interferometry guys but also the ones from the neutron group and other scientific groups at the Atominstitut – for all the enlightening coffee breaks and the intellectual and especially not so intellectual discussions and exchange. Of course Elisabeth Kreuzgruber, Niels Geerits and 帶金一真 (Kazuma Obigane) have a special role, as we shared an office for many months and therefore suffering and laughter in a more intimate way than with the guys next door.

A very special thank goes out to Carina Killian who beat me to finish our masters about around 2 months. How dare you! She contributed pretty much nothing to this thesis, but as it also marks the end of six years of studying, she definitely deserves an honorable mention. We met at the first day of university when we were still young and fresh and I am glad that she did not only become my prime study buddy but also one of the closest friends I have today. I am very grateful for that.

Vito Pecile

Vienna, March 2020