Quantum mechanics: completely unhidden?
A critical review of the work of Colbeck and Renner

Gijs Leegwater

August 14, 2013

Thesis for the master History and Philosophy of Science
Supervised by prof. dr. D. G. B. J. Dieks
Institute for History and Foundations of Science,
Faculty of Science,
Utrecht University

Abstract

In this thesis a recent claim by Roger Colbeck and Renato Renner is reviewed [18]. This claim concerns the completeness of quantum mechanics. Colbeck and Renner state that no theory can be constructed that is compatible with quantum mechanics, and gives more detailed predictions through supplemental variables added to the description of the state of a system. Their only assumption is that measurement settings can be choosen freely. We will argue that their result is for the most part valid and interesting, but the freedom assumption is not so innocent as it is presented, and parts of the result need to be clarified. Therefore, we will present a reformulation of the result. In our formulation the result is a no-go theorem for hidden variable theories satisfying Parameter Independence. This is a generalization of Bell's theorem.

Performed in the period 01-11-2012 - 14-08-2012

Universiteit Utrecht

Contents

Chapter 1

Introduction

1.1 Historical overview

The probabilistic nature of quantum mechanics has puzzled many. Famously, in a letter sent to Max Born in 1926 Einstein wrote [11]

> 'I, in any case, am convinced that *He* [God] does not play dice'[1].

Einstein himself was one of the originators of quantum mechanics: by discovering the photo-electric effect [25] in 1905, he set in motion the development of quantum mechanics. However, he was unsatisfied with one aspect of the theory—it gives probabilistic predictions, even if the quantum mechanical state of a system is completely known. In classical mechanics randomness is always related to ignorance about a system; at its core the theory is deterministic. However, in quantum mechanics, randomness is a fundamental part of the theory; it appears in its basic postulates. Einstein did not believe that Nature is probabilistic at its most fundamental level.

1.1.1 The Einstein-Podolsky-Rosen argument

In 1935 Albert Einstein, Boris Podolsky and Nathan Rosen (EPR) published an article to show that quantum mechanics is *incomplete* [26]. They argued that there are situations in which particles can have real properties ('elements of reality') that are not part of the quantum mechanical description of the state of the system.

By considering measurements on two entangled particles, performed at space-like separation, EPR showed that if it is assumed that the real properties of a particle are independent of what measurement is performed on the other particle, then the particle must have real properties which determine both the outcome of a position measurement and the outcome of a momentum measurement. But the quantum mechanical state of a particle obeys Heisenberg's uncertainty principle, which means that if the quantum mechanical state provides a single value for the position of a particle, it does not provide a value at all for the momentum. Therefore the quantum mechanical description is incomplete.

[1] 'Jedenfalls bin ich überzeugt, daßder nicht würfelt.'

In the final section of the article the authors expressed their belief that a deterministic completion of quantum mechanics is possible: a theory which does not give probabilistic predictions for every measurement outcome, but definite predictions.

1.1.2 Bohm's nonlocal deterministic theory

In the derivation of the incompleteness of quantum mechanics, Einstein and his co-authors used a locality assumption:

> '... since at the time of measurement the two systems no longer interact, no real change can take place in the second system in consequence of anything that may be done to the first system. This is, of course, merely a statement of what is meant by the absence of an interaction between the two systems.'

Reading the last sentence of this quote, it appears that EPR consider this as an assumption that can hardly be denied. So, when the belief is expressed that a completion of quantum mechanics is possible, they aim at a completion which satisfies this assumption. Einstein was looking for a *local* deterministic theory.

That a deterministic theory compatible with quantum mechanics is possible was shown by David Bohm in 1952 [8]. Bohm found such a theory, which was a rediscovery of the 'pilot-wave theory' discovered in 1927 by Louis de Broglie [10]. It is known as Bohmian mechanics. While John von Neumann thought he had proven that any deterministic theory is incompatible with quantum mechanics [51], Bohm's example proved that this is not the case. Bohm's theory is however nonlocal, therefore it is not the kind of theory Einstein dreamt of.

1.1.3 Bell's no-go theorem

The question remained whether a local deterministic theory was possible. This was answered in 1964 by John S. Bell [5]. Bell had read about von Neumann's no-go theorem, and when he discovered that Bohm provided a counterexample, he wondered whether nonlocality was a necessary ingredient for deterministic theories. Using a mathematical argument, he showed this to be the case. Bell derived an inequality involving outcomes of certain experiments which must be satisfied by any local deterministic theory underlying quantum mechanics. However, quantum mechanics predicts that those inequalities are violated. Therefore, quantum mechanics can not be substituted by a local deterministic theory, if the quantum mechanical predictions are correct.

Bell's original inequality was slightly generalized by Clauser, Holt, Shimony and Horne (CHSH) [16], who derived inequalities that must be satisfied by *any* local deterministic theory. These inequalities are now also known as 'Bell inequalities'. They have a theory-independent value: if they are violated experimentally, all local deterministic theories are ruled out, even if quantum mechanics is wrong! The most famous experimental violation of a Bell inequality is the one realized by Aspect et al. in 1982 [2].

1.1.4 The birth of quantum computation

The work of Bell clearly made visible the peculiar nature of entangled systems. This raised interest in the question whether entanglement can be *used as a resource* to perform tasks.

In 1981, Richard Feynman pondered over the fact that it is very hard to simulate quantum mechanics using a classical computer [28]. Turning this around this means that using quantum systems it might be possible to perform calculations much more efficient than with a classical computer. The idea of quantum computing was born, as was the challenge of building a quantum computer.

The state spaces of classical and quantum bits

In a classical computer, the fundamental unit of information is a bit that can have the value 0 or 1. This means that the state space of a classical bit consists of two points, while the state space of n classical bits consists of 2^n points. In a quantum computer, the fundamental unit is a quantum bit (qubit). A qubit is a 2-dimensional quantum system which means that the associated state space is the 2-dimensional Hilbert space $\mathbb{C}^2$. For n qubits, the state space is $\mathbb{C}^{2n}$.

It is obvious that the state space of qubits is vastly larger than the state space of quantum bits. This suggests that quantum bits carry much more information than classical bits. However, when reading out n quantum bits, the result consists of only n classical bits. So the *accessible* information of quantum bits still equals that of classical bits. There are however some computational tasks that can be performed much more efficiently using qubits than with classical bits. In 1993 Peter Shor showed that quantum computers would be able to factorize large numbers much faster than classical computers [48]. In 1996 Lov Grover found an efficient search algorithm for quantum computers [34]. Fourier transformations can also be performed efficiently on quantum computers [35].

While the factorization of big numbers poses a threat to security protocols, since these are based on such factorizations, at the same time quantum mechanics offers a more secure protocol. Artur Ekert [27] showed in 1991 that quantum systems can be used to send information which is impossible to intercept without the receiver noticing. This result can be placed in a field of physics which is closely related to quantum computation, namely *quantum information theory.*

Practicing quantum foundations

As illustrated above, the promise of a quantum computer has motivated research in quantum computation and quantum information in the last few decades. An interesting side-effect is that because these fields are about applications of fundamental features of quantum mechanics like entanglement and nonlocality, there has also been an offspring of results in the foundations of quantum theory. An example is the *no-cloning theorem*, formulated in 1982 by Dieks [23], and independently by Wootters and Zurek [52], which states that a quantum state cannot be cloned. Another example is *quantum teleportation* [7].

In other lines of research attempts are made to reformulate quantum mechanics in information-theoretic terms and to derive it from information-theoretic principles [36]. Some have even claimed that 'information is physical' [38].

In 2012, Matthew Pusey, Jonathan Barrett and Terry Rudolph (PBR), who also work in the fields of quantum information and computation, published a

surprisingly simple result about the relation between the quantum state of a system and its possible 'ontic states' [42]. It was found that if one assumes that there is some real ('ontic') state underlying the quantum state, then it is not possible to assign two different quantum states to two systems of which the ontic state is the same. Therefore, it seems that the quantum state is a real property of a system.

1.2 Overview of the work by Colbeck and Renner

The work of Roger Colbeck and Renato Renner (CR), which is the focus of this thesis, is another example of foundational offspring from quantum computation and information. CR have published results about the completeness of quantum mechanics [18][20], and, like PBR, about the relation between quantum mechanical states and reality [19].

Colbeck and Renner claim that quantum mechanics is complete: there is no compatible theory reducing the randomness of its predictions. From this they also reach a conclusion even stronger than PBR's: that there is a 1-to-1 correspondence between quantum states and ontic states. Not only is the quantum state a real property of a system, it is the whole reality of it!

The claims of Colbeck and Renner are quite strong and therefore interesting. If correct, this would be an extensive generalization of Bell's result: not only local deterministic theories are incompatible with quantum mechanics, but *any* extension of it.

Colbeck and Renner have tried to make their argument as general as possible by trying to derive it from only one assumption: that experimenters are free in choosing measurement settings. From this assumption they derive a no-signaling condition which plays a crucial role in their argument. Unfortunately this move is not so strong, since it appears there are assumptions hidden in their definition of 'freedom'. Therefore, almost all reactions to their results have been aimed at this part. This is at the expense of the more interesting parts of their result.

Another drawback of the results of CR is that they are published only in short form. The details of their results can not be found in the main articles, but only in separate supplementary sections. Also, the explanations in these supplementary are hard to follow. Therefore the accessibility of their results could be a lot better. Not surprisingly, some authors have completely misunderstood their result: see the quote in the next section and Chapter 5.

1.3 Aim and structure of this thesis

A mathematical proof of a theorem is usually not discovered only by trying out mathematical relations. Often there is some underlying idea, a story that can be told in words. However, when a proof is published, a necessary ingredient is the mathematical proof itself, while the underlying idea can be omitted. Getting the idea of a proof next to its mathematical form is however of great value for the insight, and for inspiration for future proofs.

The published versions of the result of Colbeck and Renner regarding possible extensions of quantum theory are a bit like such a mathematical proof

without the conceptual context. The only way to appreciate the underlying idea is to wade through all the mathematical details, which are hidden away in supplementary sections, spread over multiple publications, and sometimes missing. Therefore, the accessibility of the result could be improved a lot if the underlying idea is described more explicitly. Therefore, in this thesis the result has been reformulated and presented step-by-step. We hope that this makes the result more accessible, and also that this provides a better insight in the result.

That a better understanding of Colbeck and Renner's result is necessary is also manifest in the following quote by Valerio Scarani, who is group leader at the Centre for Quantum Technologies, National University of Singapore:

> '... Beyond the case of the maximally entangled state, which had been settled in a previous paper, they prove something that I honestly have not fully understood. Indeed, so many other colleagues have misunderstood this work, that the authors prepared a page of FAQs (extremely rare for a scientific paper) and a later, clearer version [20].
>
> Comment: the statement "Colbeck and Renner have proved that quantum theory cannot be extended" is amazingly frequent in papers, referee reports and discussions. Often, it comes in the version: "why are people still working on [whatever][*sic*], since Colbeck and Renner have conclusively proved...?" It is pretty obvious however that many colleagues making that statement are not aware of the "details" of what Colbeck and Renner have proved: they have simply memorized the bumper sticker statement [("No extension of quantum theory can have improved predictive power")]. I really dont have a problem with Colbeck and Renner summarizing their work in a catchy title; what is worrisome is other experts repeat the catchy title and base decisions solely on it.'

In order to provide conceptual clarity, we start with a 'pedagogical prelude' in Chapter 2.

After that, in Chapter 3 we will present a reformulation of the result of CR. In our opinion, the reformulated result is what is left when the right assumptions are made. It includes a compact derivation of the results presented in Chapter 2.

In Chapter 4 we present the main differences between our reformulation and the original results published by CR, in order to establish a link between their work and the reformulation. Aside from criticism to CR's 'Freedom of Choice' assumption, this chapter also treats the subject of assigning probability distributions to settings, outcomes and hidden variables, which, we argue should only be done with great care. Also some interesting thoughts regarding no-collapse interpretations of quantum mechanics are formulated.

An example of a theory which falls under the scope of the CR result will be given in Chapter 5. This theory belongs to a class of hidden variable models proposed by Antonio Di Lorenzo [22]. While Di Lorenzo claims that his models are not in conflict with the CR result, we show this is a misunderstanding: when correctly applying the CR result, Di Lorenzo's models are ruled out.

Finally we will summarize the results of this thesis, highlight some difficulties and propose some ideas for further research in Chapter 6.

1.4 Preliminaries and notation

1.4.1 Probability

We will distinguish between two kinds of variables: *random* variables and *context* variables. A random variable is a variable assigned to the result of some stochastic process, for example a measurement outcome. A context variable is a variable which defines the context of a probability distribution. For example, this can be a measurement setting. Random variables always appear within the parenthesis of a probability distribution $P(\ldots)$ while context variables always appear outside the parenthesis, as a subscript. In Section 4.1 we will discuss why we make this distinction.

In this thesis many different probability distributions are discussed. Therefore we need a notation which is flexible. However, specifying each variable and each particular value of it every time results in cluttered notation like $P(X = x, Y = y|A = a, B = b, Z = z)$. The following notation is proposed to improve readability while remaining flexible.

Uppercase letters are used for both random and context variables. Mostly we will use X, Y, Z for random variables and U, V, W for context variables. Lowercase letters $u, v, w, x, y, \ldots$ are used for particular values of the variable of the corresponding uppercase letter. If these lowercase letters are used, the uppercase letter is omitted. In some situations we need to use different lowercase letters.

For example, we write

$$P(x)_v \tag{1.1}$$

for $P(X = x)_{V=v}$, which is the probability of X having the value x, given that the context variable V equals v.

An example of a case where we need to use different lowercase letters is when we consider the probability of two random variables having the same value. For example,

$$P(X = i, Y = i)_{V=j,W=j} \tag{1.2}$$

is the probability that X and Y both equal i, given that the context variables V and W both equal j.

Now, we list some further definitions and probability laws:

- Domain of X:

$$\tilde{X} \tag{1.3}$$

  is the set of values X can take.

- Negation:

$$P(\overline{x}) := P(X \neq x) = 1 - P(X = x) \tag{1.4}$$

  For dichotomous variables with domain $\{-1, +1\}$, we will also write

$$\overline{A} := -A. \tag{1.5}$$

- Conditional probability:

$$P(X = x|Y = y) = \frac{P(X = x, Y = y)}{P(Y = y)} \tag{1.6}$$

- Law of total probability:

$$\sum_y P(x, y) = P(x); \tag{1.7}$$

$$\sum_z P(x|z)P(z) = P(x); \tag{1.8}$$

$$\sum_z P(x|y, z)P(z|y) = P(x|y). \tag{1.9}$$

- Subadditivity

$$P(x, y) \leq P(x). \tag{1.10}$$

Also, sometimes we want to define a new random variables in terms of an existing one, given a specific context. For this we use a vertical bar. For example, if we have a probability distribution

$$P(x)_v, \tag{1.11}$$

we may define

$$A := X\big|_{V=1}, \tag{1.12}$$

$$B := X\big|_{V=2}. \tag{1.13}$$

Then, probability distributions are related in the following way:

$$P(a) = P(x)_{V=1}, \tag{1.14}$$

$$P(b) = P(x)_{V=2}. \tag{1.15}$$

Note that in this case the joint probability distribution $P(a, b)$ is generally not well-defined, since it includes variables with different values for V.

1.4.2 Postulates of quantum mechanics

We will present a limited version of the postulates of quantum mechanics, since only projective measurements on finite-dimensional systems of observables with non-degenerate eigenvalues are considered in this thesis. This version of the postulates of Von Neumann [51] is based on [37] and [40].

- State postulate. For each physical system there is a corresponding Hilbert space $\mathcal{H}$. The state of a system is completely described by a unit vector in $\mathcal{H}$. A composite physical system corresponds to the tensor product of the Hilbert spaces of its constituent systems.

- Schrödinger postulate. If no measurement is being performed upon a system, the evolution of the state of the system in time is described by a unitary transformation:

$$|\psi(t)\rangle = U(t - t_0)|\psi(t_0)\rangle. \tag{1.16}$$

- Measurement postulate. A projective measurement on a system in the state $|\psi\rangle \in \mathcal{H}$ corresponds to a Hermitian operator A on $\mathcal{H}$. The possible outcomes are the eigenvalues $\{a_i\}$ of A. The probability of obtaining the outcome a_i is

$$|\langle\psi|i\rangle|^2, \tag{1.17}$$

  where $|i\rangle$ is a normalized eigenvector corresponding to the eigenvalue a_i. We will also talk about 'measuring a system S in the basis $|i\rangle_S$,' which means measuring an observable with eigenvectors $|i\rangle_S$.

- Projection postulate. If a measurement is performed on a physical system in the state $|\psi\rangle$ of an observable A, and the outcome is a_i, then directly after the measurement the system is in the eigenstate

$$|\psi\rangle \rightsquigarrow \frac{|i\rangle\langle i|\psi\rangle}{\big|\big||i\rangle\langle i|\psi\rangle\big|\big|} = |i\rangle. \tag{1.18}$$

1.4.3 Other language and notation

Aside from using P for probability distributions, we will also write $P(|\phi\rangle)$ for the one-dimensional projector $|\phi\rangle\langle\phi|$. Since the argument of this P is always a vector in the Dirac notation, no confusion will arise between projectors and probability distributions.

When we write down the state of a system, we include the label of the system as a subscript. Because of this, we can change the order of taking tensor products. For example, if system A is in the state $|i\rangle_A$ and system B is in the state $|j\rangle_B$, we have

$$|i\rangle_A \otimes |j\rangle_B \equiv |j\rangle_B \otimes |i\rangle_A. \tag{1.19}$$

We will also use the subscripts when writing down unitary operators to make clear on what part of Hilbert space the operator acts. Because of this we can write down expressions like

$$\left(U_{AA'} \otimes U_{BB'}\right)\left(|\psi\rangle_{AB} \otimes |\phi\rangle_{A'B'}\right). \tag{1.20}$$

We also often omit the symbol '$\otimes$' when taking tensor products:

$$|i\rangle_A \otimes |j\rangle_B = |i\rangle_A|j\rangle_B. \tag{1.21}$$

For definitions, we use ':=' when the object to be defined is at the left, and '=:' when it is at the right. For example,

$$\begin{aligned} f(x) &:= 5x; \\ N^2 - \sin(\pi/N) &=: J_N. \end{aligned} \tag{1.22}$$

Chapter 2

Pedagogical prelude: using chained Bell inequalities

This section is meant to be a pedagogical introduction into the subject. The aim is to provide a conceptually clear story which makes it easier to understand the rest of this thesis. In this section, I will show how the so-called 'chained Bell inequalities' lead to a strategy which rules out a bigger class of hidden variable theories than just local hidden variable theories. Large parts and notation of this section are inspired from [13] and [20].

2.1 Hidden variable theories

First we clarify what we mean with a hidden variable theory. A 'hidden variable theory' is a theory which assigns, in addition to the quantum mechanical state, supplemental variables to a system. These supplemental variables may give more information about the measurement outcomes. For example, if quantum mechanics predicts the probability $1/2$ for an outcome of a certain measurement, then, given the supplemental variables, this probability might be 1 or 0. The supplemental variables are called 'hidden' because they are not part of quantum mechanics, so from the viewpoint of this theory they are hidden. We do not make a commitment on the accessibility or the controllability of the hidden variables. If indeed it would turn out that there is a hidden variable theory underlying quantum mechanics, it is conceivable that the additional variables can be known, and are not hidden anymore. Nor do we assume realism or determinism when talking about hidden variable theories, unless this is clearly specified.

A hidden variable is said to be *compatible* with quantum mechanics or *QM-compatible* if, when averaging over the supplemental variables, the predictions are the same as the predictions of quantum mechanics. When discussing hidden variable theories the usual strategy is to impose constraints on such theories and see if hidden variable theories satisfying such constraints can still be compatible with quantum mechanics.

A hidden variable is said to be *trivial* if the supplemental variables are uncorrelated with the measurement outcomes, and therefore, given those variables, the probabilities of measurement outcomes are the same as in quantum mechan-

ics. In such a theory the supplemental variables might as well be thrown away so that we are left with ordinary quantum mechanics.

Throughout this thesis, when we write down probability distributions for outcomes of measurements, we take the quantum state of the system to be measured as a given, also when discussing hidden variable theories. So, the hidden variable is not a replacement for the quantum state, but an addition to it.

2.2 Measurements on a Bell state

In the context of hidden variable theories usually the EPRB (EPR-Bohm) thought experiment is discussed [9], which is a modified version of the measurement described by EPR (Section 1.1.1). Instead of measuring the position or momentum of a particle, the spin of electrons is measured. Since this experiment is hard to perform in practice, the first real measurements of this type have been measurements of the polarization of photons [16].

In the field of quantum computation, quantum mechanical systems are usually discussed without referring to how such a system could be implemented in the real world. In the context of Bell inequalities this means that there is no talk of electron spins or photon polarizations, but just of qubits. We will follow this practice and talk about measurements on qubits without specifying a specific implementation. The measurements we describe below are of the EPR type, so we can keep in mind they can be implemented using photons.

We will just focus on a simple system. This system consists of subsystems A and B. Each subsystem is a qubit: a two-level system associated with a two-dimensional Hilbert space. Let $\{|0\rangle_A, |1\rangle_A\}$ and $\{|0\rangle_B, |1\rangle_B\}$ be bases of the Hilbert spaces associated with qubits A and B, respectively. We will just focus on a single state of this system:

$$|\Psi\rangle_{AB} = \frac{1}{\sqrt{2}} \left(|0\rangle_A|0\rangle_B + |1\rangle_A|1\rangle_B\right). \tag{2.1}$$

This state is called a *Bell state* [40]. Note that it is entangled: it cannot be written as a single tensor product of states of A and B.

We will discuss projective measurements performed on the individual qubits. A projective measurement is characterized by an observable (for details, see Section 1.4.2). In order to define the relevant observables, we first define the qubit states

$$|\theta\rangle = \cos\frac{\theta}{2}|0\rangle + \sin\frac{\theta}{2}|1\rangle. \tag{2.2}$$

for both qubits. On the Bloch sphere (see Appendix A.3), the state $|\theta\rangle$ is obtained by rotating the unit vector in the z-direction by an angle θ around the y-axis, in the direction of the positive x-axis.

Now, the observables we consider are

$$O_\theta = -1 \cdot |\theta\rangle\langle\theta| + 1 \cdot |\theta + \pi\rangle\langle\theta + \pi|. \tag{2.3}$$

This means that measuring the observable O_v on a qubit in the state $|v\rangle$ results in the outcome -1 with certainty, while measuring the same observable on a

qubit in the state $|v+\pi\rangle$, which is orthogonal to $|v\rangle$, results in the outcome $+1$ with certainty.

For qubit A, we associate the *measurement setting* v with the observable O_v. Similarly, for B we associate w with O_w. The measurements settings V and W are *context* variables, not *random* variables. Now, we consider a joint measurement of $|\Psi\rangle_{AB}$ which consists of two local measurements of A and B with settings v and w. Associating random variables X and Y with the outcome of the measurement on A and B respectively, we can write down the quantum mechanical prediction for the outcome of this joint measurement, and for the individual measurements. The calculation is done in Section 2.7.

$$P_{QM}(x,y)_{v,w} = \frac{1}{4}(1 + x \cdot y \cdot \cos(v-w));$$
$$P_{QM}(x)_v = P_{QM}(y)_w = \frac{1}{2}. \tag{2.4}$$

As discussed in Section 1.4.1, we can also associate random variables with measurement outcomes for fixed settings. In this way, we define the following random variables

$$A = X\big|_{V=0}; \qquad B = Y\big|_{W=\pi/4};$$
$$A' = X\big|_{V=2\pi/4}; \qquad B' = Y\big|_{W=3\pi/4}. \tag{2.5}$$

Now we have, for example,

$$P_{QM}(a,b) = P_{QM}(x,y)_{V=0,W=\pi/4} = \frac{1}{2} + \frac{a \cdot b}{2\sqrt{2}}. \tag{2.6}$$

.

2.3 Alternative derivation of the CHSH inequalities

Let us now look at hidden variable theories. For simplicity we consider theories which provide, for the measurement on $|\Psi\rangle_{AB}$ (see (2.1)), a single random hidden variable Z which can take a finite amount of possible values. It is assumed that given the value of the hidden variable Z and the measurement settings V and W there is a well-defined joint probability distribution of the outcomes

$$P(x,y|z)_{v,w}. \tag{2.7}$$

We also assume that the hidden variable is independent of the measurement settings. This is called Source Independence (SI):

$$P(z)_{v,w} = P(z). \tag{2.8}$$

As noted in Section 2.1, given a particular value for the variable Z the distribution over measurement outcomes may differ from the quantum mechanical distribution, but when averaging over Z, the quantum mechanical prediction should be recovered, if the hidden variable theory is to be compatible with quantum mechanics:

$$\sum_z P(z)P(x,y|z)_{v,w} = P(x,y)_{v,w} = P_{QM}(x,y)_{v,w}. \tag{2.9}$$

Let us now turn to *local* hidden variable theories for this measurement. Following Shimony [47] we define such theories as theories obeying the following two conditions:

- Parameter Independence[1]:

$$\begin{aligned} \forall w', w'' : P(x|z)_{v,w'} = P(x|z)_{v,w''} =: P(x|z)_v; \\ \forall v', v'' : P(y|z)_{v',w} = P(y|z)_{v'',w} =: P(y|z)_w. \end{aligned} \tag{2.10}$$

- Outcome Independence:

$$\begin{aligned} P(x|y,z)_{v,w} = P(x|z)_{v,w}; \\ P(y|x,z)_{v,w} = P(y|z)_{v,w}. \end{aligned} \tag{2.11}$$

From these two conditions we can derive

$$\begin{aligned} P(x,y|z)_{v,w} &= P(x|y,z)_{v,w} P(y|z)_{v,w} = P(x|z)_{v,w} P(y|z)_{v,w} \\ &= P(x|z)_v P(y|z)_w, \end{aligned} \tag{2.12}$$

which is called *Factorizability* [46]. Factorizability is similar to what Bell has called 'local causality' [6].

A violation of Parameter Independence means that the probabilities of measurement outcomes at one subsystem depend on a freely chosen setting at the other subsystem. This means that there is a casual influence from the setting to the probabilities. However, the settings and measurements can be spacelike separated, and therefore this would imply faster-than-light causation. Usually this is regarded as being in conflict with the special theory of relativity.

A violation of Outcome Independence can be a sign that the supplemental variables do not give a complete description. If the outcomes are causally unrelated then, according to Reichenbach's common cause principle [43] any correlation between them can be screened of by a common cause. In this case this would mean there are additional hidden variables such that Outcome Independence holds. However, the use of Reichenbach's common cause principle for these type of correlations is controversial [1].

In terms of the random variables A, A', B, B' defined in (2.5), Factorizability means

$$\begin{aligned} P(a,b|z) &= P(a|z)P(b|z); \\ P(a,b'|z) &= P(a|z)P(b'|z); \\ P(a',b|z) &= P(a'|z)P(b|z); \\ P(a',b'|z) &= P(a'|z)P(b'|z). \end{aligned} \tag{2.13}$$

If a hidden variable theory obeys the Factorizability condition, then the joint probability distributions $P(a,b)$, $P(a',b)$, $P(a,b')$ and $P(a',b')$ can be derived from a common probability distribution $P(a,a',b,b')$ [29] if we define the common distribution as follows:

$$P(a,a',b,b') := \sum_z P(z)P(a|z)P(a'|z)P(b|z)P(b'|z). \tag{2.14}$$

[1] At the beginning of Chapter 3 we will slightly extend the definition of Parameter Independence.

Each of the four distributions can now be obtained by summing $P(a, a', b, b')$ over a/a' and b/b'. This imposes constraints on these distributions. We will now derive some of these constraints.

2.3.1 Inequalities from joint distributions

We prove a simple relation involving three random variables A, B, C which share a common probability distribution, which is similar to a triangle inequality:

$$P(A \neq C) \leq P(A \neq B) + P(B \neq C). \tag{2.15}$$

The proof goes as follows. We assume that there is a common probability distribution $P(A, B, C)$.

$$\begin{aligned} P(A \neq B) &= P(A \neq B, A \neq C, B \neq C) \\ &\quad + P(A \neq B, A = C, B \neq C) \\ &\quad + P(A \neq B, A \neq C, B = C); \qquad (2.16) \\ P(B \neq C) &= P(A \neq B, A \neq C, B \neq C) \\ &\quad + P(A \neq B, A = C, B \neq C) \\ &\quad + P(A = B, A \neq C, B \neq C); \qquad (2.17) \\ P(A \neq C) &= P(A \neq B, A \neq C, B \neq C) \\ &\quad + P(A = B, A \neq C, B \neq C) \\ &\quad + P(A \neq B, A \neq C, B = C). \qquad (2.18) \end{aligned}$$

Adding (2.16) and (2.17),

$$\begin{aligned} P(A \neq B) + P(B \neq C) &= 2(P(A \neq B, A \neq C, B \neq C) + P(A \neq B, A = C, B \neq C)) \\ &\quad + P(A \neq B, A \neq C, B = C) + P(A \neq B, A \neq C, B = C) \\ &= P(A \neq C) + P(A \neq B, A \neq C, B \neq C) \\ &\quad + 2P(A \neq B, A = C, B \neq C) \\ &\geq P(A \neq C). \square \end{aligned} \tag{2.19}$$

We can now apply this inequality to the random variables A, A', B and B', which take values ± 1. Because any three of these share a common probability distribution we have

$$\begin{aligned} P(A \neq B') &\leq P(A \neq B) + P(B \neq B'), \\ P(B \neq B') &\leq P(B \neq A') + P(A' \neq B'). \end{aligned} \tag{2.20}$$

Combining these,

$$P(A \neq B') \leq P(A \neq B) + P(B \neq A') + P(A' \neq B'). \tag{2.21}$$

Instead of using probabilities like $P(A \neq B')$ we can use the expectation value of the product of the variables:

$$\begin{aligned} E(A, B) &:= \sum_{a,b \in \{-1,+1\}} a \cdot b \cdot P(a, b) \\ &= P(A = B) - P(A \neq B) \\ &= 1 - 2P(A \neq B) \end{aligned} \tag{2.22}$$

Rewriting (2.21), we then get

$$E(A,B) + E(A',B) - E(A,B') + E(A',B') \leq 2, \tag{2.23}$$

which is the well-known CHSH inequality [16]. We see that the CHSH inequality is equivalent to an inequality about random variables which share a common probability distribution.

Using the definitions (2.5), quantum mechanics yields $2\sqrt{2}$ for the left-hand side, hence it violates this inequality. Since any local deterministic theory obeys Factorizability, there is no such theory compatible with quantum mechanics.

It is also known that the settings for the random variables A, A', B, B' we have chosen in (2.5) give the *maximum* violation of this inequality. In other words, there are no settings which give a bigger value for the left-hand side of (2.23) than $2\sqrt{2}$ [15]. Therefore, in the context of discussions of the Bell inequalities, usually the focus is on these measurement settings, where the difference between the settings are multiples of $\pi/4$.

However, let us go back to the representation of the inequality using probabilities:

$$P(A \neq B') \leq P(A \neq B) + P(B \neq A') + P(A' \neq B'). \tag{2.24}$$

Evaluating this expression using the quantum mechanical prediction, we get approximately

$$0.85 \leq 0.44. \tag{2.25}$$

What would now be a good way to quantify the amount of violation? We define this, admittedly rather arbitrary, to be the ratio between the left-hand side and the right-hand side. A violation means that this ratio is greater than 1. In the above case, it equals approximately 1.94. Can we find settings which make this ratio larger? Instead of the values of V and W used in the definition of A, A', B, B' (see (2.5)), let us consider other settings with equal angular separation. We define

$$\begin{aligned} A &= X\big|_{V=0}; & B &= Y\big|_{W=\theta}; \\ A' &= X\big|_{V=2\theta}; & B' &= Y\big|_{W=3\theta}. \end{aligned} \tag{2.26}$$

Then the inequality becomes

$$\sin^2 \frac{3\theta}{2} \leq 3 \sin^2 \frac{\theta}{2}. \tag{2.27}$$

The ratio becomes larger in the limit of θ going to zero. For small θ we have $\sin\theta \approx \theta$, so then we approximately get

$$\frac{9\theta^2}{4} \leq \frac{3\theta^2}{4}, \tag{2.28}$$

which gives a ratio 3. This suggests that it might be interesting to look at a small separation of settings.

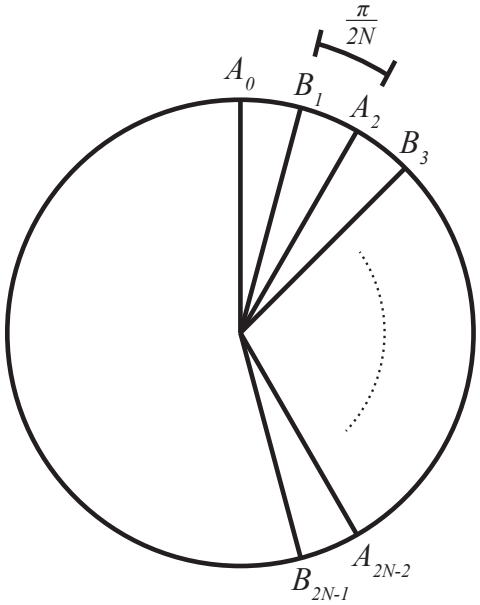

Figure 2.1: Graphical representation of the settings corresponding to the random variables A_i and B_j defined in (2.29).

2.4 Chaining CHSH inequalities

To further explore violations at such small separations, we are going to 'chain' the CHSH inequality. This was first done by Pearle [41], and more extensively by Braunstein and Caves [13]. Let N be a positive integer. We define the following random variables:

$$\begin{aligned} A_i &= X\big|_{v=i\cdot\frac{\pi}{2N}} & i &\in \{0,2,\ldots,2N-2\} \\ B_j &= Y\big|_{w=j\cdot\frac{\pi}{2N}} & j &\in \{1,3,\ldots,2N-1\} \end{aligned} \tag{2.29}$$

Interpreting the settings as angles, they can be graphically represented on a unit circle, as in Figure 2.1. Note that the separation between the settings becomes smaller as N becomes larger.

Because of the way the random variables $\{A_i\}$ and $\{B_j\}$ are defined, the joint distribution of a pair $\{A_k, B_l\}$ is automatically well-defined:

$$P(a_k, b_l) = P(x,y)_{V=k\cdot\frac{\pi}{2N}, W=l\cdot\frac{\pi}{2N}}. \tag{2.30}$$

Furthermore, if Factorizability holds, then we can define a joint distribution of

all the $\{A_i\}$ and $\{B_j\}$ from which any distribution $P(a_k, b_l)$ can be derived:

$$P(a_0, b_1, a_2, \ldots, a_{2n-2}, b_{2n-1}) := \sum_z P(z)P(a_0|z)P(b_1|z)\ldots P(b_{2n-1}|z); \tag{2.31}$$

$$P(a_k, b_l) = \sum_{\{a_i|i\neq k\}} \sum_{\{b_j|j\neq l\}} P(a_0, b_1, \ldots, b_{2n-1}). \tag{2.32}$$

Now consider the following inequality

$$1 = P(A_0 \neq \overline{A_0}) \leq P(A_0 \neq B_1) + P(B_1 \neq \overline{A_0}). \tag{2.33}$$

Here $\overline{A_0} = -A_0$, as defined in (1.5). This inequality is satisfied trivially, since by definition it says

$$1 \leq P(A_0 \neq B_1) + (1 - P(A_0 \neq B_1)) = 1. \tag{2.34}$$

Now we consider the following CHSH inequalities:

$$\begin{aligned}
P(B_1 \neq \overline{A_0}) &\leq P(B_1 \neq A_2) + P(A_2 \neq B_3) + P(B_3 \neq \overline{A_0}) \\
P(B_3 \neq \overline{A_0}) &\leq P(B_3 \neq A_4) + P(A_4 \neq B_5) + P(B_5 \neq \overline{A_0}) \\
&\vdots \\
P(B_{2N-3} \neq \overline{A_0}) &\leq P(B_{2N-3} \neq A_{2N-2}) + P(A_{2N-2} \neq B_{2N-1}) \\
&\quad + P(B_{2N-1} \neq \overline{A_0}).
\end{aligned} \tag{2.35}$$

By substituting each inequality into the inequality directly above it, and the first one into (2.33), we get the chained inequality

$$1 \leq P(A_0 \neq B_1) + \ldots + P(B_{2N-1} \neq \overline{A_0}) =: I_N. \tag{2.36}$$

This is known as a 'chained Bell inequality'. Remember that the definition of the A_i and B_j depends on N. Defining I_N as the right-hand side of this inequality, we see that local hidden variable theories obey $1 \leq I_N$, since all the CHSH inequalities used are satisfied by such theories (remember that in this case there is a joint distribution (2.31) of all the A_i and B_j). However, quantum mechanics predicts

$$I_N = 2N \sin^2 \frac{\pi}{4N}, \tag{2.37}$$

(as calculated in Section 2.7. Noting that

$$\forall x > 0\text{: } \sin x < x, \tag{2.38}$$

we have

$$I_N < \frac{\pi^2}{8N}. \tag{2.39}$$

We see that I_N can be chosen to be arbitrarily close to 0, just by picking N large enough. Quantum mechanics violates the inequality (2.36) for $N \geq 2$. Looking

at the amount of violation, i.e. the ratio between the left-hand side and the right-hand side, we see that this goes to infinity as N goes to infinity. So, in a sense, we seem to have found an inequality, satisfied by local hidden variable theories, which in the limit $N \to \infty$ is grossly violated by quantum mechanics. Also note that due to the positivity of probabilities the algebraic minimum of the right-hand side is 0, and quantum mechanics comes arbitrarily close to that value.

We should however remember that the chained Bell inequality (2.36) is just built up from many CHSH inequalities. So a violation of the chained inequality just means that one or more standard CHSH inequalities are violated. In fact, in this case, all the CHSH inequalities (2.35) which are used to build up (2.36) are violated by quantum mechanics. So, the large violation of (2.36) can be seen as a simultaneous violation of many CHSH inequalities, each corresponding to different settings. But, of course, any hidden variable theory that obeys (2.24) for all settings also obeys (2.36). So (2.36) by itself can not be used to rule out more hidden variable theories than (2.24) can. However, we will see below that the quantum mechanical prediction for the right-hand side of (2.36), which is I_N, *can* do this.

2.5 Theories satisfying Parameter Independence

In the previous sections we considered local hidden variable theories, which satisfy both Parameter Independence and Outcome Independence. We saw that such theories satisfy Factorizability, and because of that a joint distribution of all the A_i and B_j could be defined. This in turn allowed the derivation of CHSH inequalities, which are violated by quantum mechanics.

We are now going to focus on a larger class of hidden variable theories: we drop the condition of Outcome Independence while still demanding Parameter Independence. Note that $P(x|z)_v$ is well-defined for all measurements by virtue of this assumption. To show this, suppose that (2.10) is not obeyed, i.e. there are v, w', w'' with $P(x|z)_{v,w'} \neq P(x|z)_{v,w''}$. The state $|\Psi\rangle_{AB}$ is prepared 200 times. Upon the first set of 100 states measurements are performed with settings v and w', upon the second set of 100 states measurements are performed with settings v and w''. In the context of the first set of measurements, we have $P(x|z)_v = P(x|z)_{v,w'}$. In the context of the second set of measurements, we have $P(x|z)_v = P(x|z)_{v,w''}$. This means that there is no consistent probability distribution $P(x|z)_v$ for the whole run of 200 experiments.

We again define random variables A_i and B_j according to (2.29). Note that we use Parameter Independence here: in the definition of the variables A_i no value of W needs to be specified, and in the definition of the variables B_j no value of V needs to be specified.

Again, the joint distribution of a pair $\{A_k, B_l\}$ is automatically well-defined. But, since Outcome Independence is not assumed, Factorizability generally does not hold. Therefore, we cannot define a joint distribution of all $\{A_i\}$ and $\{B_j\}$, and the CHSH inequalities cannot be derived. This is a good thing, because this time we are looking for hidden variable theories that are compatible with quantum mechanics, and such theories violate the CHSH inequalities. However, we are going to derive some other interesting inequalities.

Take for example $N = 64$. In this case, the settings corresponding to the

random variables $A_0, B_1, A_2, B_3, \ldots$ are separated by an angle $\pi/128$. Now consider the probabilities $P(A_0 = +1)$ and $P(B_1 = +1)$. Deriving an inequality for these probabilities:

$$\begin{aligned} |P(A_0 = +1) - P(B_1 = +1)| &= \left| \sum_{b_1} P(A_0 = +1, b_1) - \sum_{a_0} P(a_0, B_1 = +1) \right| \\ &= \big| P(A_0 = +1, B_1 = -1) - P(A_0 = -1, B_1 = +1) \big| \\ &\leq \big| P(A_0 = +1, B_1 = -1) \big| + \big| P(A_0 = -1, B_1 = +1) \big| \\ &= P(A_0 \neq B_1), \end{aligned} \tag{2.40}$$

where we have used the triangle inequality. The inequality (2.40) looks intuitive: the difference in probabilities is bounded from above by the probability that the random variables are different. Note that we use the fact that there is a common probability distribution for A_0 and B_1. Also note that this inequality is trivially satisfied in quantum mechanics (as it should be) because then

$$P(A_0 = +1) = P(B_1 = +1) = 1/2, \tag{2.41}$$

so the left-hand side equals zero, while the calculation in Section 2.7 gives

$$P(A_0 \neq B_1) = \sin^2(\pi/256) \approx 1.5 \cdot 10^{-4}. \tag{2.42}$$

But now consider the corresponding inequality for a hidden variable theory satisfying Parameter Independence. For a specific value z' of the hidden variable Z:

$$\big| P(A_0 = +1|z') - P(B_1 = +1|z') \big| \leq P(A_0 \neq B_1|z') \tag{2.43}$$

Of course, if we know nothing more about the hidden variable theory, we do not know what the values of these probabilities are. But, since we assume that there is a finite set of hidden variables, we know that $P(z')$ is nonzero. Therefore, we have

$$P(A_0 \neq B_1|z') = \frac{P(A_0 \neq B_1, z')}{P(z')} \leq \frac{P(A_0 \neq B_1)}{P(z')}. \tag{2.44}$$

where we used subadditivity (1.10). So, we have

$$\begin{aligned} &P(z') > P(A_0 \neq B_1) \Rightarrow \\ &\qquad 1 > P(A_0 \neq B_1|z') \geq \big| P(A_0 = +1|z') - P(B_1 = +1|z') \big|. \end{aligned} \tag{2.45}$$

And, since the calculation in Section 2.7 shows

$$P(A_0 \neq B_1) = \sin^2 \frac{\pi}{4N} < \frac{\pi^2}{16N^2}, \tag{2.46}$$

we can always pick an N big enough so that $P(z') > P(A_0 \neq B_1)$. So, for some N we have

$$\big| P(A_0 = +1|z') - P(B_1 = +1|z') \big| < 1. \tag{2.47}$$

This is a nontrivial constraint on some probabilities of the hidden variable theory. We see that the strong correlation at small angles puts some constraints on hidden variable theories satisfying Parameter Independence.

But that is not all. In a similar way to the derivation of the chained Bell inequality, we can find another constraint. We start with

$$\begin{aligned}\big|P(A_0=+1)-P(\overline{A_0}=+1)\big| \leq\ & \big|P(A_0=+1)-P(B_1=+1)\big| \\ & +\big|P(B_1=+1)-P(A_2=+1)\big| \\ & +\ldots \\ & +\big|P(B_{2N-1}=+1)-P(\overline{A_0}=+1)\big|. \end{aligned} \tag{2.48}$$

Here we have just used an ordinary triangle inequality. Now using (2.40), we have

$$\big|P(A_0=+1)-P(\overline{A_0}=+1)\big| \leq P(A_0\neq B_1)+\ldots+P(B_{2N-1}\neq\overline{A_0})=I_N. \tag{2.49}$$

Notice that the right-hand side equals the right-hand side of the chained Bell inequality (2.36). Again this is trivially satisfied by quantum mechanics, since the left-hand side vanishes. But, by repeating the steps with probabilities conditioned on z', the same inequality holds for hidden variable theories satisfying Parameter Independence:

$$\begin{aligned}&\big|P(A_0=+1|z')-P(\overline{A_0}=+1|z')\big| \leq \\ &P(A_0\neq B_1|z')+\ldots+P(B_{2N-1}\neq\overline{A_0}|z')=:I_N^{z'}. \end{aligned} \tag{2.50}$$

where we have defined $I_N^{z'}$ for convenience. Because of (2.44), we have $I_N^{z'} \leq I_N/P(z')$. But this means that also $I_N^{z'}$ vanishes in the limit $N\to\infty$. In other words, for every non-zero value on the left-hand side of (2.50), we can find an N such that the inequality is violated. Therefore, we must have

$$\begin{aligned}&\big|P(A_0=+1|z')-P(\overline{A_0}=+1|z')\big|=0 \\ &\Rightarrow P(A_0=+1|z')=P(\overline{A_0}=+1|z')=1-P(A_0=+1|z')=1/2 \end{aligned} \tag{2.51}$$

Also note that $P(A_0=-1|z')=1-P(A_0=+1|z')=1/2$. We see that the probability distribution of A_0 given z' is equal to the quantum mechanical probability distribution, and independent of z'. So, for this particular measurement the hidden variable theory gives no better prediction than quantum mechanics: it is trivial.

Using symmetry arguments it can be shown that this holds for every projective measurement on qubit A, as well as on qubit B. We will demonstrate this in Section 3.1. Note that up to this point, we have only talked about projective measurements on subsystems A and B. Hence, the above argument does not rule out that a hidden variable theory satisfying Parameter Independence gives different probability distributions for some projective measurement on the combined system $A+B$. However, this will be ruled out when we generalize the result in Chapter 3.

The main ingredient of this argument consists of the very strong correlation between measurement outcomes on both sides when the separation of settings

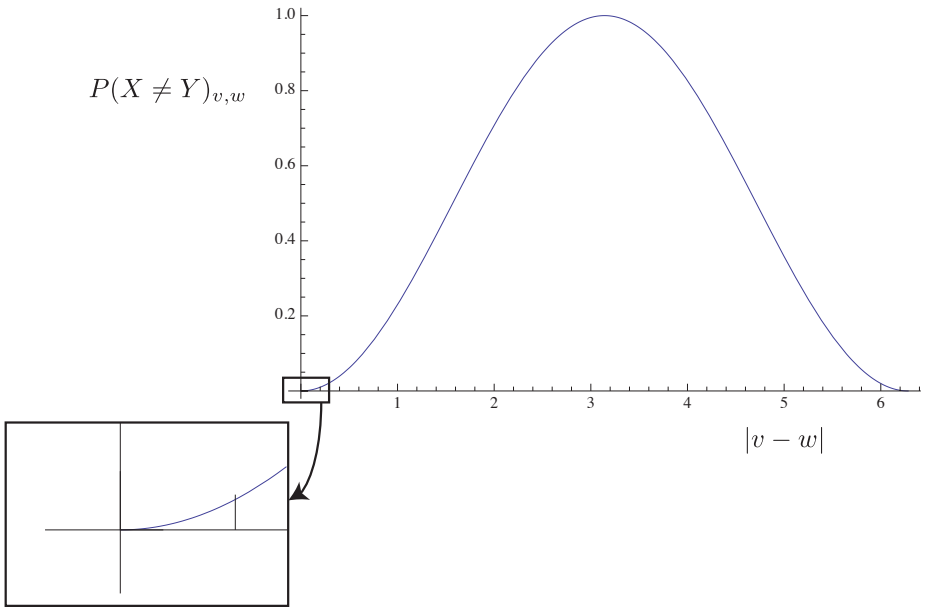

Figure 2.2: Strong correlation at small angles

is small. Each term of $I_N^{z'}$, which we defined in (2.50), is bounded from above by

$$\frac{1}{P(z')} \sin^2 \frac{\pi}{4N} \leq \frac{\pi^2}{16N^2 P(z')} \propto \frac{1}{N^2}, \tag{2.52}$$

and the number of terms is $2N$. Hence, the right-hand side has as upper bound proportional to $N/N^2 = 1/N$. It is the fact that $P(A_i \neq B_{i\pm1})$ is approximately quadratic in the angle of separation for small angles that indicates that there is a very strong correlation. This ensures that in the limit $N \to \infty$, $I_N^{z'}$ vanishes. See also Figures 2.2 and 2.3.

2.6 Generalization of the result

We will now give an idea how the above result can be extended from measurements on parts of a Bell state to arbitrary measurements on pure states. This generalization will be discussed in detail in Chapter 3.

Consider some entangled state of the system $AB = A + B$. According to Schmidt's biorthogonal decomposition theorem (see Appendix A.1), this state can be expressed as

$$|\phi\rangle_{AB} = \sum_{i=0}^{d-1} c_i |i\rangle_A |i\rangle_B \tag{2.53}$$

with c_i nonnegative, $|i\rangle_A$ and $|i\rangle_B$ orthonormal bases, and d the Schmidt order of the state. Now suppose two of the coefficients are equal (and nonzero). Without

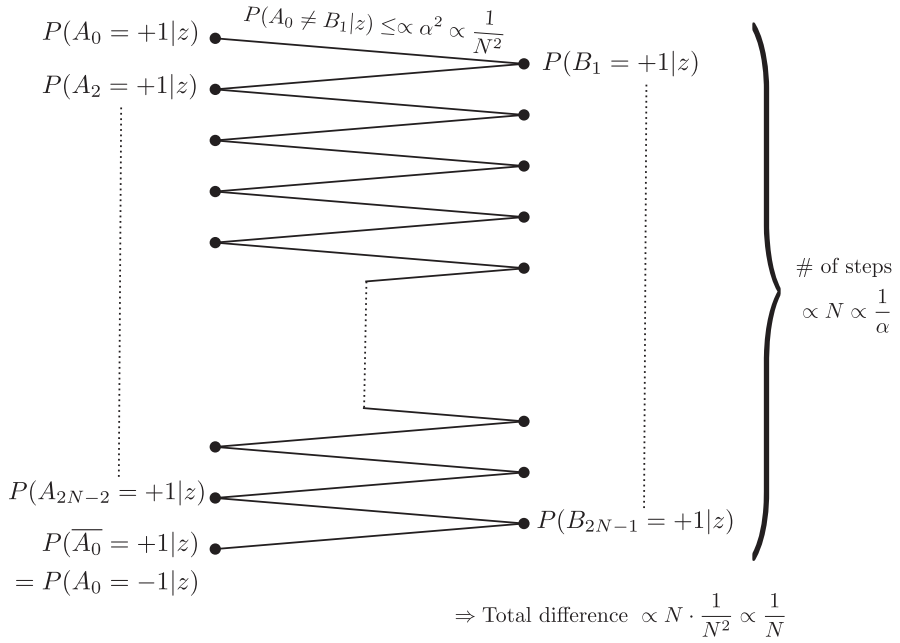

Figure 2.3: Limiting differences between probabilities using the probability of being different. α equals the separation between measurement settings $|v - w|$. We see that by taking N steps where the upper bound of the difference at each step is proportional to $1/N^2$, the total difference is proportional to $1/N$. Since this holds for every $N \in \mathbb{N}$, the total difference, which is the difference between $P(A_0 = +1|z)$ and $P(A_0 = -1|z)$, must be zero.

loss of generality we can assume $c_0 = c_1$, since the ordering of the basis vectors is arbitrary. We consider measurements on A and B in the bases $|i\rangle_A$ and $|i\rangle_B$, respectively. We also consider measurements on A and B in the bases where the first two elements $|0\rangle$ and $|1\rangle$ are 'rotated' as follows:

$$\begin{aligned} |0\rangle &\mapsto \cos\frac{\theta}{2}|0\rangle + \sin\frac{\theta}{2}|1\rangle; \\ |1\rangle &\mapsto -\sin\frac{\theta}{2}|0\rangle + \cos\frac{\theta}{2}|1\rangle. \end{aligned} \tag{2.54}$$

Now we restrict our attention to the cases that the outcomes are both eigenvalues corresponding to the two eigenstates (2.54) of the observable; if this is not the case, we neglect the measurement outcomes. Then, the correlations between the outcomes are the same as the correlations of the Bell state we have discussed in Section 2.5. We can then use the same argument to show that the corresponding probabilities must be equal. The following follows:

Theorem 2. *Consider the state of a system $A+B$, expressed in Schmidt bases of A and B:*

$$|\Phi\rangle_{AB} = \sum_i c_i |i\rangle_A |i\rangle_B. \tag{2.55}$$

Consider a projective measurement on subsystem A, in the basis $|i\rangle_A$. In any QM-compatible hidden variable theory satisfying Parameter Independence, the probabilities of the outcomes are equal to the probabilities predicted by quantum mechanics and independent of any supplemental variables.

This will be proven in Section 3.1.

For the next generalization, which will be discussed in Section 3.2, we again consider measurements on A and B in the bases $|i\rangle_A$ and $|i\rangle_B$. Using a set of special states, called 'embezzling' states, it is possible to transform this state to a state which has all Schmidt coefficients (almost) equal. We can then apply the above result to show that all probabilities are (almost) equal. Going back to the original state with unequal Schmidt coefficients, we can then show that the probabilities are equal to those given by quantum mechanics. The result is:

Theorem 3. *Consider the state of a system $A+B$, expressed in Schmidt bases of A and B:*

$$|\Phi\rangle_{AB} = \sum_i c_i |i\rangle_A |i\rangle_B. \tag{2.56}$$

Consider a projective measurement on subsystem A, in the basis $|i\rangle_A$. In any QM-compatible hidden variable theory satisfying Parameter Independence, if two Schmidt coefficients are equal, then the probabilities corresponding to the corresponding Schmidt basis vectors are also equal.

Finally, the measurement scheme of John von Neumann is invoked. In an ideal projective measurement, at some point the system to be measured S becomes entangled with (a part of) the measurement apparatus D. The measurement of S can then at the same time be interpreted as a measurement of D. But this is a measurement in the Schmidt basis of S+D. Therefore, we can apply the previous result, and conclude the following:

Theorem 4. *Consider a projective measurement performed on a pure quantum state. In any QM-compatible hidden variable theory satisfying Parameter Independence, the probabilities of the outcomes are independent of any supplemental variables. Therefore, any such theory is trivial.*

This will be shown in Section 3.4.

In this chapter we have seen that strong correlations at small angular separations rule out certain hidden variable theories satisfying Parameter Independence. In the next chapter we will extensively show and discuss the generalizations just mentioned.

2.7 Calculation of $P(x,y)_{v,w}$ and I_N

First, we calculate $P(x,y)_{v,w}$. The correspondence between eigenvalues and eigenvectors is

$$\begin{aligned} -1 &\leftrightarrow |v\rangle/|w\rangle, \\ +1 &\leftrightarrow |v+\pi\rangle/|w+\pi\rangle. \end{aligned} \tag{2.57}$$

This can be written as

$$\begin{aligned} x &\leftrightarrow \left| v + \frac{x+1}{2}\pi \right\rangle, \\ y &\leftrightarrow \left| w + \frac{y+1}{2}\pi \right\rangle. \end{aligned} \tag{2.58}$$

Defining

$$\begin{aligned} v' &= \frac{v}{2} + \frac{x+1}{4}\pi, \\ w' &= \frac{w}{2} + \frac{y+1}{4}\pi, \end{aligned} \tag{2.59}$$

we have

$$\begin{aligned} &P(x,y)_{v,w} = |_{AB}\langle\phi|\left(|2v'\rangle_A \otimes |2w'\rangle_B\right)|^2 \\ &= \left| \frac{1}{\sqrt{2}} \left({}_A\langle 0|_B\langle 0| + {}_A\langle 1|_B\langle 1|\right)\left(\cos v'|0\rangle_A + \sin v'|1\rangle_A\right) \otimes \left(\cos w'|0\rangle_B + \sin w'|1\rangle_B\right) \right|^2 \\ &= \frac{1}{2}\left(\cos v' \cos w' + \sin v' \sin w'\right)^2 = \frac{1}{2}\left(\cos v' \cos -w' - \sin v' \sin -w'\right)^2 \\ &= \frac{1}{2}\cos^2\left(\frac{v-w}{2} + \frac{x-y}{4}\pi\right) = \begin{cases} \frac{1}{2}\cos^2\left(\frac{v-w}{2}\right) & x = y \\ \frac{1}{2}\sin^2\left(\frac{v-w}{2}\right) & x \neq y \end{cases} \\ &= \frac{1}{4}\left(1 + x \cdot y \cdot \cos\left(v-w\right)\right), \end{aligned} \tag{2.60}$$

where in the last line we have used

$$\cos^2\theta = \frac{1+\cos 2\theta}{2}; \quad \sin^2\theta = \frac{1-\cos 2\theta}{2}. \tag{2.61}$$

I_N is defined as

$$I_N := P(A_0 \neq B_1) + P(B_1 \neq A_2) + \ldots + P(B_{2N-1} \neq \overline{A_0}). \tag{2.62}$$

Except for the last term, each term equals

$$P(A_i \neq B_{i\pm1}) = P(X \neq Y)_{v=i\frac{\pi}{2N}, w=(i\pm1)\frac{\pi}{2N}} = \sin^2 \frac{\pi}{4N}, \tag{2.63}$$

while the last term equals

$$\begin{aligned} P(B_{2N-1} \neq -A_0) &= P(B_{2N-1} = A_0) = P(X = Y)_{v=0, w=(2N-1)\frac{\pi}{2N}} \\ &= \cos^2\left(\frac{\pi}{2} - \frac{\pi}{4N}\right) = \sin^2 \frac{\pi}{4N}. \end{aligned} \tag{2.64}$$

Since I_N has $2N$ terms, we have

$$I_N = 2N \sin^2 \frac{\pi}{4N} < \frac{\pi}{8N}, \tag{2.65}$$

since $\sin x < x$ for $x > 0$.

Chapter 3

Reformulation of the Colbeck-Renner result

In this chapter we derive a modified version of the Colbeck-Renner result. While many steps are similar to those taken by CR, there are also differences, mainly in the assumptions that are made. A discussion of these differences is the subject of Chapter 4. In this chapter, we will not refer to CR's work and just present our formulation of the result. In this formulation the result is a no-go theorem for hidden variable theories satisfying Parameter Independence.

Throughout this chapter, Source Independence is assumed, so that the same probability distribution $P(z)$ can be used for the hidden variable, independent of the measurement settings.

Furthermore, we extend the definition of Parameter Independence slightly. Let $A + B$ be a composite system, where the systems A and B do not interact. In (2.10) we defined Parameter Independence as stating that probabilities at A are independent of settings at B and vice versa. We now also assume that probabilities at A are independent of any transformation performed at B. Just like with the initial definition, this can be motivated by the fact that a violation of this allows superluminal signaling if the transformation at B and the measurement at A are spacelike separated events.

3.1 Compact rederivation of the triviality result for Bell states

In Chapter 2 we explained at length how strong correlations at small separations of settings rule out more hidden variable theories than the standard Bell inequality does. In this section we will present the result completely and compactly. That is, we will prove the following:

Theorem 1. *Consider the following Bell state:*

$$|\Psi\rangle_{AB} = \frac{1}{\sqrt{2}} \left(|0'\rangle_A|0'\rangle_B + |1'\rangle_A|1'\rangle_B\right). \tag{3.1}$$

Consider a projective measurement on subsystem A. In any QM-compatible hidden variable theory satisfying Parameter Independence, the probabilities of the

outcomes are $1/2$*, i.e. equal to the probabilities predicted by quantum mechanics and independent of any hidden variables.*

Defining X as the outcome of the measurement and Z as the hidden variable, we have to show that

$$\forall x, z : P(x|z) = P(x) = \frac{1}{2}. \tag{3.2}$$

We define the basis in which the measurement on A is performed as $\{|0\rangle_A, |1\rangle_A\}$, and we label the corresponding outcomes $-1, +1$. Then, the measurement is represented by the observable

$$O = -1 \cdot |0\rangle\langle 0| + 1 \cdot |1\rangle\langle 1|. \tag{3.3}$$

According to the symmetry property for this Bell state (Section A.2), there is a basis $\{|0\rangle_B, |1\rangle_B\}$ of $\mathcal{H}_B$ such that

$$|\Psi\rangle_{AB} = \frac{1}{\sqrt{2}}(|0\rangle_A|0\rangle_B + |1\rangle_A|1\rangle_B). \tag{3.4}$$

We see that, when expressed in the measurement basis $\{|0\rangle_A, |1\rangle_A\}$, $|\Psi\rangle_{AB}$ has the same form as in (3.1).

To prove (3.2), we have to show that for any z

$$P(X = +1|z) = P(X = -1|z) = \frac{1}{2}. \tag{3.5}$$

We start by considering different measurements on A and B. First, we define

$$|\theta\rangle = \cos\frac{\theta}{2}|0\rangle + \sin\frac{\theta}{2}|1\rangle \tag{3.6}$$

and the observables

$$O_\theta = -1 \cdot |\theta\rangle\langle\theta| + 1 \cdot |\theta + \pi\rangle\langle\theta + \pi|. \tag{3.7}$$

For qubit A, we associate the *measurement setting* v with the observable O_v. Similarly, for B we associate w with O_w. Note that the original measurement on A of the observable O (see (3.3)) is equal to a measurement on A of the observable $O_{V=0}$.

Let $N \in \mathbb{N}$. We define the following random variables:

$$A_i = X\big|_{V=i\cdot\frac{\pi}{2N}}, \qquad i \in \{0, 2, \ldots, 2N-2\}; \tag{3.8}$$

$$B_j = Y\big|_{W=j\cdot\frac{\pi}{2N}}, \qquad j \in \{1, 3, \ldots, 2N-1\}. \tag{3.9}$$

We now prove a lemma which states that the difference between the probabilities of two random variables taking some value w is bounded from above by the probability that the random variables take different values.

Lemma 1. *Let X and Y be random variables with joint probability distribution $P(x, y)$, and let w be a possible value of both X and Y. Then the following inequality holds:*

$$|P(X = w) - P(Y = w)| \leq P(X \neq Y). \tag{3.10}$$

Proof:

$$
\begin{aligned}
|P(X=w) - P(Y=w)| &= \left| \sum_y P(X=w,y) - \sum_x P(x,Y=w) \right| \\
&= \left| \sum_{y \neq w} P(X=w,y) - \sum_{x \neq w} P(x,Y=w) \right| \\
&\leq \sum_{y \neq w} P(X=w,y) + \sum_{x \neq w} P(x,Y=w) \\
&\leq P(X \neq Y). \square
\end{aligned} \tag{3.11}
$$

We now consider probabilities in the hidden variable theory. We have, for any z,

$$
\begin{aligned}
|P(X=+1|z) - P(X=-1|z)| &= |P(X=+1|z) - P(\overline{X}=+1|z)| \\
&= |P(A_0=+1|z) - P(\overline{A_0}=+1|z)| \\
&\leq |P(A_0=+1|z) - P(B_1=+1|z)| \\
+|P(B_1=+1|z) - P(A_2=+1|z)| + \ldots + |P(B_{2N-1}=+1|z) - P(\overline{A_0}=+1|z)|.
\end{aligned} \tag{3.12}
$$

Now, applying Lemma 1 we have

$$
\begin{aligned}
&|P(X=+1|z) - P(X=-1|z)| \\
&\qquad \leq P(A_0 \neq B_1|z) + \ldots + P(B_{2N-1} \neq \overline{A_0}|z) =: I_N^z.
\end{aligned} \tag{3.13}
$$

Averaging both sides over Z, we have

$$
\begin{aligned}
&\sum_z P(z)|P(X=+1|z) - P(X=-1|z)| \\
&\qquad \leq P(A_0 \neq B_1) + P(B_1 \neq A_2) + \ldots + P(B_{2N-1} \neq -A_0) \\
&\qquad = I_N < \frac{\pi^2}{8N}.
\end{aligned} \tag{3.14}
$$

Implicitly we have used Source Independence here, since the same distribution $P(z)$ is used for each term on the right-hand side of the inequality, independently of the settings V and W. The calculation of I_N is done in Section 2.7. Since (3.14) holds for every N, we get

$$
\sum_z P(z)|P(X=+1|z) - P(X=-1|z)| = 0. \tag{3.15}
$$

For every $z \in \tilde{Z}$, $P(z) > 0$, and therefore

$$
\begin{aligned}
&|P(X=+1|z) - P(X=-1|z)| = 0 \\
&\Rightarrow P(X=+1|z) = P(X=-1|z) = \frac{1}{2}. \square
\end{aligned} \tag{3.16}
$$

3.2 Extension to states with equal Schmidt coefficients

We now extend Theorem 1 to probabilities of measurement outcomes corresponding to equal Schmidt coefficients:

Theorem 2. *Consider the state of a system $A+B$, expressed in Schmidt bases of A and B:*

$$|\Phi\rangle_{AB} = \sum_i c_i |i\rangle_A |i\rangle_B. \tag{3.17}$$

Consider a projective measurement on subsystem A, in the basis $|i\rangle_A$. In any QM-compatible hidden variable theory satisfying Parameter Independence, if two Schmidt coefficients are equal, then the probabilities corresponding to the corresponding Schmidt basis vectors are also equal.

The strategy of this proof is as follows. Like in Section 3.1 we consider additional measurements on A and B with the measurement bases 'rotated'. Then, we consider the probability distribution of measurement outcomes, only focussing on the probabilities of the outcomes corresponding to the two equal Schmidt coefficients. It turns out that we then obtain probability distributions equal to those of Section 3.1. Therefore we can apply the same derivation.

Assume that two coefficients of (3.17) are equal. Since the ordering of basis elements is arbitrary, we can without loss of generality assume $c_0 = c_1$. The measurement on A corresponds to an observable:

$$O = -1 \cdot |0\rangle\langle 0| + 1 \cdot |1\rangle\langle 1| + \sum_i i \cdot |i\rangle\langle i|. \tag{3.18}$$

We want to prove that, for any z,

$$P(X = +1|z) = P(X = -1|z). \tag{3.19}$$

Again we define

$$|\theta\rangle = \cos\frac{\theta}{2}|0\rangle + \sin\frac{\theta}{2}|1\rangle, \tag{3.20}$$

and observables

$$O_\theta = -1 \cdot |\theta\rangle\langle\theta| + 1 \cdot |\theta+\pi\rangle\langle\theta+\pi| + \sum_i i|i\rangle\langle i|. \tag{3.21}$$

For $A(B)$, we associate the setting $v(w)$ with the observable $O_v(O_w)$. The original measurement of O on A is equal to a measurement of $O_{V=0}$.

We focus only on cases where the outcomes are ± 1. Therefore, we define, for $x, y \in \{-1, +1\}$,

$$P_{|1|}(x,y)_{v,w} = P(x,y||X| = |Y| = 1)_{v,w}. \tag{3.22}$$

As we will show in Section 3.2.1, a similar calculation as in Section 2.7 gives

$$P_{|1|}(x,y)_{v,w} = \frac{1}{4}\left(1 + x \cdot y \cdot \cos\left(v - w\right)\right), \tag{3.23}$$

which is equal to the probability distribution $P(x,y)_{v,w}$ of the measurement outcomes for the Bell state discussed in Section 3.1. Therefore, we can apply the same steps (3.8)-(3.16), giving, for any z,

$$P_{|1|}(X=+1|z)_{V=0} = P_{|1|}(X=-1|z)_{V=0} = \frac{1}{2}. \tag{3.24}$$

Since

$$P(X=+1|z)_v = P_{|1|}(X=+1|z)_v \cdot P(|X|=|Y|=1|z)_v, \tag{3.25}$$

we have

$$\begin{aligned} P(X=+1|z) &= P_{|1|}(X=+1|z)_{V=0} \cdot P(|X|=|Y|=1|z)_{V=0} \\ &= P_{|1|}(X=-1|z)_{V=0} \cdot P(|X|=|Y|=1|z)_{V=0} \\ &= P(X=-1|z). \square \end{aligned} \tag{3.26}$$

Note that this result does not say that the two probabilities are independent of Z, only that they are equal for any value of Z. However, in the case that all Schmidt coefficients are equal, applying the result says that all probabilities are equal. Because for any measurement the probabilities of different outcomes add up to 1, every probability equals $1/d$, where d is the Schmidt order of the state. Therefore, in this case, the probabilities are independent of Z:

Corollary 1. *Consider a system $A+B$ of dimension $d \times d$ in the state:*

$$|\Phi\rangle_{AB} = \sum_{i=0}^{d-1} \frac{1}{\sqrt{d}} |i\rangle_A |i\rangle_B. \tag{3.27}$$

Consider a projective measurement on subsystem A in the basis $|i\rangle_A$. In any QM-compatible hidden variable theory satisfying Parameter Independence the probabilities of the outcomes are $1/d$ where d is the Schmidt order of $|\Phi\rangle_{AB}$, i.e. equal to the probabilities predicted by quantum mechanics and independent of any supplemental variables.

A state with Schmidt order $d \geq 2$ and all Schmidt coefficients equal is called a *maximally entangled state*.

3.2.1 Calculation of $P_{|1|}(x,y)_{v,w}$

We want to calculate

$$P_{|1|}(x,y)_{v,w} = \frac{P(x,y,|X|=|Y|=1)_{v,w}}{P(|X|=|Y|=1)_{v,w}}. \tag{3.28}$$

where the state to be measured is

$$|\Phi\rangle_{AB} = c_0(|0\rangle_A|0\rangle_B + |1\rangle_A|1\rangle_B) + \sum_{i\geq 2} c_i |i\rangle_A |i\rangle_B. \tag{3.29}$$

and the settings v and w correspond to observables O_v for A and O_w for B with

$$O_\theta = -1 \cdot |\theta\rangle\langle\theta| + 1 \cdot |\theta+\pi\rangle\langle\theta+\pi| + \sum_i i|i\rangle\langle i|. \tag{3.30}$$

First we concentrate on the numerator of (3.28). We only have to consider probabilities of outcomes ± 1, because for the other outcomes $|X| = |Y| = 1$ does not hold, and therefore the probability is zero.

The calculation is almost exactly the same as in Section 2.7. Using the same definitions, instead of (2.60) we have

$$\begin{aligned} &P(x, y||X| = |Y| = 1)_{v,w} = |_{AB}\langle\phi| \left(|2v'\rangle_A \otimes |2w'\rangle_B\right)|^2 \\ &= |c_0 \left({}_A\langle 0|_B\langle 0| + {}_A\langle 1|_B\langle 1|\right) \left(\cos v'|0\rangle_A + \sin v'|1\rangle_A\right) \otimes \left(\cos w'|0\rangle_B + \sin w'|1\rangle_B\right)|^2 \\ &= \dots . \end{aligned} \tag{3.31}$$

The only difference is that $1/\sqrt{2}$ is replaced by c_0. Repeating the rest of the steps of (2.60), we get

$$P(x, y||X| = |Y| = 1)_{v,w} = \frac{c_0^2}{2} \left(1 + x \cdot y \cdot \cos\left(v - w\right)\right). \tag{3.32}$$

For the denominator of (3.28), we have

$$P(|X| = |Y| = 1)_{v,w} = \sum_{x,y=\pm 1} P(x, y, |X| = |Y| = 1)_{v,w} = 2 \cdot c_0^2. \tag{3.33}$$

Therefore,

$$P_{|1|}(x, y)_{v,w} = \frac{1}{4} \left(1 + x \cdot y \cdot \cos\left(v - w\right)\right). \tag{3.34}$$

3.3 Extension to states with unequal Schmidt coefficients

The next step is to extend Theorem 2 further, to all probabilities of outcomes of measurements in the Schmidt basis, even if the corresponding Schmidt coefficients are unequal. The idea is to first convert the entangled state to an (almost) maximally entangled state, by splitting up the terms, so that all Schmidt coefficients of the state become approximately equal. This is done using 'embezzling states' [50][21], which are states from which any entangled state can be extracted with arbitrary precision, using only local unitary operations on each subsystem. Then, Theorem 2 can be applied to this (almost) maximally entangled state, which means that the probabilities of measurement outcomes on this state are independent of Z.

Finally, the probabilities of the outcomes of measurements on the (almost) maximally entangled state can be related to the probabilities of outcomes of measurements on the original state, which has unequal Schmidt coefficients.

Theorem 3. *Consider the state of a system $A + B$, expressed in Schmidt bases of A and B:*

$$|\Phi\rangle_{AB} = \sum_{i=0}^{d-1} c_i |i\rangle_A |i\rangle_B. \tag{3.35}$$

Consider a projective measurement on subsystem A, in the basis $|i\rangle_A$. In any QM-compatible hidden variable theory satisfying Parameter Independence, the probabilities of the outcomes are equal to the probabilities predicted by quantum mechanics and independent of any supplemental variables.

For this proof, we need *embezzling states*. The family of embezzling states is denoted as $\{|\mu(n)\rangle_{A''B''}|n \in \mathbb{N}\}$. An embezzling state is a bipartite state that can be used to transform some initial state $|0\rangle_{A'}|0\rangle_{B'}$ to any desired entangled state $\sum_i c_i|i\rangle_{A'}|i\rangle_{B'}$, using only local transformations, which correspond to unitary operators on $\mathcal{H}_{A''} \otimes \mathcal{H}_{A'}$ and $\mathcal{H}_{B''} \otimes \mathcal{H}_{B'}$. The transformation is generally not exact, but can be chosen to be of arbitrary precision, by picking a state from the family with n big enough. We will elaborate on the details of the transformation and its precision in Sections 3.3.2 and 3.3.3.

We will use an embezzling state to create maximally entangled states. There exist unitary operators $U(n)^m_{A''A}$, $U(n)^m_{B''B}$ and an embezzling state $|\mu(n)\rangle_{A''B''}$, such that

$$\begin{aligned}(U(n)^m_{A''A'} \otimes U(n)^m_{B''B'})\,(|\mu(n)\rangle_{A''B''} \otimes |0\rangle_{A'}|0\rangle_{B'}) \\ \approx |\mu(n)\rangle_{A''B''} \otimes \frac{1}{\sqrt{m}} \sum_{j=0}^{m-1} |j\rangle_{A'}|j\rangle_{B'},\end{aligned} \tag{3.36}$$

where $|0\rangle_{A'}|0\rangle_{B'}$ is some initial state. Of course, the systems A' and B' should both have at least dimension m in order to contain the maximally entangled state with Schmidt order m.

As mentioned at the beginning of this section, we want to convert $|\Phi\rangle_{AB}$ to a state which is almost maximally entangled. This can be done by coupling each term to a maximally entangled state with the Schmidt order proportional to the square of the coefficient of the term. More formally, we associate with each coefficient c_i an integer m_i, such that

$$\frac{c_0^2}{m_0} \approx \frac{c_1^2}{m_1} \approx \ldots \approx \frac{c_{d-1}^2}{m_{d-1}} \approx \frac{1}{g}, \tag{3.37}$$

where $g = \sum_i m_i$. Now we define the unitary operators

$$\begin{aligned}\mathbf{U}_{A''A'A} &:= \sum_{i=0}^{d-1} U(n)^{m_i}_{A''A'} \otimes |i\rangle\langle i|_A; \\ \mathbf{U}_{B''B'B} &:= \sum_{i=0}^{d-1} U(n)^{m_i}_{B''B'} \otimes |i\rangle\langle i|_B.\end{aligned} \tag{3.38}$$

Applying these operators, we get

$$\begin{aligned}\mathbf{U}_{A''A'A} \otimes \mathbf{U}_{B''B'B}(|\mu(n)\rangle_{A''B''} \otimes |0\rangle_{A'B'} \otimes |\Phi\rangle_{AB}) \\ \approx |\mu(n)\rangle_{A''B''} \otimes \sum_{i=0}^{d-1} \sum_{j=0}^{m_i-1} \frac{c_i}{\sqrt{m_i}} |j\rangle_{A'}|i\rangle_A|j\rangle_{B'}|i\rangle_B \\ \approx |\mu(n)\rangle_{A''B''} \otimes \sum_{i=0}^{d-1} \sum_{j=0}^{m_i-1} \frac{1}{\sqrt{g}} |j,i\rangle_{A'A}|j,i\rangle_{B'B},\end{aligned} \tag{3.39}$$

where $|j,i\rangle_{A'A} := |j\rangle_{A'}|i\rangle_A$.

This is approximately a maximally entangled state. For simplicity, we now assume that these transformations can be performed perfectly; we postpone a

discussion of the approximations involved to Section 3.3.1. According to Theorem 2, when performing a measurement on $B + B'$ in the $|j, i\rangle$ basis, the probability for every outcome is then $1/g$, in any hidden variable theory satisfying Paramater Independence.

Now consider the situation where the embezzlement transformation has been performed only on system $B+B'+B''$. This means that the transformation corresponding to $\mathbf{U}_{B''B'B}$ has been performed, but not the one corresponding to $\mathbf{U}_{A''A'A}$. Now there is still a perfect correlation between a measurement on A and a measurement on $B+B'$: at A the measurement produces an eigenvalue corresponding to $|i\rangle_A$ if and only if the measurement at $B+B'$ produces an eigenvalue corresponding to one of the eigenvectors $\{|j, i\rangle_{B'B}|0 \leq j \leq m_i - 1\}$.

We now invoke Parameter Independence: the probabilities at $B+B'$ are independent of what local transformations are performed on $A+A'+A''$. If the embezzlement transformation is performed at both sides, the system $A+A'+B+B'$ is approximately in a maximally entangled state of Schmidt order g, and therefore the probability corresponding to any eigenvector $|j, i\rangle_{B'B}$ is $1/g$. But this should then also be true when the embezzlement transformation is only performed at $B+B'+B''$ and not at $A+A'+A''$. Therefore, also in this case the probability corresponding to a single $|j, i\rangle_{B'B}$ is $1/g$, and the sum of probabilities corresponding to the eigenvectors $\{|j, i\rangle_{B'B}|0 \leq j \leq m_i - 1\}$ is m_i/g. But then the probability at A to get the eigenvector corresponding to $|i\rangle_A$ is also m_i/g.

Now we again invoke Parameter Independence: because the probabilities at A do not depend on any transformation performed on systems at B, the probability associated to $|i\rangle_A$ is always m_i/g, also without any embezzlement operation performed at $B+B'+B''$. Since this probability equals c_i^2, we see indeed that the probabilities are equal to the quantum mechanical ones, and independent of any hidden variables.□

3.3.1 About the approximations involved

There are two approximations that we have neglected. First, the states created using embezzling states are not exactly the desired states. Second, not all sets of coefficients allow a transformation to a maximally entangled state: when, for two coefficients c_i, c_j the ratio c_i^2/c_j^2 is irrational, there are no integers m_i and m_j such that $c_i^2/m_i = c_j^2/m_j$. Thus in general (3.37) can only hold approximately.

Embezzlement

To discuss the precision of the embezzlement transformation, we introduce a distance measure between two quantum states called the *fidelity* [40]. For pure states, the fidelity is simply

$$F(|\psi\rangle, |\xi\rangle) = |\langle\xi|\psi\rangle|. \tag{3.40}$$

The fidelity of a transformation is then defined as the fidelity between the desired state and the actual state after the transformation. A fidelity close to 1 means that the two states are 'close' to each other. In Section 3.3.3 we show that the fidelity of the embezzlement transformation can be as close to 1 as desired. It is also easy to see that if the fidelity of each of the transformations $U(n)^{m(k)}_{A''A'} \otimes$

$U(n)^{m(k)}_{B''B'}$ is greater than $1-\epsilon$, then also the fidelity of the total transformation $\mathbf{U}_{A''A'A} \otimes \mathbf{U}_{B''B'B}$ is greater than $1-\epsilon$.

However, we need to know how this relates to probabilities of outcomes. We will now show how the fidelity relates to differences between probabilities.

For a non-degenerate projective measurement on a state $|\psi\rangle$, the probability of the outcome corresponding to the eigenvector $|i\rangle$ is $|\langle\psi|i\rangle|^2$. Now, let $|\phi\rangle$ be another state such that

$$F(|\phi\rangle, |\psi\rangle) > 1 - \epsilon \tag{3.41}$$

for some $\epsilon > 0$. We are going to derive an upper bound for the difference in the probabilities $|\langle\psi|i\rangle|^2$ and $|\langle\phi|i\rangle|^2$. By the triangle inequality, for $y, z \in \mathbb{C}$,

$$\left.\begin{aligned} |y| = |y - z + z| \leq |y - z| + |z| \Rightarrow |y| - |z| \leq |y - z| \\ |z| = |z - y + y| \leq |z - y| + |y| \Rightarrow |z| - |y| \leq |y - z| \end{aligned}\right\} \Rightarrow ||y| - |z|| \leq |y - z|. \tag{3.42}$$

Since the probabilities $|\langle\psi|i\rangle|^2$ are independent of the phase of $|\psi\rangle$, we can redefine $|\psi\rangle$ such that $|\langle\phi|\psi\rangle| = \langle\phi|\psi\rangle > 1 - \epsilon$. Now

$$\begin{aligned} \big||\langle\phi|i\rangle|^2 - |\langle\psi|i\rangle|^2\big| &= \big||\langle\phi|i\rangle| + |\langle\psi|i\rangle|\big| \cdot \big||\langle\phi|i\rangle| - |\langle\psi|i\rangle|\big| \\ &\leq 2 \cdot |\langle\phi|i\rangle - \langle\psi|i\rangle| = 2 \cdot |\langle\phi - \psi|i\rangle| \leq 2 \cdot ||\phi - \psi||, \end{aligned} \tag{3.43}$$

where we used the Cauchy-Schwarz inequality [40]

$$|\langle i|j\rangle| \leq ||i|| \cdot ||j||, \tag{3.44}$$

where $||i|| = \sqrt{\langle i|i\rangle}$ is the norm of $|i\rangle$. Now,

$$||\phi - \psi|| = \sqrt{\langle\phi - \psi|\phi - \psi\rangle} = \sqrt{2 - 2\mathrm{Re}(\langle\phi|\psi\rangle)} < \sqrt{2\epsilon}. \tag{3.45}$$

So,

$$\big||\langle\phi|i\rangle|^2 - |\langle\psi|i\rangle|^2\big| < 2\sqrt{2\epsilon} =: \delta. \tag{3.46}$$

Therefore, we can make the probability of ϕ and ψ as close to each other as we want: if we want the difference to be no more than δ, we choose

$$\epsilon = \frac{\delta^2}{8}. \tag{3.47}$$

Creating a maximally entangled state

Even when the embezzling transformation would be perfect, still not always a maximally entangled state can be created from an entangled state with unequal Schmidt coefficients. As already mentioned, this is the case when the ratio between the squares of two coefficients is irrational. Let c_i and c_j be such coefficients. We want to find integers m_i and m_j such that

$$\frac{c_i^2}{m_i} \approx \frac{c_j^2}{m_j} \Leftrightarrow \frac{c_i^2/m_i}{c_j^2/m_j} \approx 1. \tag{3.48}$$

By picking integers m_i and m_j big enough, the above ratio can be made arbitrarily close to 1. This also means that the probabilities $P_{|1|}$ defined in (3.22) can be made arbitrarily close to its desired values.

How do the above considerations help in justifying Theorem 3? Assume that Theorem 3 does *not* hold. Then there is some hidden variable theory satisfying Parameter Independence for which some probabilities are dependent on a hidden variable. Call this 'theory M'. We can roughly say that by choosing a high enough fidelity for the embezzlement, as well as choosing high enough integers to create the almost maximally entangled state, all the probabilities involved can be as close as we want to the desired ones. But this means that we can derive bounds as tight as we want on how much the probabilities depend on hidden variables. By making the bounds tight enough, it can be shown that theory M cannot be correct. So, for every such theory M we can derive a contradiction, and therefore Theorem 3 holds.

While not being mathematically rigorous here, we think that we have made it plausible that Theorem 3 holds. A mathematically complete proof of this would be welcome, but falls outside the scope of this thesis.

3.3.2 Embezzlement Explained

Embezzling states are a family of states defined as[1]

$$|\mu(n)\rangle_{A'B'} := \frac{1}{\sqrt{C(n)}} \sum_{j=1}^{n} \frac{1}{\sqrt{j}} |j\rangle_{A'} |j\rangle_{B'}, \tag{3.49}$$

where $C(n) = \sum_{j=1}^{n} 1/j$ is a normalization constant. The higher n, the more accurately entangled states can be extracted. So, how are states extracted from this state? Suppose we want to extract an entangled state with Schmidt order m:

$$|\phi\rangle_{AB} = \sum_{i=1}^{m} v_i |i\rangle_A |i\rangle_B; \quad v_i \in \mathbb{R}_+. \tag{3.50}$$

Consider the tensor product of this state and $|\mu(n)\rangle_{A'B'}$,

$$|\mu(n)\rangle_{A'B'} \otimes |\phi\rangle_{AB} = \sum_{i=1}^{m} \sum_{j=1}^{n} \frac{1}{\sqrt{C(n)}} \frac{v_i}{\sqrt{j}} |i,j\rangle_{AA'} |i,j\rangle_{BB'}. \tag{3.51}$$

where we have defined $|i,j\rangle = |i\rangle|j\rangle$. Now the idea is to only keep the n terms with the largest coefficients, and replace the values of the coefficients with the values $\{1/j | j = 1, \ldots, n\}$. We define the pairs $\{(i_1, j_1), (i_2, j_2), \ldots\}$ of values of i and j where the value of $\frac{v_i}{\sqrt{j}}$ is ordered from high to low, i.e.

$$\frac{v_{i_1}}{\sqrt{j_1}} \geq \frac{v_{i_2}}{\sqrt{j_2}} \geq \frac{v_{i_3}}{\sqrt{j_3}} \geq \ldots. \tag{3.52}$$

Now consider the vector

$$\frac{1}{\sqrt{C(n)}} \sum_{r=1}^{n} \frac{v_{i_r}}{\sqrt{j_r}} |i_r, j_r\rangle_{AA'} |i_r, j_r\rangle_{BB'}. \tag{3.53}$$

[1] In this section we let the basis vectors start at $|1\rangle$ instead of $|0\rangle$, since this simplifies the expressions considerably.

and replace the coefficients as follows:

$$\frac{1}{\sqrt{C(n)}}\frac{v_{i_r}}{\sqrt{j_r}} \mapsto \frac{1}{\sqrt{C(n)}}\frac{1}{\sqrt{r}}, \tag{3.54}$$

resulting in the state

$$|\lambda(n)\rangle = \frac{1}{\sqrt{C(n)}}\sum_{r=1}^{n}\frac{1}{\sqrt{r}}|i_r, j_r\rangle_{AA'}|i_r, j_r\rangle_{BB'}. \tag{3.55}$$

It seems that we have botched up the desired state (3.51) quite a lot by throwing away and changing coefficients. However, it turns out that for n large enough, (3.55) is arbitrarily close to (3.51), i.e.: for every $\epsilon > 0$ there is a N such that for all $n > N$, $\langle\lambda(n)|(|\mu(n)\rangle \otimes |\phi\rangle_{AB}) > 1-\epsilon$, which we show in the next section. Also note that (3.55) has the same coefficients as (3.49). Therefore, (3.55) can be easily obtained from (3.49) by performing only unitary operations on A+A' and B+B'. This is done as follows.

We start with a bipartite system A+B which is in some initial state $|0\rangle_A|0\rangle_B$. The Hilbert space of both systems A and B must have at least m dimensions. Consider unitary operators $U_{A'A}, U_{B'B}$ which perform the following transformations:

$$\begin{aligned} U_{A'A}|r\rangle_{A'}|0\rangle_A &= |j_r\rangle_{A'}|i_r\rangle_A; \\ U_{B'B}|r\rangle_{B'}|0\rangle_B &= |j_r\rangle_{B'}|i_r\rangle_B. \end{aligned} \tag{3.56}$$

Applying $U_{A'A} \otimes U_{B'B}$ to $|\mu(n)\rangle_{A'B'} \otimes |0\rangle_A|0\rangle_B$:

$$\begin{aligned} &(U_{A'A} \otimes U_{B'B})\,\frac{1}{\sqrt{C(n)}}\sum_{r=1}^{n}\frac{1}{\sqrt{r}}|r\rangle_{A'}|0\rangle_A|r\rangle_{B'}|0\rangle_B \\ &\quad = \frac{1}{\sqrt{C(n)}}\sum_{r=1}^{n}\frac{1}{\sqrt{r}}|j_r\rangle_{A'}|i_r\rangle_A|j_r\rangle_{B'}|i_r\rangle_B, \end{aligned} \tag{3.57}$$

as desired.

3.3.3 Fidelity of the embezzlement transformation

We will prove that for any $\epsilon > 0$ the fidelity $1-\epsilon$ can be achieved, for the case that is relevant for us: when the state $|\phi\rangle_{AB}$ to be extracted is a maximally entangled state with Schmidt order f:

$$|\xi(f)\rangle_{AB} := \frac{1}{\sqrt{f}}\sum_{i=0}^{f-1}|i\rangle_A|i\rangle_B. \tag{3.58}$$

In this case, the desired state after the transformation is

$$|\mu(n)\rangle_{A'B'} \otimes |\xi(f)\rangle_{AB} = \sum_{i=0}^{f-1}\sum_{j=1}^{n}\frac{1}{\sqrt{C(n)\cdot j\cdot f}}|j\rangle_{A'}|i\rangle_A|j\rangle_{B'}|i\rangle_B \tag{3.59}$$

The n highest coefficients are, in decreasing order:

$$\frac{1}{\sqrt{C(n)}}\left\{\underbrace{\frac{1}{\sqrt{f}},\ldots,\frac{1}{\sqrt{f}}}_{f\text{ numbers}},\underbrace{\frac{1}{\sqrt{2f}},\ldots,\frac{1}{\sqrt{2f}}}_{f\text{ numbers}},\ldots,\underbrace{\frac{1}{\sqrt{\lfloor\frac{n}{f}\rfloor\cdot f}},\ldots,\frac{1}{\sqrt{\lfloor\frac{n}{f}\rfloor\cdot f}}}_{f\text{ numbers}},\right.$$
$$\left.\underbrace{\frac{1}{\sqrt{\left(1+\lfloor\frac{n}{f}\rfloor\right)\cdot f}},\ldots,\frac{1}{\sqrt{\left(1+\lfloor\frac{n}{f}\rfloor\right)\cdot f}}}_{n\bmod f=f\cdot\left(\frac{n}{f}-\lfloor\frac{n}{f}\rfloor\right)\text{ numbers}}\right\}, \quad (3.60)$$

where $\lfloor z\rfloor$ is the *floor* of z: the largest integer not greater than z. The coefficients of the actual state after the transformation $|\lambda(n)\rangle$ are

$$\frac{1}{\sqrt{C(n)}}\left\{\frac{1}{1},\frac{1}{2},\frac{1}{2},\ldots,\frac{1}{n}\right\}. \quad (3.61)$$

So, for the inner product we have

$$\begin{aligned}
&\langle\lambda(n)|\left(|\mu(n)\rangle_{A'B'}\otimes|\xi(f)\rangle_{AB}\right)=\frac{1}{C(n)}\left(\frac{1}{\sqrt{1f}}+\frac{1}{\sqrt{2f}}+\ldots+\frac{1}{\sqrt{f\cdot f}}\right.\\
&+\frac{1}{\sqrt{(f+1)2f}}+\frac{1}{\sqrt{(f+2)2f}}+\ldots+\frac{1}{\sqrt{2f\cdot 2f}}+\ldots\\
&+\frac{1}{\sqrt{\left(\left(\lfloor\frac{n}{f}\rfloor-1\right)f+1\right)\lfloor\frac{n}{f}\rfloor f}}+\ldots+\frac{1}{\sqrt{\lfloor\frac{n}{f}\rfloor f\lfloor\frac{n}{f}\rfloor f}}\\
&\left.+\frac{1}{\sqrt{\left(\lfloor\frac{n}{f}\rfloor f+1\right)\left(\lfloor\frac{n}{f}\rfloor+1\right)f}}+\ldots+\frac{1}{\sqrt{n\left(\lfloor\frac{n}{f}\rfloor+1\right)f}}\right)\\
&\geq\frac{1}{C(n)}\left(f\cdot\left(\frac{1}{f}+\frac{1}{2f}+\ldots+\frac{1}{\lfloor\frac{n}{f}\rfloor f}+\left(\frac{n}{f}-\lfloor\frac{n}{f}\rfloor\right)\cdot\frac{1}{\left(\lfloor\frac{n}{f}\rfloor+1\right)f}\right)\right)\\
&=\frac{1}{C(n)}\left(\frac{1}{1}+\frac{1}{2}+\ldots+\frac{1}{\lfloor\frac{n}{f}\rfloor}+\left(\frac{n}{f}-\lfloor\frac{n}{f}\rfloor\right)\cdot\frac{1}{\lfloor\frac{n}{f}\rfloor+1}\right). \quad (3.62)
\end{aligned}$$

Now we will need some bounds involving $C(n)$, which is also known als the n-th *harmonic number*. Note that $C(n)$ equals the grey area between $x=0$ and $x=n$ in Figure 3.1. We now define $D(z)$ for $z\geq 1$ which is the area of the bars between $x=0$ and $x=z$. Unlike $C(n)$, $D(z)$ is also defined for non-integers. For integers we have $D(n)=C(n)$. For non-integers we have

$$D(z)=\frac{1}{1}+\frac{1}{2}+\ldots+\frac{1}{\lfloor z\rfloor}+(z-\lfloor z\rfloor)\frac{1}{\lfloor z\rfloor+1}. \quad (3.63)$$

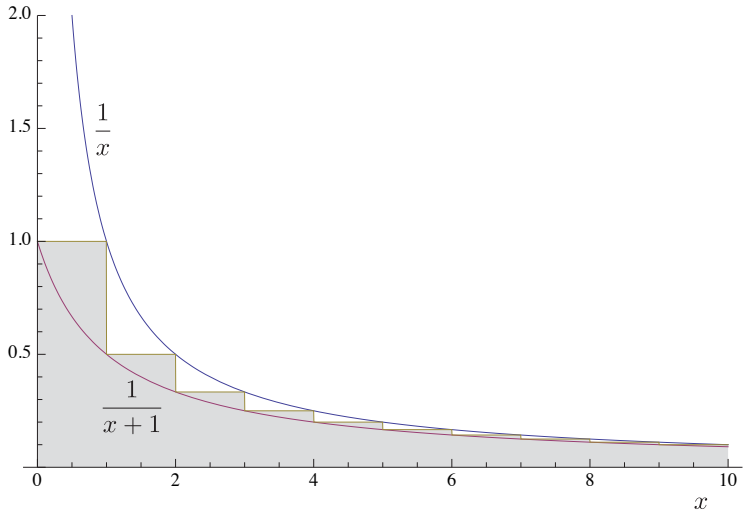

Figure 3.1: Using the two plotted graphs, bounds for the grey area can be derived.

In Figure 3.1 we have also plotted the graphs $1/x$ and $1/(x+1)$. Comparing the areas, we can deduce the following inequalities

$$D(z) \leq 1 + \int_1^z \frac{1}{x}\mathrm{d}x = 1 + \ln x \tag{3.64}$$

$$D(z) \geq \int_0^z \frac{1}{x+1}\mathrm{d}x = \ln(x+1) \tag{3.65}$$

$$D(z_2) - D(z_1) \leq \int_{z_1}^{z_2} \frac{1}{x}\mathrm{d}x = \ln z_2 - \ln z_1. \tag{3.66}$$

Now, the right-hand side of (3.62) can be written as

$$\frac{D(n/f)}{D(n)} = 1 - \frac{D(n) - D(n/f)}{D(n)}. \tag{3.67}$$

Using

$$D(n) \geq \ln(1+n) > \ln n \tag{3.68}$$

and

$$D(n) - D(n/f) \leq \ln n - \ln(n/f) = \ln f, \tag{3.69}$$

we have

$$\langle\lambda(n)|\left(|\mu(n)\rangle_{A'B'} \otimes |\xi(f)\rangle_{AB}\right) > 1 - \frac{\ln f}{\ln n}. \tag{3.70}$$

So, for any f and $\epsilon > 0$, if $n > f^{1/\epsilon}$, then the fidelity is greater than $1 - \epsilon$.

3.4 Extension to arbitrary projective measurements

We will now perform the final step which extends Theorem 3 to arbitrary projective measurements:

Theorem 4. *Consider a projective measurement performed on a pure quantum state. In any QM-compatible hidden variable theory satisfying Parameter Independence, the probabilities of the outcomes are independent of any supplemental variables. Therefore, any such theory is trivial.*

3.4.1 Von Neumann standard description of measurement

For this last step we use the standard measurement scheme of quantum mechanics, first presented by Von Neumann in 1932 [51]. In this treatment, not only the object to be measured but also the measuring device is treated as a quantum mechanical system. A measurement involves a unitary interaction between the system to be measured and the measurement device, where eigenstates of the observable are coupled to orthogonal states of the measurement device. To illustrate this, consider a quantum system in the state

$$\sum_i c_i |i\rangle_S, \tag{3.71}$$

which is expressed in the eigenbasis of the observable corresponding to the projective measurement to be performed. Suppose that the measurement device is initially in the state $|0\rangle_D$. Now, the unitary interaction that couples the object to the measurement device acts like

$$U\left(\sum_i c_i |i\rangle_S |0\rangle_D\right) = \sum_i c_i |i\rangle_S |i\rangle_D. \tag{3.72}$$

If more than one of the c_i are nonzero, this interaction results in an entangled state of the system and the measuring device. This state contains multiple terms, corresponding to multiple outcomes. However, in reality there is only a single outcome. Therefore it is usually supposed that the following transition occurs:

$$\sum_i c_i |i\rangle_S |i\rangle_D \rightsquigarrow |i\rangle_S |i\rangle_D \text{ with probability } |c_i|^2. \tag{3.73}$$

The nature of this transition is unclear. How this problem is dealt with depends on the interpretation of quantum mechanics. We can at least say that the measurement yields as an outcome the eigenvalue corresponding to $|i\rangle_S$ with probability $|c_i|^2$. Also, directly after the measurement, to the system can be assigned the eigenstate corresponding to the outcome.

3.4.2 Measuring D is measuring S

According to the ideal measurement scheme presented above, when a measurement is performed on S at some stage the following state is produced:

$$\sum_i c_i |i\rangle_S \otimes |i\rangle_D. \tag{3.74}$$

In fact, this coupling can be extended. Assume D is only a small part of the measurement device, which first interacts with S, and D' is the next part which is involved in the interaction. Then, we have

$$\sum_i c_i|i\rangle_S \mapsto \sum_i c_i|i\rangle_S|i\rangle_D \mapsto \sum_i c_i|i\rangle_S|i\rangle_D|i\rangle_{D'} \mapsto \ldots$$
$$\mapsto |i\rangle_S|i\rangle_D|i\rangle_{D'} \text{ with probability } |c_i|^2 \tag{3.75}$$

Here we can see that a measurement performed on D *is* in fact *by definition* a measurement performed on S. Starting from the second step, (3.75) describes a measurement on D. Starting from the first step, it describes a measurement on S. However, this is a single measurement with a single outcome, which is therefore a measurement of S and of D simultaneously. In other words: every ideal projective measurement performed on a system is in fact also a measurement performed on a small part of the measurement device.

The measurement on D is a measurement on a part of an entangled state, in the Schmidt basis. Therefore, Theorem 3 applies. This means that the probabilities of the outcomes of the measurement on D are independent of any hidden variables. But we have just seen that the measurement on S is actually at the same time such a measurement on D. Therefore, for any projective measurement on a pure state, the outcomes are independent of any hidden variables. This concludes the proof.

Another way to see this is that during a measurement of S, in principle the entangled state (3.74) can be isolated. Then, D can be spatially separated from S. Now, the measurement of S can be completed by a measurement of D in the basis $|i\rangle_D$. But, for this measurement Theorem 3 holds, for which the main ingredients are the strong correlations between hypothetical measurements on S and D.

3.5 Summary

In this chapter we have proved several theorems, each one building on the previous one. All results apply to QM-compatible hidden variable theories satisfying Parameter Independence.

First we showed that for projective measurements on parts of Bell states, such hidden variable theories give the same probabilities as quantum mechanics. Then we showed that for projective measurements in a Schmidt basis, probabilities corresponding to equal Schmidt coefficients are equal. Then we extended this to all probabilities of a measurements in the Schmidt basis. Finally, we extended the result to arbitrary measurements on pure states, by noting that during an ideal measurement entanglement is created, and therefore any ideal measurement is at the same time a measurement in the Schmidt basis on a part of an entangled state.

The result is a no-go theorem for QM-compatible hidden variable theories satisfying Parameter Independence. In the next chapter we will discuss some differences between the formulation of the result in this chapter and CR's original formulation of their result.

Chapter 4

Some differences between our reformulation and the original formulation

The version of the Colbeck-Renner result given in the previous chapter differs from the original result as presented in [18] and [20] on several points. CR's statement of their result is 'No extension of quantum theory can have improved predictive power'. Instead, we have presented the result as a no-go theorem for hidden variable theories satisfying Parameter Independence.

In this chapter we will discuss the main differences and defend the choices we have made in our reformulation. To begin with, we will argue that it is not a good idea to treat measurement settings as random variables. Instead, they should be treated as context variables. Also, we will see that Colbeck and Renner have an operational approach in mind, where all variables discussed are assumed to be accessible. In our opinion, this approach makes the result less general. This also relates to the difference between the 'No-Signaling' condition discussed by CR and the 'Parameter Independence' condition we have used. CR's statement of their result looks stronger than our reformulation, since we add the condition of Parameter Independence. However, as we will argue in Section 4.1, such a condition is also implicitly assumed by Colbeck and Renner by the way they define 'Freedom of Choice'. Finally, we will discuss how CR have used the measurement process to realize the generalization to arbitrary measurements, which in our reformulation is done in Section 3.4.

4.1 Treating measurement settings as random variables

In Section 1.4.1 we introduced a type of variables called *context variables*, which should be distinguished from random variables. We will now explain why this distinction is made.

When random variables are associated to outcomes of experiments, it is important that when we talk about probability distributions over such outcomes, the context is clear, and that all relevant factors are taken into account. This is

to ensure that the probability distributions are well-defined. This is independent of whether the probability is interpreted objectively or subjectively.

We explain this using an example. Suppose an experimenter two dice, each with 4 faces. One die (A) contains only the odd numbers 1, 3, 5 and 7. The other (B) contains the even numbers 2, 4, 6 and 8. A measurement consists of throwing a die of choice and writing down the outcome. A single random variable X is associated to the outcome of this 'experiment'. We try to define a probability distribution over the outcomes $P(x)$.

The experimenter may choose only to roll die A, for which a typical set of outcomes is

$$\{3, 1, 1, 7, 3, 5 \ldots\}. \tag{4.1}$$

He may also choose only to roll die B, which might give the outcomes

$$\{2, 4, 8, 6, 6, 4, \ldots\}. \tag{4.2}$$

Or, he can alternate between the dies, which might result in the outcomes

$$\{3, 4, 1, 8, 5, 2, \ldots\}. \tag{4.3}$$

We see that we cannot consistently associate a single probability distribution to X which is compatible with each of these three runs of throws. For example, in the first run we would want a distribution with $P(X = 3) = 1/4$, in the second run we want $P(X = 3) = 0$ and in the final run we want one of these two distributions, depending on the number of the throw (odd or even). We see that $P(x)$ is not well-defined.

We now introduce the context variable D, which for each throw gives the die that is thrown: $D = 1$ for die A and $D = 2$ for die B. Now, given a value of D, there are well-defined probability distributions of X: we can use $P(x)_{D=1}$ and $P(x)_{D=2}$ which are the probability distributions for die A and die B, respectively.

Now again consider this last game, but now let the experimenter throw a regular six-face die before each measurement. If this die gives an odd number, die A will be thrown, otherwise die B will be thrown. The random variable X will again be associated with the outcome of this last throw. Because the choice of which die to throw is now itself the result of a stochastic process (with uniform probability 1/2), we can treat D as a random variable, with $P(D = 1) = P(D = 2) = 1/2$. Instead of writing $P(x)_d$, with D a context variable, we can now write $P(x|d)$. Now, $P(x)$ *is* well-defined:

$$\forall x : P(x) = \sum_d P(x|d)P(d) = \frac{1}{8}. \tag{4.4}$$

We see the dangers of assigning a probability distribution to the choice of die D in the general case. First, there is no reason to assume that this choice is the result of a stochastic process: the experimenter is free to choose any pattern of picking dice. Second, it makes the probability distribution $P(x)$ automatically well-defined, which does not always make sense considering the runs (4.1)-(4.3) we have considered above. In general, D is a *context* variable, and only when a value of D is given, the probability distribution of X is well-defined. There

is only an exception when it is clear that the context variable D is assigned a value which results from some stochastic process. In this case, we can treat D as a random variable. Then, the probability distribution of X is well-defined also when a value of D is not given.

We connect these considerations to the argument of Colbeck and Renner. They associate probability distributions $P(a)$ and $P(b)$ to the setting variables A and B (which correspond to the variables V and W in our formulation, respectively), thereby treating them as random variables. They use this to define independence between these variables and other random variables. For example, the random variable A is independent of the random variable X iff $P(a, x) = P(a)P(x)$. Instead, we prefer to treat A and B as *context* variables. There is no problem defining independence between a random and a context variable: independence of a context variable A and a random variable X can be defined as

$$\forall a', a'' \in \tilde{A} : P(x)_{a'} = P(x)_{a''}. \tag{4.5}$$

An objection might be that it is not clear how to define independence between two context variables if they cannot be treated as random variables. But we think it is not needed to define such an independence. If we have two context variables A and B, it is possible to let them depend on each other in any way. For example, we can always choose $A = B$.

There is another unwanted consequence when context variables are treated as random variables. Suppose we express Parameter Independence (2.10) using only random variables:

$$\forall w', w'' : P(x|z, v, w') = P(x|z, v, w''). \tag{4.6}$$

Suppose this condition is violated:

$$\exists w', w'' : P(x|z, v, w') \neq P(x|z, v, w''). \tag{4.7}$$

Now, suppose Z and X are accessible by Alice, V is controllable by Alice, and W is controllable by Bob. In this case this violation is usually interpreted as allowing the possibility for Bob to signal to Alice, because the distribution $P(x|z, v)$ at Alice's side seems to depend on the value of W. But in fact, since W is treated as a random variable here, $P(x|z, v)$ is fixed:

$$P(x|z, v) = \sum_w P(x|z, v, w)P(w|z, v). \tag{4.8}$$

It is unclear now how Bob could signal to Alice. This is not too surprising, since Bob's 'free choice' of a value for W has to obey a fixed probability distribution $P(w)$. However, if we treat W as a context variable, (4.7) becomes

$$\exists w', w'' : P(x|z)_{v,w'} \neq P(x|z)_{v,w''}. \tag{4.9}$$

Now it is really clear how Bob can signal to Alice: there is not a single distribution $P(x|z)_v$, but there are different ones for different values of W.

The discussion about treating settings as random variables is not seen often in the literature. Usually settings are implicitly treated as random variables

because they are used in conditional probabilities. For example, writing down $P(x|a)$ implies that A is a random variable since

$$P(x|a) = \frac{P(x,a)}{P(a)}. \tag{4.10}$$

However, in the writings of Colbeck and Renner the treatment of settings as random variables is very explicit, since distributions over those settings are defined. We will come back to this in Section 4.1.

We are not the first to criticize the assignment of probability distributions to measurement settings. In [14], Jeremy Butterfield writes

> '[A]n act of measurement surely need not have a probability. Why should every proposition or event have a probability? And since [the setting] a is a feature of a complex apparatus, and is fixed or at least influenced by the choice of the experimenter, it seems quite a good candidate for not having a probability.'

4.2 The operational approach of Colbeck and Renner

In Chapter 3, we have not assumed anything about the accessibility of the hidden variable Z. Therefore, the class of hidden variable theories we have studied includes theories where such variables can be accessible, but also theories where such variables might be inaccessible. Colbeck and Renner, however, emphasize that their approach is operational. That is, the variables they discuss are all thought to be accessible. This is especially clear from the following answer, which can be found at [17]:

> 'Q: What about models which provide higher explanation of the correlations, but for which certain parameters or systems remain hidden (as a physical principle), either forever or until after the quantum measurements have been performed?
>
> A: We do not rule out this possibility. However, this does not contradict our claim since such hidden systems do not provide additional predictive power at the level of experimentalists, and hence are not extensions in our terminology.'

This also explains why we use the expression 'hidden variable theories' instead of 'extensions of quantum mechanics'. Colbeck and Renner use the latter, and part of their definition of such theories is that the extra variables in the extended theory are accessible. The advantage of our choice is that the result then holds not only for theories in which the supplemental variables are accessible, but also for theories where these variables are partially accessible, or not accessible at all.

4.2.1 'Spacetime Variables' and causal structure

The operational approach of Colbeck and Renner also allows them to assign spacetime coordinates to variables. The coordinates indicate a time and place at which the corresponding variable is accessible.

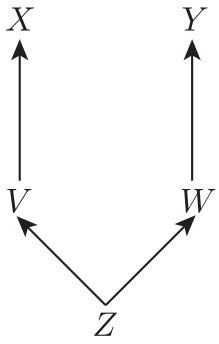

Figure 4.1: The causal order used

Because we want to explain how CR derive their result, we will now follow their operational approach. Also, we will follow them in assigning independent probability distributions over settings. In accordance with the discussion above, we interpret this as follows: the settings V and W are determined by the outcome of some stochastic process, for example independent random number generators.

CR consider an EPRB measurement: a joint measurement on a state of two maximally entangled qubits. Random variables X and Y are associated with measurement outcomes, and random variables V and W are assigned to the respective measurement settings. Z is the supplemental random variable.

The notion of a 'causal order' is introduced. Spacetime coordinates are attached to the random variables, now called 'Spacetime Variables' (SV's), so that their causal relation can be defined. As mentioned above, these coordinates refer to the spacetime point at which the variable can be accessed. The causal order is a preorder relation between variables, denoted with the symbol '$\rightsquigarrow$'. A preorder relation is reflexive ($\forall X : X \rightsquigarrow X$) and transitive ($\forall X, Y, Z : X \rightsquigarrow Y, Y \rightsquigarrow Z : X \rightsquigarrow Z$). The relation '$A \rightsquigarrow B$ is defined as 'B is in the future lightcone of A'. We write $A \not\rightsquigarrow B$ if $A \rightsquigarrow B$ is not true.

The local measurements are assumed to take place in spacelike separated regions of space-time, while Z is assumed to be in the backward lightcone of V and W. This gives a causal order generated by the relations $V \rightsquigarrow X$, $W \rightsquigarrow Y$, $Z \rightsquigarrow V$ and $Z \rightsquigarrow W$ (by transitivity, this means that also $Z \rightsquigarrow X$ and $Z \rightsquigarrow Y$). Graphically, this can be represented as in Figure 4.1.

4.3 No-Signaling from Freedom of Choice

In this part the authors attempt to derive a no-signaling condition from an assumption called 'Freedom of Choice' which they present as a very reasonable assumption. We will follow the argument as presented in [20].

4.3.1 Freedom of Choice

A random variable A is *free* if it is independent of the set of random variables outside of its causal future $\Gamma := \{X : A \not\rightsquigarrow X\}$:

$$\forall X_1, X_2, \ldots \in \Gamma : P(a, x_1, x_2, \ldots) = P(a)P(x_1, x_2, \ldots). \tag{4.11}$$

The 'Freedom of Choice' (FR) assumption says that the settings V and W are free variables with respect to the causal order defined above (see Figure 4.1). This means that

$$\begin{aligned} P(y, v, w, z) &= P(v)P(y, w, z); \\ P(x, v, w, z) &= P(w)P(x, v, z). \end{aligned} \tag{4.12}$$

Now, a condition called 'No-Signalling' is derived from this assumption. No-Signalling (NS) is defined as

$$\begin{aligned} P(x, z|v, w) &= P(x, z|v), \\ P(y, z|v, w) &= P(x, z|w). \end{aligned} \tag{4.13}$$

The derivation is simple. We have

$$\begin{aligned} P(x, z|v, w) &= \frac{P(x, z, v, w)}{P(v, w)} = \frac{P(x, z, v)P(w)}{P(v)P(w)} = P(x, z|v); \\ P(y, z|v, w) &= \frac{P(y, z, v, w)}{P(v, w)} = \frac{P(y, z, w)P(v)}{P(w)P(v)} = P(y, z|w). \end{aligned} \tag{4.14}$$

The analogue of Source Independence (2.8), with random variables instead of context variables, can also be derived:

$$P(z|v, w) = \frac{P(z, v, w)}{P(v, w)} = \frac{P(z, v)P(w)}{P(v)P(w)} = \frac{P(z)P(v)}{P(v)} = P(z). \tag{4.15}$$

Then we have also, for any w', w'',

$$P(x|z, v, w') = \frac{P(x, z|v, w')}{P(z|v, w')} = \frac{P(x, z|v)}{P(z)} = \frac{P(x, z|v, w'')}{P(z|v, w'')} = P(x|z, v, w'), \tag{4.16}$$

and similarly $P(y|z, v', w) = P(y|z, v'', w)$. That is, we have

$$\begin{aligned} \forall w', w'' : P(x|z, v, w') &= P(x|z, v, w''); \\ \forall v', v'' : P(y|z, v', w) &= P(y|z, v'', w). \end{aligned} \tag{4.17}$$

Like (4.6), this is just Parameter Independence (2.10) with random variables. The fact that CR call this No-Signaling makes sense, given that they follow an operational approach where all variables, including Z, are accessible. However, as discussed above, it is confusing to talk about signaling when settings are treated as random variables and therefore have fixed probability distributions.

4.3.2 The 'Freedom of Choice' assumption

The Freedom of Choice assumption is presented as a mild assumption. For example in [19] they claim their result is derived 'only assuming the freedom of choice of the experimenters'. This makes it sound like an assumption that can hardly be denied. However, there has been a lot of response to this: many authors disagree with the way this assumption is defined[31][32][33][39]. We also think that the assumption is stronger than its name suggests. The problem lies in the causal order that is implicit in the assumption. Using this specific causal order, a free variable is defined as 'independent of all variables not in the future lightcone'. The motivation of using this causal order is the theory of relativity, where influences at spacelike distances are usually assumed to be prohibited, since by using a Lorentz transformation such an influence would in some frame be an influence backwards in time. In the words of CR [18]:

> 'the motivation for assumption FR is that, when interpreted within the usual relativistic spacetime structure, it is equivalent to demanding that $[V]$ can be chosen such that it is uncorrelated with any pre-existing values in any reference frame. (...)
>
> We also remark that Assumption FR is consistent with a notion of relativistic causality in which an event B cannot be the cause of A if there exists a reference frame in which A occurs before B. In fact, our criterion for A to be a free choice is satisfied whenever anything correlated to A could potentially have been caused by A. However, in an alternative world with a universal (frame-independent) time, one might reject assumption FR and replace it with something weaker, for example, that A is free, if it is uncorrelated with anything in the past with respect to this universal time. Nevertheless, since experimental observations indicate the existence of relativistic spacetime, we use a notion of free choice consistent with this.'

If NS would be violated, free variables would have influences to variables at spacelike separation. But if free variables can not have such influences *by definition*, then it is no surprise that NS is satisfied.

It is clear that No-Signaling has more to do with the causal structure that is assumed than with the assumption that V and W are free variables. From the above quotation, it is also clear that CR admit that when another causal structure is assumed, No-Signaling does not follow. However, at other times (for example in the first quote of this paragraph) they do not mention this causal structure and only talk about the freedom of the experimenters.

It is especially peculiar that CR justify their use of the relativistic causal structure by the last sentence of the above quotation. Their result is about possible extensions of quantum theory, which are theories that include variables that have never been observed. In such a context it is strange to rule out beforehand any theory which has a different causal structure than that suggested by relativity, such as Bohmian mechanics. If we wonder whether quantum mechanics can be superseded by another theory, why would we rule out theories which could supersede relativity?

Moreover, there are other reasons why combining causal structures like the above with classical probability theory is suspect. Adhering to such a causal structure suggests a common cause principle can be applied to it: that any

correlation between two events that are not causally related can be screened of by a common cause. But the violation of the Bell inequality in the EPRB setting is exactly what prevents a common cause explanation for the correlations.

To conclude, we think that assuming the relativistic causal structure and the way Freedom of Choice is defined is too restrictive: it excludes influences over spacelike distances from the start. At the least it is a controversial move which makes the result vulnerable to criticism. Indeed, their result has been mostly criticized for the definition of 'Freedom of Choice', while little attention has been paid to the rest of the result.

Therefore, in our reformulation we have not made such a move, and just assumed Parameter Independence, which is similar to No-Signaling, from the beginning. We admit this makes the result less general: it now only applies to hidden variable theories with Parameter Independence, instead of all 'extensions of quantum theory'. However, in our eyes it makes the result stronger, because it does not rely on the suspicious move of deriving No-Signaling from a freedom assumption.

4.4 Relation between the measurements on S and D

In proving Theorem 4 in Section 3.4, we have argued that measurements of different parts of a system can actually be the same measurement by definition. Closely reading CR's work seems to tell a different story. This is most apparent in the following passage[1]:

> 'For later reference, we also note that, according to quantum theory, any possible evolution of a quantum system, S, corresponds to a unitary mapping on a larger state space (that may include the environment of the system). In the case of a measurement process, this larger state space includes the measurement device, D. Specifically, a projective measurement, say $\{E_x^a\}_x$, would correspond to a unitary of the form
>
> $$|\Psi\rangle \mapsto \sum_x \sqrt{E_x^a}|\Psi\rangle_S \otimes |x\rangle_D, \tag{4.18}$$
>
> where $\{|x\rangle_D\}$ are orthonormal states of the measurement device (and possibly also its environment) that encode the outcome. The outcome X of the original measurement may then be *recovered* by a *subsequent* projective measurement on D in the basis $\{|x\rangle_D\}$.' (my emphasis)

Colbeck and Renner seem to talk about consecutive measurements: first the initial measurement on S is performed (call this 'measurement 1'), and then another measurement on D may be performed (call this 'measurement 2'). However, in order to apply Theorem 3, it is crucial that when measurement 2 is performed S and D are entangled as in (4.18). Hereby CR seem to adhere to a no-collapse interpretation of quantum mechanics: states always evolve unitarily as long as the system considered is big enough so that it is isolated.

[1] [20], p. 4

This touches the problem of reconciling the projection postulate with the Schrödinger postulate (see Section 1.4.2): If there is collapse during a measurement, then which interactions count as measurements? And if there is no collapse during a measurement, then how can it be justified that the collapse postulate can still be applied for practical purposes? How this is dealt with depends on the interpretation of quantum mechanics. Are Colbeck and Renner correct in stating that the measurement on D gives the same outcome as the measurement on S? We will now address this question in the context of different interpretations.

4.4.1 No-collapse interpretations

In so-called no-collapse interpretations it is assumed that any isolated system always evolves unitarily, even if a measurement has taken place within that system. We discuss three variants: Bohmian mechanics, the modal interpretations and the many-worlds interpretation.

Bohmian mechanics [8] is a deterministic theory, giving the same predictions as ordinary quantum mechanics. It is usually formulated in the context of non-relativistic particle quantum mechanics. The positions of the particles are assumed to have definite values at all times. And indeed, if measurement 1 has been performed on S, then measurement 2 on D will give the same outcome. However, Bohmian mechanics is an example of a hidden variable theory where, given values of the hidden variables, the theory *does* give different probability distributions over outcomes than quantum mechanics. Therefore it must violate one of the assumptions that has been made, and indeed does: Parameter Independence is violated. Note however that signaling is still not possible in Bohmian mechanics because the hidden variables, i.e. the particle positions, are not directly accessible.

In modal interpretations [24] it is assumed that the state vector assigned to a system does not give all the properties of the system. Rather, for some observables, the system has one definite eigenvalue as its property, although the state vector is a superposition of different eigenstates of the observable. In this way, the gap between the superposition of the state vector and the definiteness of observed properties is closed. However, there is no consensus on what the dynamics is of the properties: how do the properties evolve when the state vector evolves? Therefore, it is not clear whether the condition of Colbeck and Renner is obeyed: that a second measurement in the same basis gives the same outcome as the first measurement.

In the many-worlds interpretation [44] it is assumed that when the combined system $S+D$ has evolved into a superposition, all the corresponding outcomes have actually occurred, but in different 'branches of reality'. Also in this case it does not follow that the second measurement gives the same outcome, since the first measurement does not even have a single outcome.

An inconsistency in CR's derivation?

If we follow CR in saying that after the first measurement there is a single definite outcome, but on the other hand evolution is always unitary when considering a system large enough, then there seems to be an inconsistency in their argument. Since they assume that measurement 2 in the basis $|x\rangle_D$ gives the

same outcome as measurement 1, a variable X denoting the outcome of this first experiment is, from the perspective of measurement 2, a supplemental variable which gives information about the measurement outcome. But should X be included in the causal structure? Is it a 'Spacetime Variable' in CR's sense? X is an example of a variable to which there is limited access from the perspective of measurement 2: the only way to get information about X is to perform measurement 2 in the basis $|x\rangle_D$. Is is not clear what X says about measurement outcomes if the second measurement is performed in another basis than $|x\rangle_D$. Furthermore, nothing can be known about X before the measurement, unless the extended theory provides some other mechanism to know something about X before the measurement was made.

In [18] we read[2]:

> 'assumption *FR* is that the input, A, of a measurement process can be chosen such that it is uncorrelated with certain other SVs, namely all those whose coordinates lie outside the future lightcone of the coordinates of A. (...) it is equivalent to demanding that A can be chosen such that it is uncorrelated with any pre-existing values in any reference frame.'

From the perspective of the experimenter performing measurement 2, X surely looks like a value which is pre-existing in some reference frame, even stronger: it is pre-existing in *any* reference frame.

It is also noteworthy that Colbeck and Renner *use* their result to show that a 'system's wave function is in one-to-one correspondence with its elements reality' [19]. This claim is directly in contradiction with their treatment of the measurement process quoted at the beginning of Section 4.4, which says that if a measurement is performed within a closed system, the whole system can on the one hand be described by a state which is the result of unitary evolution, but on the other hand the measurement has a definite outcome X. Since in general multiple outcomes are compatible with the same quantum state, X is an additional property of the system, and therefore there cannot be a one-to-one correspondence between quantum state and its elements of reality.

More on modal interpretations

According to modal interpretations, if during a measurement the measurement device D has interacted with the system to be measured S, as long as $S+D$ remains isolated, its evolution is unitary. Therefore, the state of $S+D$ will generally end up in a superposition like (4.18), which has terms corresponding to different outcomes. However, in modal interpretations, the system $S+D$ has then an additional property that singles out one of the terms in the superposition. This property can be interpreted as a hidden variable, since it is not included in the state of $S+D$. We define the hidden variable P, of which the value indicates the term that is singled out. We would now expect that when the experimenter interacts with $S+D$ to find out the outcome of the measurement, this outcome will be the one corresponding to the term singled out by P. But this would mean that P is a hidden variable which tells something about measurement outcomes because the outcome is not independent of P. Following the result of Chapter 3, this means that a theory which, in the spirit of

[2] p. 2

the modal interpretation, assigns extra values to systems which say something about measurement outcomes, violates Parameter Independence. It would be interesting to work out this issue with modal interpretations in more detail, but for now we leave it at this.

4.4.2 Relation to collapse interpretations

In collapse interpretations, sooner or later the combined system $S+D$ will stop evolving unitarily, and 'collapse' to one of the terms in (4.18). An example of a collapse interpretation is the GRW interpretation [30], where the Schrödinger equation is modified such that there is a very high probability of collapse when number of particles involved in an entangled system becomes large. Another example is an interpretation where the consciousness of the observer plays a role: once the measurement outcome is registered by a conscious observer, the total state collapses. This interpretation is usually attributed to Von Neumann and Wigner [3].

In such interpretations, at any point in time during a measurement, either the state has collapsed and there is a single outcome, or the state has evolved only unitarily. But it cannot be that there is a definite outcome while the state is still entangled between object and measurement device. So, in this case, it matters *when* measurement 2 takes place. If it takes place *before* the collapse, then one cannot really speak anymore about the outcome of the initial measurement, since this measurement has not finalized, and only the second measurement has an outcome. If the second measurement takes place *after* the collapse, then indeed the same outcome is obtained. However, since the state is not in a superposition anymore, the strong correlations are lost and therefore Theorem 3 cannot be applied anymore.

4.4.3 Alternative interpretation of the relation between the two measurements

In the above we have seen that the claim that the outcome of measurement 1 can be recovered by performing measurement 2 is not in agreement with some of the most common interpretations of quantum mechanics. We have contacted Roger Colbeck regarding this claim. In his reply, he explained that in the context of collapse theories measurement 2 should be seen as a measurement on only a small part of the measurement device, before any collapse has occurred. Then, measurement 2 is by definition also a measurement on S, and therefore measurement 1 and measurement 2 become one and the same measurement. We have used this argument in Section 3.4. However, this explanation seems to differ a bit from the reasoning in the quotation at the beginning of Section 4.4.

4.5 Summary

In this chapter we have highlighted some differences between our formulation and the original formulation of CR. The most important point is that CR try to derive the result from very little assumptions: a no-signaling condition is derived from an assumption about the freedom of experimenters. Although it is

an interesting discussion whether such a derivation is justified, it is a step vulnerable to criticism. Therefore, in our reformulation we have chosen to assume a condition similar to No-Signaling, namely Parameter Independence, instead of trying to derive it. Furthermore, we have not restricted ourselves to an operational approach: we have not assumed anything about the accessibility of the supplemental variables. Also, we have argued that it is better not to treat measurement settings, and other circumstances that can be chosen at will by experimenters, as random variables.

Finally, there is another difference that deserves mentioning. In CR's work there is no analogue of Theorem 2. Instead, they only derive a result similar to Theorem 1, which is about maximally entangled states of Schmidt order 2, and then suggest that this result can be extended to maximally entangled states of Schmidt order 2^n for any $n \in \mathbb{N}$, because the tensor product of n maximally entangled states of Schmidt order 2 equals a maximally entangled state of Schmidt order 2^n [18][3]. It is however not made explicit how this follows. Furthermore, Theorem 2 is more general because it applies to entangled states of any Schmidt order, not only of Schmidt order 2^n.

[3]Supplementary Information, p. 6

Chapter 5

Example: Di Lorenzo's models are in conflict with CR's result

Recently a paper by Antonio Di Lorenzo has appeared [22] in which he defines a class of hidden variables for spin measurements performed on a system in the singlet state

$$|\Psi\rangle_S = \frac{1}{\sqrt{2}}(|0\rangle_{S_1}|1\rangle_{S_2} - |1\rangle_{S_1}|0\rangle_{S_2}). \tag{5.1}$$

The models in this class satisfy Parameter Independence (Di Lorenzo: 'setting independence'), and Source Independence (Di Lorenzo: 'measurement independence'). However, the models are not trivial. Therefore, there seems to be a conflict with the result of Colbeck and Renner. Di Lorenzo explicitly states that there is no conflict:

> '..., the results presented here show that quantum mechanics can be extended through the specification of additional parameters $[z]$, and that this extension has improved predictive power, (...), and consequently the predicted joint probability for given $[z]$ differs from the quantum mechanical one:
>
> $$[P(x,y|z)_{\vec{a},\vec{b}} \neq P^{QM}(x,y)_{\vec{a},\vec{b}}].$$
>
> This seems to contradict the findings of [Colbeck and Renner]. However, in [their work], the impossibility to have an improved predicted power refers to the marginal probability $[P(x|z)_{\vec{a}}]$, not to the joint probability $[P(x,y|z)_{\vec{a},\vec{b}}]$. The models discussed in the present work predict marginal probabilities of $[P(x|z)_{\vec{a}}] = 1/2$ and hence do not contradict [Colbeck and Renner's work]. In other words, the apparent tension is due to the definition of "extension of quantum theory."'

This seems to be wrong. Indeed, a part of Colbeck and Renner's result is that the marginal probabilities in this case are $1/2$ (as shown in Section 3.1). But

the end result is that for every projective measurement the probabilities are independent of any supplemental variables like Z (as shown in Section 3.4.2).

We will now explicitly show where Di Lorenzo's result is in conflict with the result of Colbeck and Renner (or rather: our reformulation of it). We will consider one simple model from the class presented by Di Lorenzo and show step-by-step how this model is ruled out. This is a good opportunity to give a concrete example, which helps in understanding the results of Chapter 3.

The fact that Di Lorenzo considers models for a two-particle spin singlet state, where the measurement outcomes for both particles are strongly correlated, may cause some confusion. Colbeck and Renner's result is based on strong correlations for measurements on entangled systems. But we will not use the strong correlations of the singlet state to disprove Di Lorenzo's model. Instead, we consider an entanglement between the whole 2-particle system and another 2-particle system. The strong correlations between measurements on both 2-particle systems is what we will use to prove that Di Lorenzo's model is in conflict with CR's result.

The following is a simple model from the class of models presented by Di Lorenzo[1]:

$$\begin{aligned} |\Psi\rangle_S &= \frac{1}{\sqrt{2}}(|0\rangle_{S_1}|1\rangle_{S_2} - |1\rangle_{S_1}|0\rangle_{S_2}); \\ \tilde{Z} &= \{-1, +1\}; \\ P(Z=+1) &= P(Z=-1) = 1/2; \\ P(x_1, x_2|z)_{\vec{a},\vec{b}} &= \frac{1}{4}\left(1 - x_1 \cdot x_2 \left(\vec{a}\cdot\vec{b} - (1-(\vec{a}\cdot\vec{b})^2)\cdot\frac{z}{4}\right)\right). \end{aligned} \tag{5.2}$$

Here, the settings $\vec{a}$ and $\vec{b}$ are unit vectors corresponding to observables $\vec{a}\cdot\vec{\sigma}$ and $\vec{b}\cdot\vec{\sigma}$, where $\vec{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$ is a vector consisting of Pauli matrices. S_1 is a single-particle system ('particle 1') as is S_2 ('particle 2'), and S is the composite system $S_1 + S_2$. Note that we only consider the spins of the particle, not the positions. X_1 and X_2 are the outcomes of the measurements on particle 1 and particle 2, respectively.

We only have to consider a single measurement setting. An easy choice is $\vec{a} = \hat{z}$ and $\vec{b} = \hat{x}$, so that $\vec{a}\cdot\vec{b} = 0$, and the observables are simply σ_z and σ_x. In that case,

$$P(x_1, x_2|z) = \frac{1}{4}\left(1 + \frac{x_1 \cdot x_2 \cdot z}{4}\right) = \frac{1}{4} + \frac{x_1 \cdot x_2 \cdot z}{16}. \tag{5.3}$$

The probabilities given a value for Z are 3/16 or 5/16, while the probabilities when averaging over Z are all 1/4, which is in accordance with the quantum

[1] This model can be obtained from Di Lorenzo's class of models by choosing $\mu(\lambda) = \frac{1}{2}(\delta(\lambda - 1) + \delta(\lambda + 1))$, $G(\lambda) = \lambda/4$ in equation (5) and identifying $x_1 = \sigma$, $x_2 = \tau$.

mechanical prediction. In particular, we have

$$\begin{aligned}
P(X_1 = +1, X_2 = +1 | Z = +1) &= \frac{5}{16};\\
P(X_1 = +1, X_2 = -1 | Z = +1) &= \frac{3}{16};\\
P(X_1 = +1, X_2 = +1 | Z = -1) &= \frac{3}{16};\\
P(X_1 = +1, X_2 = -1 | Z = -1) &= \frac{5}{16}.
\end{aligned} \tag{5.4}$$

We will show that these probabilities lead to a contradiction when taking into account the measuring process.

The joint measurement of particles 1 and 2 can be described as a single projective measurement. To do this, we define an observable with the following eigenvectors and eigenvalues:

Outcomes	Eigenvector	Eigenvalue
$X_1 = +1, X_2 = +1$	$\|0\rangle_{S_1}\|0_x\rangle_{S_2}$	-1
$X_1 = +1, X_2 = -1$	$\|0\rangle_{S_1}\|1_x\rangle_{S_2}$	$+1$
$X_1 = -1, X_2 = +1$	$\|1\rangle_{S_1}\|0_x\rangle_{S_2}$	$+2$
$X_1 = -1, X_2 = -1$	$\|1\rangle_{S_1}\|1_x\rangle_{S_2}$	$+3$

where

$$\begin{aligned}
|0_x\rangle &:= \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle);\\
|1_x\rangle &:= \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle).
\end{aligned} \tag{5.5}$$

We will also write $|0_z 1_x\rangle_S$ for $|0\rangle_{S_1}|1_x\rangle_{S_2}$, and similarly for the other combinations. The observable for this projective measurement is then

$$O = -1 \cdot P(|0_z 0_x\rangle) + 1 \cdot P(|0_z 1_x\rangle) + 2 \cdot P(|1_z 0_x\rangle) + 3 \cdot P(|1_z 1_x\rangle). \tag{5.6}$$

Also, the singlet state $|\Psi\rangle_S$ can be written

$$\begin{aligned}
|\Psi\rangle_S &= \frac{1}{\sqrt{2}}\left(|0\rangle_{S_1} \otimes \frac{1}{\sqrt{2}}(|0_x\rangle_{S_2} - |1_x\rangle_{S_2}) - |1\rangle_{S_1} \otimes \frac{1}{\sqrt{2}}(|0_x\rangle_{S_2} + |1_x\rangle_{S_2})\right)\\
&= \frac{1}{2}(|0_z 0_x\rangle_S - |0_z 1_x\rangle_S - |1_z 0_x\rangle_S - |1_z 1_x\rangle_S).
\end{aligned} \tag{5.7}$$

Following the von Neumann measuring scheme (Section 3.4.1), during a measurement of the observable O the following state is created

$$\begin{aligned}
|\phi\rangle_{SD} = \frac{1}{2}(&|0_z 0_x\rangle_S|0_z 0_x\rangle_D - |0_z 1_x\rangle_S|0_z 1_x\rangle_D\\
&- |1_z 0_x\rangle_S|1_z 0_x\rangle_D - |1_z 1_x\rangle_S|1_z 1_x\rangle_D).
\end{aligned} \tag{5.8}$$

where D is (part of the) measuring device. In order to obtain a Schmidt decomposition with positive coefficients, we define

$$\begin{aligned}
|0_z 0'_x\rangle_D &:= |0_z 0_x\rangle_D;\\
|0_z 1'_x\rangle_D &:= -|0_z 1_x\rangle_D;\\
|1_z 0'_x\rangle_D &:= -|1_z 0_x\rangle_D;\\
|1_z 1'_x\rangle_D &:= -|1_z 1_x\rangle_D.
\end{aligned}$$

so that

$$|\phi\rangle_{SD} = \frac{1}{2}(|0_z 0_x\rangle_S |0_z 0'_x\rangle_D + |0_z 1_x\rangle_S |0_z 1'_x\rangle_D + |1_z 0_x\rangle_S |1_z 0'_x\rangle_D + |1_z 1_x\rangle_S |1_z 1'_x\rangle_D). \quad (5.9)$$

As explained in Section 3.4.2, the measurement on S can also be interpreted as a measurement on D in the corresponding basis

$$\{|0_z 0'_x\rangle_D, |0_z 1'_x\rangle_D, |1_z 0'_x\rangle_D, |1_z 1'_x\rangle_D\}, \quad (5.10)$$

with corresponding observable

$$-1 \cdot P(|0_z 0'_x\rangle) + 1 \cdot P(|0_z 1'_x\rangle) + 2 \cdot P(|1_z 0'_x\rangle) + 3 \cdot P(|1_z 1'_x\rangle). \quad (5.11)$$

This step can be made more explicit by considering an *indirect* measurement. The spin of particle 1 is coupled to another particle ($1'$) in the z-basis, while the spin of particle 2 is coupled to another particle ($2'$) in the x-basis. This also creates the state $|\phi\rangle_{SD}$, but now D is simply another system consisting of two particles. Now, particle $1'$ is measured in the z-basis, while particle $2'$ is measured in the x-basis. This is effectively a measurement of particles 1 and 2.

We now consider different measurements on S and D after the state $|\phi\rangle_{SD}$ has been created as part of the first measurement of S. Note that this first measurement has not been finished at this point. First defining

$$|\theta^{(\prime)}\rangle = \cos\frac{\theta}{2}|0_z 0_x{}^{(\prime)}\rangle + \sin\frac{\theta}{2}|0_z 1_x{}^{(\prime)}\rangle, \quad (5.12)$$

we consider measurements of the observables

$$\begin{aligned} O_v &= -1 \cdot P(|v\rangle) + 1 \cdot P(|v+\pi\rangle) + 2 \cdot P(|1_z 0_x\rangle) + 3 \cdot P(|1_z 1_x\rangle); \\ O_w &= -1 \cdot P(|w'\rangle) + 1 \cdot P(|w+\pi'\rangle) + 2 \cdot P(|1_z 0'_x\rangle) + 3 \cdot P(|1_z 1'_x\rangle), \end{aligned} \quad (5.13)$$

on S and D, respectively. The setting v (w) corresponds to the observable O_v (O_w). To the outcomes we associate random variables X and Y, respectively. As in Section 3.2, we look at the probability distribution conditioned on the outcomes being ± 1:

$$P_{|1|}(x,y)_{v,w} = P(x,y||X| = |Y| = 1)_{v,w} = \frac{1}{4}\left(1 + x \cdot y \cdot \cos(v-w)\right). \quad (5.14)$$

In the Di Lorenzo model, we have

$$\begin{aligned} P_{|1|}(Y=-1|Z=+1)_{W=0} &= \frac{5}{8}; \\ P_{|1|}(Y=+1|Z=+1)_{W=0} &= \frac{3}{8}; \\ P_{|1|}(Y=-1|Z=-1)_{W=0} &= \frac{3}{8}; \\ P_{|1|}(Y=+1|Z=-1)_{W=0} &= \frac{5}{8}. \end{aligned} \quad (5.15)$$

We will show that (5.14) and (5.15) are in contradiction. We define

$$\begin{aligned}
B_0 &= Y\big|_{W=0}; & A_1 &= X\big|_{V=1\cdot\frac{\pi}{10}};\\
B_2 &= Y\big|_{W=2\cdot\frac{\pi}{10}}; & A_3 &= X\big|_{V=3\cdot\frac{\pi}{10}};\\
B_4 &= Y\big|_{W=4\cdot\frac{\pi}{10}}; & A_5 &= X\big|_{V=5\cdot\frac{\pi}{10}};\\
B_6 &= Y\big|_{W=6\cdot\frac{\pi}{10}}; & A_7 &= X\big|_{V=7\cdot\frac{\pi}{10}};\\
B_8 &= Y\big|_{W=8\cdot\frac{\pi}{10}}; & A_9 &= X\big|_{V=9\cdot\frac{\pi}{10}}.
\end{aligned} \tag{5.16}$$

Note that $P_{|1|}(y)_{W=0} = P_{|1|}(b_0)$. Now,

$$\begin{aligned}
\frac{1}{4} &= \Big|P_{|1|}(Y=+1|Z=+1)_{W=0} - P_{|1|}(Y=-1|Z=+1)_{W=0}\Big|\\
&= \Big|P_{|1|}(B_0=+1|Z=+1) - P_{|1|}(\overline{B_0}=+1|Z=+1)\Big|\\
&\leq \Big|P_{|1|}(B_0=+1|Z=+1) - P_{|1|}(A_1=+1|Z=+1)\Big|\\
&+ \Big|P_{|1|}(A_1=+1|Z=+1) - P_{|1|}(B_2=+1|Z=+1)\Big|\\
&+ \ldots + \Big|P_{|1|}(A_9=+1|Z=+1) - P_{|1|}(\overline{B_0}=+1|Z=+1)\Big|\\
&\leq P_{|1|}(B_0\neq A_1|Z=+1) + \ldots + P_{|1|}(A_9\neq\overline{B_0}|Z=+1).
\end{aligned} \tag{5.17}$$

Similarly, for $Z=-1$:

$$\begin{aligned}
\frac{1}{4} &= \Big|P_{|1|}(Y=+1|Z=-1)_{W=0} - P_{|1|}(Y=-1|Z=-1)_{W=0}\Big|\\
&\leq P_{|1|}(B_0\neq A_1|Z=-1) + \ldots + P_{|1|}(A_9\neq\overline{B_0}|Z=-1).
\end{aligned} \tag{5.18}$$

Averaging over Z, i.e. multiplying (5.17) and (5.18) by $P(Z=+1) = P(Z=-1) = 1/2$ and adding the two, we get

$$\frac{1}{4} \leq P_{|1|}(B_0\neq A_1) + \ldots + P_{|1|}(A_9\neq\overline{B_0}) = I_5 = 10\cdot\sin^2\frac{\pi}{10\cdot 2} \approx 0.245, \tag{5.19}$$

which is false.

Starting from a model taken from the class proposed by Di Lorenzo, we reached a contradiction. Therefore, when taking into account the measurement scheme, this model cannot be in agreement with quantum mechanics. Although in this chapter we have just shown the conflict of a single model of Di Lorenzo's class, the general result of Chapter 3 shows that all models must be in conflict, except for trivial ones, in which probabilities are independent of any supplemental variables.

We did not need to use embezzlement in this example, because the Schmidt coefficients of $|\phi\rangle_{SD}$ were already all equal.

Chapter 6

Conclusion and discussion

In this chapter, the results of this thesis are summarized. Furthermore, we discuss some possible issues and suggestions for improvement, and some ideas for further research.

6.1 Overview of results

The main result is that we derived a no-go theorem for non-trivial QM-compatible hidden variable theories satisfying Parameter Independence. This result is not really new: it is a reformulation of the work of Colbeck and Renner, who presented it as a theorem saying that no extension of quantum theory can have improved predictive power. We have explained step-by-step the most important ingredient of the result: the strong correlation in measurements on maximally entangled states. Furthermore, we have included a complete description and explanation of the embezzlement transformation, as well as a proof that this transformation can be done with arbitrary precision. By evaluating critically the work of Colbeck and Renner, we have stumbled upon some interesting points. We have seen that in CR's assumption of 'Freedom of Choice', a no-signaling assumption is hidden in the definition. Another point is that variables which describe the context of an experiment should generally not be treated as random variables. Also, we have seen that interesting questions arise when trying to combine this result with interpretations of quantum mechanics. It seems that in no-collapse interpretations like the modal interpretation, Parameter Independence must somehow be violated if in addition to the state vector extra properties are assigned to systems which tell something about measurement outcomes. Finally, we have explicitly demonstrated a conflict of recent work of Antonio Di Lorenzo with the result.

6.2 Possible issues and improvements

6.2.1 Application to more general types of measurement

In this thesis we have only discussed non-degenerate projective measurements. We have not extended the result to degenerate measurements, although we guess that this is an easy extension of the result. Another question is whether the

result also holds for the more general class of POVM measurements[1]. According to CR, this is indeed the case since according to Naimark's theorem any POVM measurement can be described as a projective measurement on a larger Hilbert space[2]. But, what seems necessary then is that every POVM measurement *is* in fact a projective measurement on a larger Hilbert space. This is a point that deserves further investigation.

Also, in the generalization to arbitrary projective measurements, we have used the von Neumann measurement scheme. This scheme is highly idealized, and real measurements are usually only approximations of this. For example, the object system might be disturbed more by the measurement than is described by the von Neumann scheme, or even destroyed. Consider the measurement of a photon: the photon is destroyed when it is detected. There are other issues: the initial state before the measurement might be something different than a pure product state of the system to be measured and the apparatus. There is an extensive literature on this subject, and it would be interesting to investigate how the result relates to more realistic measurements.

6.2.2 Relation to experimental practice

In the derivations of Theorems 1-4, we have made use of measurements which are easy to write down, but hard to perform in reality. For example, to get a value of I_N close to zero, measurements are considered for which the separation between settings is very small. Furthermore, embezzling states have been used, which have never been created in reality. We have described the embezzlement transformation using unitary operators, but we have not suggested how such transformations can be implemented in practice. This is a trait of the field of quantum computation, where it is usually assumed that any unitary operation can be performed without worrying about the physical implementation of the operation. However, if quantum mechanics is correct, all those steps can be justified by stating that those measurements and transformations are at least possible in principle.

Knowing one I_N means not knowing them all

I_N, definined in (2.36), is a quantity that can be measured for any system which is in a maximally entangled of Schmidt order 2. It imposes a constraint on the probability distributions of any theory satisfying Parameter Independence, even if quantum mechanics is not correct. Experiments have been performed to find an I_N as low as possible [49]. However, the power of this theory-independent approach is limited, since there is no guarantee that I_N can be assumed to be that low in other experimental arrangements, using other systems.

For example, we might do an EPRB-type experiment with photons and find an upper bound for I_N of 0.01. This does not say what the upper bound of I_N applies for an EPRB-type experiment with electrons. While quantum mechanics predicts that I_N goes to zero in the limit $N \to \infty$, this does not mean that in other theories the value of I_N is the same in different types of experiments.

[1] See for example [40], Section 2.2.6

[2] [20], p. 4

6.2.3 Generalization to continuous hidden variables

For simplicity, we limited our class of hidden variable theories to theories which assign a single discrete hidden variable to a system. However, the generalization to continuous hidden variables should not be too hard. All probability distributions involving the hidden variable Z would then be replaced by probability density functions, and sums over Z should be replaced by integrals. The results would be reformulated a bit: instead of holding for all hidden variables, they would hold for all hidden variables except a measure-zero set. For example, Theorem 1 has been proved for continuous hidden variables in [12].

6.3 Outlook on future research

6.3.1 Relation between quantum states and reality

Colbeck and Renner have used their result to derive that there is a one-to-one correspondence between quantum states and 'its elements of reality', which is supposed to be some real underlying state [19]. We are skeptical about this result, since this is a claim about ontology, while CR adhere to an operational approach. Furthermore, something like Parameter Independence might not hold at the fundamental level of reality. And, the relation between CR's result and the different interpretations of quantum mechanics, discussed in Section 4.4, is also relevant to the question how quantum states relate to reality. Although we think that a one-to-one correspondence between quantum states and 'real' states cannot be derived directly from CR's result, there is a lot to investigate here.

6.3.2 Relation to the probabilistic framework

The framework of general probabilistic theories has been used to treat quantum mechanics as a probabilistic theory in a vast landscape of possible probabilistic theories, which also include classical mechanics and 'superquantum' theories which allow for correlations stronger than those in quantum mechanics (see for example [4]). Using this framework, the question why quantum theory has this particular place in the landscape can be investigated. It would be interesting to connect the result of this thesis to this framework. For example, in this framework the result might be stated as follows: quantum theory cannot be replaced by a non-signaling superquantum theory where the pure quantum states are mixtures of states in this higher theory. Another interesting question is whether the result of this thesis can be generalized to arbitrary theories in the probabilistic framework: can we distinguish between those theories that can be extended and those that cannot be extended?

6.4 Acknowledgements

I would like to thank Dennis Dieks for supervising this project, and for showing me the paper of Colbeck and Renner [20], which sparked my interested for this thesis subject. I would also like to thank my fellow master students for the good times we spent. Furthermore, I would like to thank Michiel Seevinck, Klaas

Landsman and Daan Hendrickx for inspiring discussions, and Roger Colbeck for answering my questions about his work. Finally, I want to thank Ilmer Rozendaal, my parents Jan and Tineke and my sisters Jeannette and Elisabeth for their love and support.

Appendix A

Quantum mechanics

A.1 Schmidt decomposition

Let $|\phi\rangle_{AB} \in \mathcal{H}_A \otimes \mathcal{H}_B$ be the state of a composite system $A + B$. Then $|\phi\rangle_{AB}$ can be written as

$$|\phi\rangle_{AB} = \sum_i \lambda_i |i\rangle_A |i\rangle_B, \tag{A.1}$$

with λ_i non-negative, $\sum_i \lambda_i^2 = 1$ and $|i\rangle_A, |i\rangle_B$ orthonormal sets of vectors. This representation is called the Schmidt decomposition. The numbers λ_i are called the Schmidt coefficients. The number of positive λ_i is called the Schmidt number of $|\phi\rangle_{AB}$. $|i\rangle_A$ and $|i\rangle_B$ are called Schmidt bases. The Schmidt basis $|i\rangle_A(|i\rangle_B)$ may have less elements than the dimension of $\mathcal{H}_A(\mathcal{H}_B)$. In that case, it is not a basis of $\mathcal{H}_A(\mathcal{H}_B)$, but of the subspace spanned by it. The set of numbers λ_i is unique (except for the number of zeroes), and if all λ_i are different then the decomposition is unique (of course, up to ordering of the vectors $|i\rangle_A|i\rangle_B$).

If at least two λ_i are equal, then the Schmidt decomposition is not unique. We can use the symmetry property explained in Section A.2 to find other decompositions. Take a set of orthonormal vectors $\{|j\rangle_A|j\rangle_B\}$ for which the Schmidt coefficients are equal. Then, for any basis $|j'\rangle_A$ of the subspace spanned by $|j\rangle_A$, there is a basis $|j'\rangle_B$ for the subspace spanned by $|j\rangle_B$ such that

$$\sum_j |j\rangle_A |j\rangle_B = \sum_j |j'\rangle_A |j'\rangle_B. \tag{A.2}$$

So, we can obtain another Schmidt decomposition by replacing the vectors $|j\rangle_A|j\rangle_B$ by $|j'\rangle_A|j'\rangle_B$.

The Schmidt decomposition was first proved in [45]. A modern version can be found in [40].

A.1.1 Perfect correlation when measuring in Schmidt bases

Suppose Alice and Bob share the state $|\phi\rangle_{AB}$. Consider the Schmidt decomposition (A.1). Consider projective measurements on the state, where Alice performs her measurement in the Schmidt basis $|i\rangle_A$ while Bob performs his in

the Schmidt basis $|i\rangle_B$. Calling Alice's eigenvalues e_i and Bob's eigenvalues f_i, the probability of getting the pair of outcomes (e_j, f_k) is

$$|(\langle j|\langle k|)|\phi\rangle|^2 = \left|\sum_i \lambda_i \langle j|i\rangle\langle k|i\rangle\right|^2 = \lambda_j^2 \delta_{jk}. \tag{A.3}$$

We see that there is perfect correlation: Alice obtains the value e_j if and only if Bob obtains the value f_j.

A.2 Symmetry property of maximally entangled states

If $|\Psi\rangle_{AB} \in \mathcal{H}_A \otimes \mathcal{H}_B$ is a maximally entangled state in a D^2-dimensional Hilbert space of Schmidt order D, then for any basis $|i'\rangle_A$ of $\mathcal{H}_A$ there is a basis $|i'\rangle_B$ of $\mathcal{H}_B$ such that

$$|\Psi\rangle_{AB} = \frac{1}{\sqrt{D}} \sum_{i=1}^{D} |i'\rangle_A \otimes |i'\rangle_B. \tag{A.4}$$

This guarantees that there is maximal correlation for any projective measurement on A: there is always a measurement on B which gives the same outcome.

Proof
By definition, we have

$$|\Psi\rangle_{AB} = \frac{1}{\sqrt{D}} \sum_{i=1}^{D} |i\rangle_A \otimes |i\rangle_B. \tag{A.5}$$

Now, let $|i'\rangle_A$ be any basis of $\mathcal{H}_A$. Let U be the unitary operator which transforms the vector $|i\rangle_A$ to $|i'\rangle_A$. This means $U_{ij}|j\rangle_A = |i'\rangle_A$. Now define $|i'\rangle_B := U^*_{ij}|j\rangle_B$. Note that U^* is also unitary since $U^{*\dagger}U^* = (U^\dagger U)^* = \mathbb{I}^* = \mathbb{I}$. Therefore, $|i'\rangle_B$ is a basis of $\mathcal{H}_B$. Furthermore,

$$\frac{1}{\sqrt{D}} \sum_{j=1}^{D} |j'\rangle_A \otimes |j'\rangle_B = \frac{1}{\sqrt{D}} \sum_{jkl} U_{jk}|k\rangle_A \otimes U^*_{jl}|l\rangle_B \tag{A.6}$$

$$= \frac{1}{\sqrt{D}} \sum_{kl} \delta_{kl}|k\rangle_A \otimes |l\rangle_B \tag{A.7}$$

$$= \frac{1}{\sqrt{D}} \sum_{k} |k\rangle_A \otimes |k\rangle_B = |\Psi\rangle_{AB}. \tag{A.8}$$

Here we have used

$$\sum_j U_{jk} U^*_{jl} = \delta_{kl}, \tag{A.9}$$

which follows from the fact that the columns of a unitary matrix form an orthonormal basis.

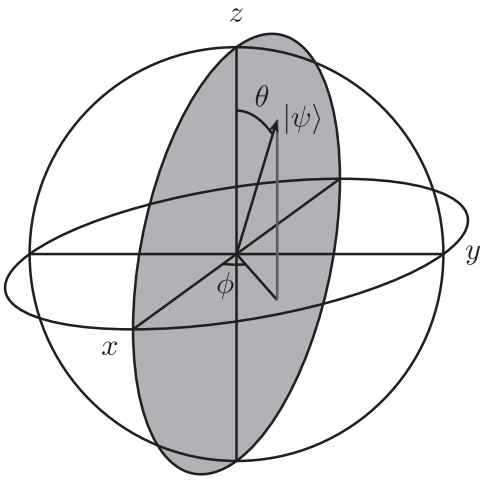

Figure A.1: The Bloch Sphere

A.3 Bloch Sphere

The Bloch sphere is a convenient representation of the qubit state space. Since the global phase of a quantum state is not physical, we can write any pure qubit state as

$$|\psi\rangle = \cos\frac{\theta}{2}|0\rangle + e^{i\phi}\sin\frac{\theta}{2}|1\rangle. \tag{A.10}$$

Such a state corresponds to a point on the surface of a unit sphere with spherical coordinates θ, ϕ and $r = 1$: This sphere is called the *Bloch sphere*, after Felix Bloch.

The circle intersecting the xz-plane, which is colored gray, corresponds to qubit states with real coefficients: the states $|\theta\rangle$ we defined in equation (2.2).

Bibliography

[1] Arntzenius, F., Reichenbach's Common Cause Principle, in E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy*, Fall 2010 edition (2010), URL: `http://plato.stanford.edu/archives/fall2010/entries/physics-Rpcc/`

[2] Aspect, A., Dalibard, J. and Roger, G., Experimental test of Bell's inequalities using time-varying analyzers, Phys. Rev. Lett., **49**, 1804 (1982)

[3] Atmanspacher, H., Quantum approaches to consciousness, in E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy*, summer 2011 edition (2011), URL: `http://plato.stanford.edu/archives/sum2011/entries/qt-consciousness/`

[4] Barrett, J., Information processing in generalized probabilistic theories, Phys. Rev. A, **75**, 032304 (2007)

[5] Bell, J. S., On the Einstein-Podolsky-Rosen paradox, Physics, **1**, 195 (1964)

[6] Bell, J. S., The theory of local beables, in *Speakable and Unspeakable in Quantum Mechanics*, Cambridge University Press (2004)

[7] Bennett, C. H., Brassard, G., Crépeau, C., Jozsa, R., Peres, A. and Wootters, W. K., Teleporting an unknown quantum state via dual classical and Einstein-Podolsky-Rosen channels, Phys. Rev. Lett., **70**, 1895 (1993)

[8] Bohm, D., A suggested interpretation of the quantum theory in terms of "hidden" variables. I and II, Physical Review, **85**, 166 (1952)

[9] Bohm, D. and Aharonov, Y., Discussion of experimental proof for the paradox of Einstein, Rosen, and Podolsky, Phys. Rev., **108**, 1070 (1957)

[10] Bohm, D. and Hiley, B., The de Broglie pilot wave theory and the further development of new insights arising out of it, Foundations of Physics, **12**, 1001 (1982)

[11] Born, M. and Einstein, A., *The Born-Einstein Letters*, Palgrave Macmillan (2004)

[12] Branciard, C., Brunner, N., Gisin, N., Kurtsiefer, C., Lamas-Linares, A., Ling, A. and Scarani, V., Testing quantum correlations versus single-particle properties within Leggett's model and beyond, Nature Physics, **4**, 681 (2008)

[13] Braunstein, S. L. and Caves, C. M., Wringing out better Bell inequalities, Nuclear Physics B - Proceedings Supplements, **6**, 211 (1989)

[14] Butterfield, J., Bell's theorem: What it takes, The British Journal for the Philosophy of Science, **43**, 41 (1992)

[15] Cirel'son, B., Quantum generalizations of Bell's inequality, Letters in Mathematical Physics, **4**, 93 (1980)

[16] Clauser, J. F., Horne, M. A., Shimony, A. and Holt, R. A., Proposed experiment to test local hidden-variable theories, Phys. Rev. Lett., **23**, 880 (1969)

[17] Colbeck, R., Frequently asked (and answered) questions for "Quantum theory cannot be extended", URL: `http://perimeterinstitute.ca/personal/rcolbeck/FAQ.html`, retrieved on August 13, 2013 (2010)

[18] Colbeck, R. and Renner, R., No extension of quantum theory can have improved predictive power, Nat Commun, **2**, 411 (2011)

[19] Colbeck, R. and Renner, R., Is a system's wave function in one-to-one correspondence with its elements of reality?, Phys. Rev. Lett., **108**, 150402 (2012)

[20] Colbeck, R. and Renner, R., The completeness of quantum theory for predicting measurement outcomes, arXiv: 1208.4123 (2012)

[21] de Oliveira Oliveira, M., Embezzlement States are Universal for Non-Local Strategies, arXiv: 1009.0771 (2010)

[22] Di Lorenzo, A., Determination of hidden-variable models reproducing the spin singlet, Phys. Rev. A, **86**, 042119 (2012)

[23] Dieks, D., Communication by EPR devices, Physics Letters A, **92**, 271 (1982)

[24] Dieks, D. and Vermaas, P. E. (Eds.), *The modal interpretation of quantum mechanics*, The University of Western Ontario Series in Philosophy of Science, Kluwer Academic Publishers, Dordrecht (1998)

[25] Einstein, A., Über einen die Erzeugung und Verwandlung des Lichtes betreffenden heuristischen Gesichtspunkt, Annalen der Physik, **322**, 132 (1905)

[26] Einstein, A., Podolsky, A. and Rosen, N., Can quantum-mechanical description of physical reality be considered complete?, Phys. Rev., **47**, 777 (1935)

[27] Ekert, A. K., Quantum cryptography based on Bell's theorem, Phys. Rev. Lett., **67**, 661 (1991)

[28] Feynman, R., Simulating physics with computers, International Journal of Theoretical Physics, **21**, 467 (1982)

[29] Fine, A., Hidden variables, joint probability, and the Bell inequalities, Phys. Rev. Lett., **48**, 291 (1982)

[30] Ghirardi, G., Rimini, A. and Weber, T., A model for a unified quantum description of macroscopic and microscopic systems, in L. Accardi and W. Waldenfels (Eds.), *Quantum Probability and Applications II*, *Lecture Notes in Mathematics*, volume 1136, 223–232, Springer Berlin Heidelberg (1985)

[31] Ghirardi, G. and Romano, R., About possible extensions of quantum theory, Foundations of Physics, **43**, 881 (2013)

[32] Ghirardi, G. and Romano, R., Ontological models predictively inequivalent to quantum theory, Phys. Rev. Lett., **110**, 170404 (2013)

[33] Ghirardi, G. C. and Romano, R., On the completeness of quantum mechanics and the interpretation of the state vector, Journal of Physics: Conference Series, **442**, 012002 (2013)

[34] Grover, L., Quantum mechanics helps in searching for a needle in a haystack, Phys. Rev. Lett., **79**, 325 (1997)

[35] Hales, L. and Hallgren, S., An improved quantum fourier transform algorithm and applications, in *Proceedings of the 41st Annual Symposium on Foundations of Computer Science*, 515–525 (2000)

[36] Hardy, L., Reconstructing quantum theory, arXiv: 1303.1538 (2013)

[37] Hilgevoord, J., *Grondslagen van de Quantummechanica*, Lecture notes, Utrecht University (2004)

[38] Landauer, R., Information is physical, Physics Today, **44**, 23 (1991)

[39] Larsson, J.-A. and Cabello, A., Hidden variable models for quantum mechanics can have local parts, arXiv: 0907.2619 (2009)

[40] Nielsen, M. A. and Chuang, I. L., *Quantum Computation and Quantum Information*, Cambridge University Press (2004)

[41] Pearle, P. M., Hidden-variable example based upon data rejection, Phys. Rev. D, **2**, 1418 (1970)

[42] Pusey, M. F., Barrett, J. and Rudolph, T., On the reality of the quantum state, Nat Phys, **8**, 475 (2012)

[43] Reichenbach, H., *The Direction of Time*, University of Califonia Press, Berkeley (1956)

[44] Saunders, S., Barrett, J., Kent, A. and Wallace, D. (Eds.), *Many worlds? Everett, quantum theory and reality.*, Oxford University Press (2010)

[45] Schmidt, E., Zur Theorie der linearen und nichtlinearen Integralgleichungen. III. Teil, Mathematische Annalen, **65**, 370 (1908)

[46] Seevinck, M. P., *Parts and Wholes. An Inquiry into Quantum and Classical Correlations*, Ph.D. thesis, Utrecht University (2008), arXiv: 0811.1027

[47] Shimony, A., Controllable and uncontrollable non-locality, in *Search for a naturalistic world view*, volume 2, Cambridge University Press (1993)

[48] Shor, P., Algorithms for quantum computation: Discrete logarithms and factoring, in *Proceedings, 35th Annual Symposium on Foundations of Computer Science*, IEEE Computer Society Press (1994)

[49] Stuart, T. E., Slater, J. A., Colbeck, R., Renner, R. and Tittel, W., Experimental bound on the maximum predictive power of physical theories, Phys. Rev. Lett., **109**, 020402 (2012)

[50] van Dam, W. and Hayden, P., Universal entanglement transformations without communication, Phys. Rev. A, **67**, 060302 (2003)

[51] von Neumann, J., *Mathematische Grundlagen der Quanten-mechanik*, Verlag Julius-Springer, Berlin (1932)

[52] Wootters, W. K. and Zurek, W. H., A single quantum cannot be cloned, Nature, **299**, 802 (1982)