

# An impossibility theorem for Parameter Independent hidden variable theories

Gijs Leegwater\*

March 31, 2016

## Abstract

Recently, Roger Colbeck and Renato Renner (C&R) have claimed that '[n]o extension of quantum theory can have improved predictive power' (Colbeck & Renner, 2011, 2012b). If correct, this is a spectacular impossibility theorem for hidden variable theories, which is more general than the theorems of Bell (1964) and Leggett (2003). Also, C&R have used their claim in attempt to prove that a system's quantum-mechanical wave function is in a one-to-one correspondence with its 'ontic' state (Colbeck & Renner, 2012a). C&R's claim essentially means that in any hidden variable theory that is compatible with quantum-mechanical predictions, probabilities of measurement outcomes are independent of these hidden variables. This makes such variables otiose. On closer inspection, however, the generality and validity of the claim can be contested. First, it is based on an assumption called 'Freedom of Choice'. As the name suggests, this assumption involves the independence of an experimenter's choice of measurement settings. But in the way C&R define this assumption, a no-signalling condition is surreptitiously presupposed, making the assumption less innocent than it sounds. When using this definition, any hidden variable theory violating Parameter Independence, such as Bohmian Mechanics, is immediately shown to be incompatible with quantum-mechanical predictions. Also, the argument of C&R is hard to follow and their mathematical derivation contains several gaps, some of which cannot be closed in the way they suggest. We shall show that these gaps can be filled. The issue with the 'Freedom of Choice' assumption can be circumvented by explicitly assuming Parameter Independence. This makes the result less general, but better founded. We then obtain an impossibility theorem for hidden variable theories satisfying Parameter Independence only. As stated above, such hidden variable theories are impossible in the sense that any supplemental variables have no bearing on outcome probabilities, and are therefore trivial. So, while quantum mechanics itself satisfies Parameter Independence, if a variable is added that changes the outcome probabilities, however slightly, Parameter Independence must be violated.

---

\*Erasmus University Rotterdam, Faculty of Philosophy, Burg. Oudlaan 50, 3062 PA Rotterdam, The Netherlands. E-mail: leegwater@fwb.eur.nl

## 1 Introduction

In 1935, Einstein, Podolsky and Rosen famously argued that quantum mechanics is incomplete and that there might be another theory that does provide a complete description of physical reality (Einstein et al., 1935). One class of candidates for such a theory is the class of so-called ‘hidden variable theories’, which supplement the quantum state with extra variables.<sup>1</sup> Hidden variable theories have indeed been developed, for example the de Broglie–Bohm theory (Bohm, 1952), which is deterministic and complete in Einstein’s sense. However, a number of impossibility theorems have been derived, showing that large classes of possible hidden variable theories are incompatible with quantum-mechanical predictions. John Bell proved such incompatibility for local deterministic hidden variable theories (Bell, 1964), as well as for local stochastic hidden variable theories (Bell, 2004), while the incompatibility of ‘crypto-nonlocal’ theories was proven by Leggett (2003). Still, a large class of hidden variable theories, like the de Broglie–Bohm theory, remains unscathed by these impossibility theorems.

The hidden variable theories shown to be incompatible by Bell are theories satisfying a criterion called Factorizability (Fine, 1982), which is equivalent to the conjunction of two locality conditions coined by Abner Shimony (1984): Parameter Independence (ParInd) and Outcome Independence (OutInd). Therefore, any hidden variable theory compatible with quantum mechanics violates at least one of these two conditions. In this article we claim something stronger: any hidden variable theory compatible with quantum mechanics violates ParInd, except for ‘trivial’ hidden variable theories, where the values of the hidden variables have no bearing on measurement outcome probabilities.

This article is based on recent work by Roger Colbeck and Renato Renner (C&R), who have claimed that they have derived an even more general impossibility theorem (Colbeck & Renner, 2011, 2012b). Stating that ‘[n]o extension of quantum theory can have improved predictive power’, they essentially claim that *any* non-trivial hidden variable theory, also if it violates ParInd (like the de Broglie–Bohm theory), is incompatible with quantum-mechanical predictions. Given the wide scope of this claim, this would be a spectacular result, which would to a great extent put constraints on any possible future theory replacing quantum mechanics.

However, C&R’s claim crucially hinges on an assumption dubbed ‘Freedom of Choice’. As the name suggests, this assumption is meant to be about the freedom of experimenters to choose their measurement settings. From this assumption, C&R derive ‘no-signalling’, which is essentially equal to ParInd. Nevertheless, when inspecting the way ‘Freedom of Choice’ is defined, it becomes apparent that ParInd is in fact part of this assumption. Most criticism of C&R’s work focuses on this issue (Ghirardi & Romano, 2013a,b; Colbeck & Renner, 2013; Leifer, 2014; Landsman, 2015). We agree with the criticism: C&R’s ‘Freedom of Choice’ assumption is much stronger than its name suggests. Therefore, while the impression is given that any hidden variable theory with free experimenters is shown to be incompatible with QM, in fact their result applies to a smaller class of hidden variable theories: those satisfying ParInd. The de Broglie–Bohm theory, which violates ParInd, is therefore not shown to be in-

---

<sup>1</sup>In the literature, sometimes the term ‘hidden variable theory’ is understood to refer only to deterministic theories. Instead, we have a very general use of the term in mind: theories that add, in addition to the quantum state, an extra variable to the description of a system.

compatible after all.

If the above issue was the only problem with C&R's work, the result of the present article could easily be achieved by adding ParInd as an explicit assumption. The theorem would then still be an interesting impossibility theorem, being more general than the theorems of Bell and Leggett. However, there are more shortcomings in the work of C&R. First, it is hard to understand, even for experts. Valerio Scarani (2013) says:

‘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 [(Colbeck, 2010)] (extremely rare for a scientific paper) and a later, clearer version [(Colbeck & Renner, 2012b)].’

The case of the maximally entangled state that Scarani refers to corresponds to the triviality claim of C&R restricted to local measurements on a Bell state. This result consists of the statement that not only the quantum-mechanical outcome probabilities, but also the outcome probabilities in any hidden variable theory equal  $1/2$  (in the present paper, this result is presented in Section 4). Some authors, for example Antonio Di Lorenzo (2012), appear to have understood C&R as deriving only this result, which, as Scarani alludes to, had been derived before. Actually, for C&R this is only the first step in proving the more general theorem that probabilities in hidden variable theories are always equal to the quantum-mechanical probabilities.

More importantly, C&R's derivation contains gaps, of which some are allegedly filled in other publications, while others remain. One example is their careless handling of limits: in more than one occasion results are derived that only hold approximately, which are then used as if they hold exactly.<sup>2</sup>

Because of these shortcomings, at present no acceptable deduction of the impossibility theorem for hidden variable theories satisfying ParInd exists in the literature. In this article we attempt to repair the shortcomings of C&R's derivation in order to establish such a deduction. A step that is not explicitly mentioned by C&R, involving the relation between measurements on entangled states and measurements on non-entangled states, is formulated explicitly. Furthermore, we give a deduction that is mathematically acceptable. We emphasize that this does not consist of simply filling some gaps. For some parts of the deduction to succeed, an entirely new strategy has to be constructed, or so we claim. This is especially the case when taking proper care of all the limits used in the proof. Also, some parts of the deduction can, in our opinion, be considerably simplified, especially the first steps. For these reasons, in this article we do not merely point out all the shortcomings in the original derivation; rather, we construct a new version of it.

C&R have also used their claim in an attempt to answer the question whether the quantum-mechanical wave function is ‘ontic’ or ‘epistemic’. Since the appearance of the Pusey-Barret-Rudolph (PBR) theorem (Pusey et al., 2012), this is a hotly debated topic. On the basis of their claim, C&R argue not only that the wave function is ontic, but also that it is in a one-to-one correspondence with its ontic state (Colbeck & Renner, 2012a). In the Discussion (Section 9), we

<sup>2</sup>Section 9 contains a more detailed treatment of these gaps.

shall consider what remains of this  $\psi$ -ontology result if C&R's claim is replaced by the weaker result deduced in this article.

The result will be deduced in several steps. The first steps are quite simple and correspond to results that existed already before the work of C&R. However, we believe that even for those whom are already familiar with this result, these steps are still of value since they are considerably simplified, only using a triangle inequality and a simple inequality from probability theory. The final steps require more mathematics and may be harder to follow. Most of the mathematics is relegated to appendices, so as not to distract the reader from the central line of reasoning. If the reader want to shorten the reading time, the best section to skip might be Section 7, because the extent of the generalization (from states with coefficients that are square roots of rational numbers to any coefficients) is relatively small compared to the amount of mathematics needed. It is however a necessary part for deriving the full theoretical result. In the Discussion, I shall mention the most important differences between our deduction and that of C&R.

## 2 Notation

Quantum-mechanical systems are referred to by the symbols  $A, B, A', B'$  etc. To denote composite systems, the symbols for the subsystems are combined, for example  $AB$  and  $AA'A''$ . The Hilbert space of system  $A$  is denoted by  $\mathcal{H}_A$ , a state as  $|\psi\rangle_A$  and an operator on  $\mathcal{H}_A$  as  $U^A$ . For notational convenience, the subscript attached to a state may be omitted when no confusion is possible, especially when large composite systems like  $AA'A''BB'B''$  are involved. The symbol  $\otimes$  for taking tensor products is also often omitted, and we freely change the order of states when combining systems, so that we can write

$$U^{AB} (|i\rangle_A \otimes |j\rangle_B) \equiv U^{AB} (|j\rangle_B \otimes |i\rangle_A) \equiv U^{AB} (|j\rangle_B |i\rangle_A). \quad (2.1)$$

We also write  $[\psi]^A := |\psi\rangle_{AA}\langle\psi|$ , and  $\mathbb{N}_r$  for the set  $\{0, 1, \dots, r-1\}$ . The cardinality of a set  $J$  is written as  $\#J$ , and sequences are notated as  $(x_n)_{n=0}^{10}$ .

In this article, mainly projective measurements with a finite number of outcomes are considered. Such measurements are defined by a complete set of orthogonal projectors  $\{\hat{E}_i^A\}_{i=0}^{d-1}$  with  $d \in \mathbb{N}$ , each projector corresponding to a possible outcome. If all projectors are 1-dimensional, the set of projectors can be written as  $\{[i]^A\}_{i=0}^{d-1}$ , in which case it is said that the measurement is performed 'in the basis  $\{|i\rangle_A\}_{i=0}^{d-1}$ '. We also allow ourselves to say this when  $\{|i\rangle_A\}_{i=0}^{d-1}$  is not a basis of  $\mathcal{H}_A$ , but of a strict subspace of  $\mathcal{H}_A$ . In this case the corresponding complete set of orthogonal projectors is  $\{[i]^A\}_{i=0}^{d-1} \cup \{\mathbb{I}^A - \sum_{i=0}^{d-1} [i]^A\}$ . Equivalently, a projective measurement can be characterized by an observable (a Hermitian operator) of which the eigenspaces corresponding to the eigenvalues equal the ranges of the projectors  $\{\hat{E}_i^A\}_{i=0}^{d-1}$ . Two observables which are equal up to their eigenvalues represent the same measurement. So, if a measurement is characterized by the complete set of orthogonal projectors  $\{\hat{E}_i^A\}_{i=0}^{d-1}$ , a corresponding observable is

$$\hat{O}^A = \sum_{i=0}^{d-1} e_i \hat{E}_i^A, \quad (2.2)$$

where the  $e_i \in \mathbb{R}$  are (distinct) eigenvalues. Probabilities of outcomes can be expressed using the projectors:

$$\Pr^{|\psi\rangle_A}(\hat{E}_i^A) = \langle \psi |_A \hat{E}_i^A | \psi \rangle_A. \quad (2.3)$$

The superscript including the state of the system may be omitted if no confusion is possible. Also, sometimes the superscript on the observable is omitted, for example when a general form of an observable is defined which can be applied to multiple systems.

Often measurements are considered on subsystems, and the pure state of the composite system is specified:

$$\Pr^{|\phi\rangle_{AB}}(\hat{E}_i^A). \quad (2.4)$$

To express probabilities using observables, we associate to any observable  $\hat{O}^A$  a random variable  $O^A$ . We want to emphasize that we only consider joint distributions of random variables if their associated observables are jointly measurable (and therefore commute). It is well known that if one defines joint distributions for non-commuting observables, then the corresponding random variables are subject to additional constraints, in the form of a Bell inequality (Fine, 1982), which we want to avoid. With this in place, probabilities can be expressed using the random variables associated with observables:

$$\Pr(O^A = e_i) = \Pr(\hat{E}_i^A). \quad (2.5)$$

We will see in the next section that the hidden variable theories we consider lead to a decomposition of the quantum probabilities. That is, an extra variable  $\lambda$  is added to each quantum probability, and there is a measure  $\mu(\lambda)$  such that, when averaging over  $\lambda$  using this measure, the quantum probability is retrieved. The probabilities in hidden variable theories and in a decomposition are called  $\lambda$ -probabilities. In decompositions, these probabilities are denoted as the quantum probabilities, with  $\lambda$  added as a subscript:

$$\Pr_\lambda(O^A = e_i) = \Pr_\lambda(\hat{E}_i^A)_{\hat{O}^A}. \quad (2.6)$$

In the notation using projectors, we have added an extra subscript indicating the measured observable. This is because a  $\lambda$ -probability might be *contextual*, depending not only on the projector  $\hat{E}_i^A$ , but also on the other projectors characterizing the measurement.<sup>3</sup>

Identities involving  $\lambda$ -probabilities appearing in this article mostly hold ‘almost everywhere’, i.e. for all  $\lambda$  in a subset  $\Omega \subset \Lambda$  with  $\mu(\Omega) = 1$ , where  $\mu$  is a measure on the measurable space  $\Lambda$ . Wherever this is the case, the symbol  $\doteq$  is used. Often we will consider  $\lambda$ -probabilities, expressed using a projector, that are almost everywhere independent of the observable that is measured. For example, we might have<sup>4</sup>

$$\forall \hat{O}^A : \Pr_\lambda(\hat{E}_i^A)_{\hat{O}^A} \doteq \frac{1}{2}. \quad (2.7)$$

<sup>3</sup>Note that this is a specific type of contextuality which is, for example, different from the contextuality considered in the Kochen–Specker Theorem (Kochen & Specker, 1975), which concerns a dependence of *values*, instead of probabilities, on the measurement context.

<sup>4</sup>The quantifier in (2.7) ranges over all observables that include  $\hat{E}_i^A$  in their sets of corresponding projectors

In such cases, we allow ourselves to drop the observable from the notation, so the above can be rewritten as

$$\Pr_\lambda \left( \hat{E}_i^A \right) \doteq \frac{1}{2}, \quad (2.8)$$

although  $\Pr_\lambda \left( \hat{E}_i^A \right)$  might not be well-defined for a measure zero subset of  $\lambda$ 's.

A  $\lambda$ -probability is called trivial if it equals the corresponding quantum-mechanical probability for almost every  $\lambda$ . A decomposition is called trivial if all the  $\lambda$ -probabilities occurring in it are trivial. Also, a hidden variable theory is called trivial if all the  $\lambda$ -probabilities occurring in it are trivial.

From Section 6 onwards, limits are taken involving multiple variables. These are always *repeated* limits. If, for example, we write

$$\lim_{l \rightarrow \infty} \lim_{n \rightarrow \infty} f(n, l) = \lim_{l \rightarrow \infty} \left( \lim_{n \rightarrow \infty} f(n, l) \right), \quad (2.9)$$

this means that first the limit  $n \rightarrow \infty$  is taken, and  $\lim_{n \rightarrow \infty} f(n, l)$  exists. Note that the order of the limits is important: it might be that in the above case  $\lim_{l \rightarrow \infty} f(n, l)$  does not exist in which case  $\lim_{n \rightarrow \infty} \lim_{l \rightarrow \infty} f(n, l)$  is not well-defined. If  $\epsilon > 0$  and  $\lim_{l \rightarrow \infty} \lim_{n \rightarrow \infty} f(n, l) = 0$  then, to find an  $n$  and an  $l$  such that  $|f(n, l)| < \epsilon$ , *first* an  $l$  large enough must be chosen, and *then* an  $n$  large enough.

### 3 Theorem

The main result of this article concerns  $\lambda$ -probabilities for measurements on a system  $A$  in the state  $|\psi\rangle_A$ , to which a hidden variable theory assigns an extra variable  $\lambda$ . This  $\lambda$  is distributed according to some measure  $\mu(\lambda)$ , which might depend not only on the quantum state  $|\psi\rangle_A$  but also on other factors, like the specific method that was used to prepare the state. Now, given a value of  $\lambda$ , the outcome probabilities of a measurement of system  $A$  in the state  $|\psi\rangle_A$  may differ from the quantum-mechanical probabilities. Also, the  $\lambda$ -probabilities might have additional dependencies compared with quantum probabilities. For example,  $\lambda$ -probabilities may be contextual, as mentioned in the previous section. They may also depend on the specific implementation of the measurement. If a  $\lambda$ -probability for a certain implementation of the measurement were non-trivial, this would generate a non-trivial decomposition. Therefore, it is enough to show that there is no non-trivial decomposition, and additional factors such as the specific implementation of the measurement do not have to be included when writing down  $\lambda$ -probabilities:  $\Pr_\lambda^{|\psi\rangle_A} \left( \hat{E}_i^A \right)_{\hat{O}^A}$ .<sup>5</sup>

Now, given a value of  $\lambda$  not only the outcome probabilities of a measurement of system  $A$  in the state  $|\psi\rangle_A$  may differ, but also outcome probabilities of

<sup>5</sup>More specifically, factors such as the specific implementation can be omitted because these can be held fixed throughout the derivation, by picking one implementation for every observable that is considered. However, the measurement context (i.e. the observable to which the projector belongs) does have to be explicitly mentioned initially, because this cannot be held fixed throughout the derivation: different contexts are sometimes considered for the same projector. For example, (5.32) is a result for a specific context, but using perfect correlation, this leads to (5.34), which holds in any context. Thanks go to Guido Bacciagaluppi for pointing out this method.

measurements on systems which have interacted with  $A$ . For example,  $A$  may be coupled to a system  $B$  in the following way:

$$|\psi\rangle_A |0\rangle_B \mapsto \sum_{j=0}^{d-1} \hat{E}_j^A |\psi\rangle_A |j\rangle_B \quad (3.10)$$

where  $\{\hat{E}_j^A\}_{j=0}^{d-1}$  is a complete set of orthogonal projectors and  $\{|j\rangle_B\}_{j=0}^{d-1}$  a set of orthonormal vectors. Then, a measurement in the basis  $\{|j\rangle_B\}_{j=0}^{d-1}$  on  $B$  can be considered, and for this measurement the  $\lambda$ -probabilities might also differ from the quantum-mechanical probabilities. Like above, there might be additional dependencies, such as on the specific implementation of the interaction between  $A$  and  $B$ . Again, a non-trivial  $\lambda$ -probability would imply a non-trivial decomposition, and therefore these additional dependencies can be omitted, so we write down the  $\lambda$ -probability as

$$\Pr_\lambda^{|\phi\rangle_{AB}} ([i]^B)_{\hat{O}_B}. \quad (3.11)$$

So, here  $\lambda$  still refers to the variable assigned to system  $A$  when it was in the state  $|\psi\rangle_A$ . We will also consider measurements on more complicated composite systems like  $AA'A''BB'B''$ , and also in these cases,  $\lambda$  always refers to the original system  $A$ , and there is only one measure  $\mu(\lambda)$  that is considered.

We can now further impose conditions on such decompositions. For example, we might impose *non-contextuality*, meaning that the  $\lambda$ -probability of a measurement outcome only depends on the corresponding projector, and not on the other projectors characterizing the measurement. This is however a strong condition, as it follows from Gleason's theorem that such decompositions are trivial if the dimension of the Hilbert space of the system is at least 3.<sup>6</sup> We will, however, impose the following conditions:

- **CompQuant: Compatibility with Quantum-Mechanical Predictions**

We consider theories where the quantum state  $|\psi\rangle_A$  is supplemented with a hidden variable  $\lambda$ , which assumes a value from the measurable space  $\Lambda$ . It is assumed that the quantum-mechanical probabilities are retrieved when averaging using the measure  $\mu(\lambda)$ . This means that, for any decomposition generated by the hidden variable theory, for every projective measurement involving state  $|\phi\rangle$ , observable  $\hat{O}$  and projector  $\hat{E}$ ,

$$\int_\Lambda d\mu(\lambda) \Pr_\lambda^{|\phi\rangle} (\hat{E})_{\hat{O}} = \Pr^{|\phi\rangle} (\hat{E}) = \langle \phi | \hat{E} | \phi \rangle. \quad (3.12)$$

<sup>6</sup>According to Gleason's theorem (Gleason, 1957), for any Hilbert space with dimension at least 3, any probability measure over the projectors corresponds to a unique density matrix  $\rho$  satisfying  $\Pr(\hat{E}) = \text{Tr}(\rho\hat{E})$  for all projectors  $\hat{E}$ . Now, if  $\Pr(\hat{E})$  could be decomposed so that  $\Pr(\hat{E}) = \int d\mu(\lambda) \Pr_\lambda(\hat{E})$  then, by Gleason's theorem, to each  $\Pr_\lambda(\hat{E})$  there would correspond a density matrix  $\rho_\lambda$  such that  $\Pr_\lambda(\hat{E}) = \text{Tr}(\rho_\lambda\hat{E})$ . But then,  $\Pr(\hat{E}) = \int d\mu(\lambda) \text{Tr}(\rho_\lambda\hat{E}) = \text{Tr}((\int d\mu(\lambda)\rho_\lambda)\hat{E})$ , so that  $\int d\mu(\lambda)\rho_\lambda = \rho$ . If  $\rho$  is a pure state, then it can, by the definition of a pure state, only be trivially decomposed, which means that  $\rho_\lambda \doteq \rho$ , and therefore  $\Pr_\lambda(\hat{E}) \doteq \Pr(\hat{E})$ .

As discussed above,  $\lambda$  is the hidden variable assigned to system  $A$  when it was in the state  $|\psi\rangle_A$ , although the probabilities considered can be for measurements on systems in other states, e.g. systems which have interacted with  $A$ . It is also assumed that the measure  $\mu$  is independent of what measurement is being performed. This is commonly justified by the fact that experimenters are free to choose their measurements, independent of the value of  $\lambda$ .<sup>7</sup>

• **ParInd: Parameter Independence**

When considering joint measurements,  $\lambda$ -probabilities for measurement outcomes on one subsystem are independent of which measurement is performed, and of whether a measurement is performed at all,<sup>8</sup> on any other subsystem.<sup>9</sup> For example, for measurements on subsystems  $A$  and  $B$  of a composite system  $AB$  this means that

$$\begin{aligned} \sum_{j \in J} \Pr_{\lambda}^{|\phi\rangle_{AB}} \left( \hat{E}_i^A, \hat{E}_j^B \right)_{\hat{O}^A \otimes \hat{O}^B} &= \sum_{k \in K} \Pr_{\lambda}^{|\phi\rangle_{AB}} \left( \hat{E}_i^A, \hat{E}_k^B \right)_{\hat{O}^A \otimes \hat{O}^B} \\ &=: \Pr_{\lambda}^{|\phi\rangle_{AB}} \left( \hat{E}_i^A \right)_{\hat{O}^A}, \end{aligned} \quad (3.13)$$

where  $\{\hat{E}_i^A\}_{i \in I}$ ,  $\{\hat{E}_j^B\}_{j \in J}$  and  $\{\hat{E}_k^B\}_{k \in K}$  are complete sets of orthogonal projectors. Note that ParInd makes the last expression well-defined: without ParInd, it could depend on the type of measurement performed on  $B$ . Note that this definition of ParInd is more general than the one standardly used in the literature, where it usually only plays a role in EPR-type measurements.

Note that it is also implicitly assumed that  $\lambda$ -probabilities for measurements on a certain system are independent of the state of other systems, e.g.

$$P_{\lambda}^{|\phi\rangle_{AB} \otimes |\zeta\rangle_{A'B'}} \left( \hat{E}^{AB} \right)_{\hat{O}_{AB}} = P_{\lambda}^{|\phi\rangle_{AB} \otimes |\chi\rangle_{A'B'}} \left( \hat{E}^{AB} \right)_{\hat{O}_{AB}} = P_{\lambda}^{|\phi\rangle_{AB}} \left( \hat{E}^{AB} \right)_{\hat{O}_{AB}}. \quad (3.14)$$

This is used at the end of Section 6. Denying this doesn't seem to make sense:  $\lambda$ -probabilities could then depend on states of arbitrary other systems, and of course there is always an infinitude of other systems of which the states are continually evolving. We now formulate the main theorem:

**Theorem 1.** *Let  $A$  be a system to which quantum mechanics assigns the state  $|\psi\rangle_A$ . Consider a projective measurement on  $A$  with projectors  $\{\hat{E}_i^A\}_{i=0}^{d-1}$ . Then, any decomposition satisfying CompQuant and ParInd is trivial, i.e.*

$$\Pr_{\lambda}^{|\psi\rangle_A} \left( \hat{E}_i^A \right) \doteq \Pr^{|\psi\rangle_A} \left( \hat{E}_i^A \right). \quad (3.15)$$

In Colbeck & Renner (2011), where C&R first present their triviality claim, the supplemental variable, analogous to  $\lambda$  in this article, is by them dubbed

<sup>7</sup>Note that this means that we do not consider retrocausal and superdeterministic models.

<sup>8</sup>Not performing a measurement on a system can equivalently be described as ‘performing a measurement of the observable  $\mathbb{1}$ ’, which always yields the outcome 1.

<sup>9</sup>Note that no requirement is imposed on the spatiotemporal relation between the two measurements constituting the joint measurement: they are not necessarily space-like separated.



‘additional information’ and represented by a discrete random variable  $Z$ , with distribution  $\Pr(Z)$ . Also, they assume that  $Z$  is accessible to experimenters, as an output of an operation that has an input represented by the variable  $C$ . Note that if this the case, an ensemble of systems in identical quantum-mechanical states can be prepared with a distribution that deviates from the distribution  $\Pr(Z)$ , simply by first preparing an ensemble with distribution  $\Pr(Z)$  and then discarding some of the systems, the selection depending on the value of  $Z$ . It is then possible that outcomes of measurements performed on such modified ensembles deviate from the quantum-mechanical predictions. However, there is still compatibility with quantum mechanics in the sense that, when averaging using the original distribution  $\Pr(Z)$ , the quantum-mechanical probabilities are recovered. The fact that C&R consider  $Z$  to be accessible also explains why the condition analogous to ParInd is by them called ‘no-signalling’. If  $\lambda$  is accessible to experimenters, a violation of ParInd would allow someone in control of one subsystem to send information to someone in control of another subsystem. In their later article (Colbeck & Renner, 2012a), C&R state that  $Z$  can alternatively be interpreted as ‘forever hidden and hence unlearnable in principle’. In the present article, nothing is assumed about the accessibility of  $\lambda$ , in order to stay as general as possible.

In the following sections, we shall derive the theorem step-by-step. In Section 4, we repeat the derivation of a known result regarding measurements on Bell states. In Section 5, the result is generalized to higher-dimensional states. In Section 6, states will be considered where the Schmidt coefficients are square roots of rational numbers, while in Section 7 the generalization is made to arbitrary entangled states. In Section 8, the final generalization is made to any projective measurement and to POVM<sup>10</sup> measurements, including measurements on systems in non-entangled states.

The following table gives an overview of the intermediate steps leading to the final result, and the assumptions made at each step.<sup>11</sup>

Section	State	Result
4	$1/\sqrt{2}( 0\rangle_A 0\rangle_B +  1\rangle_A 1\rangle_B)$	$\Pr_\lambda([0]^A) \doteq \Pr([0]^A) = 1/2$
5	$\sum_i c_i i\rangle_A i\rangle_B$	$c_j = c_k \Rightarrow \Pr_\lambda([j]^A) \doteq \Pr_\lambda([k]^A)$
	$\sum_{i=0}^{d-1} 1/\sqrt{d} i\rangle_A i\rangle_B$	$\forall i : \Pr_\lambda([i]^A) \doteq \Pr([i]^A) = 1/d$
6	$\sum_i c_i i\rangle_A i\rangle_B; \forall i : c_i^2 \in \mathbb{Q}$	$\forall i : \Pr_\lambda([i]^A) \doteq \Pr([i]^A) = c_i^2$
7	$\sum_i c_i i\rangle_A i\rangle_B; \forall i : c_i \in \mathbb{R}$	$\forall i : \Pr_\lambda([i]^A) \doteq \Pr([i]^A) = c_i^2$
8	$ \psi\rangle_A$	$\forall \hat{E}^A : \Pr_\lambda(\hat{E}^A) \doteq \Pr(\hat{E}^A) = \ \hat{E}^A \psi\rangle_A\ ^2$
	$ \psi\rangle_A$	$\forall \hat{F}^A : \Pr_\lambda(\hat{F}^A) \doteq \Pr(\hat{F}^A) = {}_A\langle\psi \hat{F}^A \psi\rangle_A$

## 4 Triviality for Bell states

Consider a bipartite system  $AB$  prepared in a Bell state:

$$|\phi\rangle_{AB} := \frac{1}{\sqrt{2}}(|0\rangle_A|0\rangle_B + |1\rangle_A|1\rangle_B) \in \mathcal{H}_A \otimes \mathcal{H}_B = \mathbb{C}^2 \otimes \mathbb{C}^2, \quad (4.16)$$

<sup>10</sup>POVM stands for Positive Operator-Valued Measure. A detailed treatment of these different types of measurements can be found in Busch et al. (1996).

<sup>11</sup> $\hat{E}^A$  is a projector, while  $\hat{F}^A$  is a positive operator.

where  $\{|0\rangle_A, |1\rangle_A\}$  and  $\{|0\rangle_B, |1\rangle_B\}$  are orthonormal bases. Consider a measurement on subsystem  $A$  in the basis  $\{|0\rangle_A, |1\rangle_A\}$ . This measurement has two possible outcomes, each with a quantum-mechanical probability  $1/2$ . The result of this section will be that in any decomposition satisfying ParInd and CompQuant, the  $\lambda$ -probabilities of these outcomes also equal  $1/2$ , rendering such a decomposition trivial with respect to this measurement.

Define a continuous set of observables for  $A$  and  $B$

$$\begin{aligned}\hat{O}_\theta &:= -1 \cdot [\theta] + 1 \cdot [\theta + \pi], \text{ where} \\ |\theta\rangle &:= \cos(\theta/2)|0\rangle + \sin(\theta/2)|1\rangle \text{ and } \theta \in [0, \pi].\end{aligned}\quad (4.17)$$

Note that  $\langle \theta | \theta + \pi \rangle = 0$ , and  $\hat{O}_\theta$  has eigenvectors  $|\theta\rangle, |\theta + \pi\rangle$  with eigenvalues  $-1, +1$ . For any  $N \in \mathbb{N}$ , define the following discrete subset of these observables:

$$\begin{aligned}\hat{A}_{N,a} &:= \hat{O}_{a\pi/2N}^A, \quad a \in \{0, 2, \dots, 2N\} =: \mathcal{A}_N; \\ \hat{B}_{N,b} &:= \hat{O}_{b\pi/2N}^B, \quad b \in \{1, 3, \dots, 2N-1\} =: \mathcal{B}_N; \\ \hat{A}_0 &:= \hat{A}_{N,0}.\end{aligned}\quad (4.18)$$

The last definition is unambiguous because  $\hat{A}_{N,0} = \hat{O}_0^A$  is independent of  $N$ . Note that the observables  $\hat{A}_{N,2N}$  and  $\hat{A}_0$  correspond to the measurement in the basis  $\{|0\rangle_A, |1\rangle_A\}$ . Also note that  $\hat{A}_{N,2N} = -\hat{A}_0$ , which means that  $A_{N,2N} = -A_0$ . Define a correlation measure<sup>12</sup>

$$I_N := \sum_{|a-b|=1} \Pr^{|\phi\rangle_{AB}}(A_{N,a} \neq B_{N,b}). \quad (4.19)$$

Before turning to the formal proof, the following should provide an intuition of the idea behind it. The correlation measure above is inspired by the use of *chained Bell inequalities*, which were introduced in Pearle (1970) and Braunstein & Caves (1989). An important ingredient of chained Bell inequalities is that correlations are considered between outcomes when jointly measuring  $\hat{O}_\theta^A$  and  $\hat{O}_{\theta'}^B$ , with a small difference between  $\theta$  and  $\theta'$ . The correlation when jointly measuring two consecutive observables from the list  $\hat{A}_0, \hat{B}_{N,1}, \dots, \hat{B}_{N,2N-1}, \hat{A}_{N,2N}$  gets stronger as  $N$  gets larger:  $\Pr(A_{N,a} \neq B_{N,b})$  gets closer to 0, while the difference between  $\theta$  and  $\theta'$  becomes smaller. Combining all the correlations between outcomes of consecutive observables, which is what happens in the definition of  $I_N$  in (4.19), an upper bound can be derived for the difference in  $\lambda$ -probabilities of the outcomes for the first and the last observables, i.e.  $\hat{A}_0$  and  $\hat{A}_{N,2N} = -\hat{A}_0$ . It turns out that as  $N$  gets larger, the correlations become so much stronger that the  $\lambda$ -probabilities for the values of  $A_0$  and  $-A_0$  must be equal, i.e. they must both equal  $1/2$ . This is related to the fact that as  $N \rightarrow \infty$ , each term in  $I_N$  is roughly proportional to  $1/N^2$ , while the number of terms is proportional to  $N$ . Therefore,  $I_N$  is roughly proportional to  $1/N$  and goes to 0 in the limit. This is illustrated in Figure 1.

Proceeding with the formal proof, omitting the superscript  $|\phi\rangle_{AB}$  for nota-

<sup>12</sup>In all summations involving  $a$  and  $b$ , it is assumed that  $a \in \mathcal{A}_N$  and  $b \in \mathcal{B}_N$ .

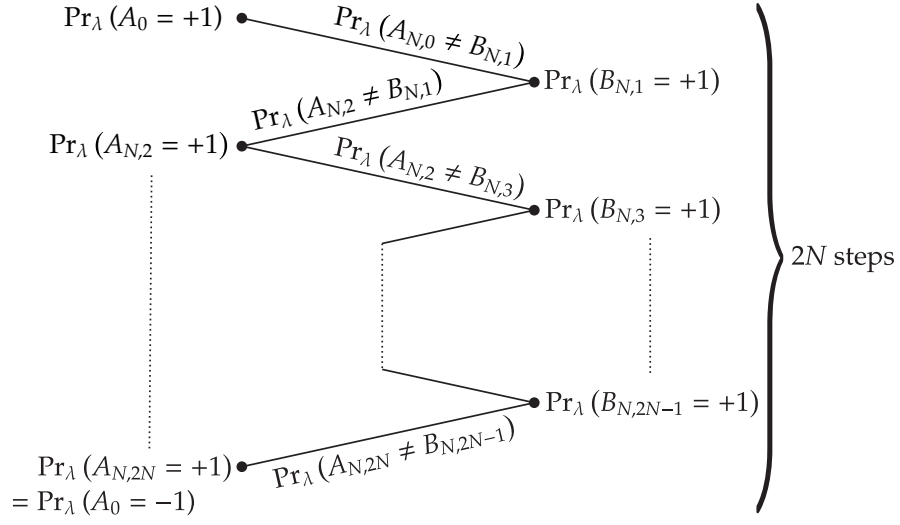


Figure 1: The probabilities displayed at two subsequent points differ no more than the probability displayed at the line between these two points. Therefore, the difference between  $\Pr_\lambda(A_0 = +1)$  and  $\Pr_\lambda(A_0 = -1)$  is bounded above by  $I_N$ , which is the sum of  $2N$  probabilities of the form  $\Pr_\lambda(A_{N,a} \neq B_{N,b})$  with  $|a - b| = 1$ . The  $\lambda$ -average of each of these probabilities is smaller than  $\pi^2/(16N^2)$ . Therefore, the  $\lambda$ -averaged difference between  $\Pr_\lambda(A_0 = +1)$  and  $\Pr_\lambda(A_0 = -1)$  is bounded above by  $2N\pi^2/(16N^2) = \pi^2/(8N)$ . Since this holds for all  $N \in \mathbb{N}$ , the  $\lambda$ -averaged difference equals zero, and therefore  $\Pr_\lambda(A_0 = +1) = \Pr_\lambda(A_0 = -1) = 1/2$  for almost all  $\lambda$ .

tional simplicity,

$$\begin{aligned}
|\Pr_\lambda(A_0 = 1) - \Pr_\lambda(A_0 = -1)| &= |\Pr_\lambda(A_{N,0} = 1) - \Pr_\lambda(A_{N,2N} = 1)| \\
&\leq \sum_{|a-b|=1} |\Pr_\lambda(A_{N,a} = 1) - \Pr_\lambda(B_{N,b} = 1)| \\
&\leq \sum_{|a-b|=1} \Pr_\lambda(A_{N,a} \neq B_{N,b}). \tag{4.20}
\end{aligned}$$

The first inequality is a simple triangle inequality, while the second follows from an inequality in probability theory:

$$\begin{aligned}
\forall z \in \text{Range}(X) \cap \text{Range}(Y): \\
|\Pr(X = z) - \Pr(Y = z)| &= |\Pr(X = z, Y = z) + \Pr(X = z, Y \neq z) \\
&\quad - \Pr(X = z, Y = z) - \Pr(X \neq z, Y = z)| \\
&\leq \Pr(X = z, Y \neq z) + \Pr(X \neq z, Y = z) \\
&\leq \Pr(X \neq Y). \tag{4.21}
\end{aligned}$$

Note that we have assumed ParInd here: this guarantees that, for example,  $\Pr_\lambda(A_{N,a} = 1)$  is well-defined and independent of what measurement is being performed on  $B$ .

Integrating both sides of (4.20) using the measure  $\mu(\lambda)$  gives, using CompQuant,

$$\begin{aligned}
\int_\Lambda d\mu(\lambda) |\Pr_\lambda(A_0 = 1) - \Pr_\lambda(A_0 = -1)| &\leq \int_\Lambda d\mu(\lambda) \sum_{|a-b|=1} \Pr_\lambda(A_{N,a} \neq B_{N,b}) \\
&= \sum_{|a-b|=1} \Pr(A_{N,a} \neq B_{N,b}) = I_N. \tag{4.22}
\end{aligned}$$

$I_N$  is a sum of quantum-mechanical probabilities and can therefore be calculated:

$$\begin{aligned}
a \in \mathcal{A}_N, b \in \mathcal{B}_N, |a-b| = 1 &\Rightarrow \Pr^{|\phi\rangle^{AB}}(A_{N,a} = \pm 1, B_{N,b} = \mp 1) = 1/2 \sin^2\left(\frac{\pi}{4N}\right), \\
\text{so } I_N = 2N \sin^2\left(\frac{\pi}{4N}\right) &\leq \frac{\pi^2}{8N} \Rightarrow \lim_{N \rightarrow \infty} I_N = 0. \tag{4.23}
\end{aligned}$$

Since (4.22) has to hold for all  $N$ , and the left-hand side is non-negative, it follows that

$$\int_\Lambda d\mu(\lambda) |\Pr_\lambda(A_0 = 1) - \Pr_\lambda(A_0 = -1)| = 0 \quad \Rightarrow \quad \Pr_\lambda(A_0 = 1) \doteq \Pr_\lambda(A_0 = -1). \tag{4.24}$$

Since both probabilities add up to 1, they both equal 1/2 almost everywhere. This can be written as<sup>13</sup>

$$\begin{aligned}
\Pr_\lambda([0]^A) &\doteq \Pr([0]^A) = 1/2; \\
\Pr_\lambda([1]^A) &\doteq \Pr([1]^A) = 1/2. \tag{4.25}
\end{aligned}$$

<sup>13</sup>The observable need not be specified here because in a 2-dimensional Hilbert space a single projector fixes the other projector associated with the measurement.

In fact, for any other basis  $\{|0'\rangle_A, |1'\rangle_A\}$  of  $\mathcal{H}_A$ , there is a basis  $\{|0'\rangle_B, |1'\rangle_B\}$  of  $\mathcal{H}_B$  such that  $|\phi\rangle_{AB} = \frac{1}{\sqrt{2}}(|0'\rangle_A|0'\rangle_B + |1'\rangle_A|1'\rangle_B)$ .<sup>14</sup> Therefore, (4.25) also holds for measurements on  $A$  in any other basis. We have proved triviality for the state  $|\phi\rangle_{AB}$  (4.16), in the 4-dimensional Hilbert space  $\mathbb{C}^2 \otimes \mathbb{C}^2$ .

## 5 Generalizing to higher dimensional state spaces

The result of the previous section can easily be generalized to maximally entangled states in higher dimensional state spaces. Consider a bipartite system  $AB$  in the state

$$|\phi_d\rangle_{AB} := \sum_{i=0}^{d-1} c_i |i\rangle_A |i\rangle_B \in \mathcal{H}_A \otimes \mathcal{H}_B = \mathbb{C}^d \otimes \mathbb{C}^d, \quad (5.26)$$

with  $c_i \in \mathbb{R}^+$  for all  $i$ ,  $d > 2$ , and  $\{|i\rangle_A\}_{i=0}^{d-1}$  and  $\{|i\rangle_B\}_{i=0}^{d-1}$  orthonormal bases. Consider a measurement on  $A$  in the basis  $\{|i\rangle_A\}_{i=0}^{d-1}$ . We shall prove that equal coefficients imply equal  $\lambda$ -probabilities:

$$\forall j, k \in \mathbb{N}_d: \quad c_j = c_k \Rightarrow \Pr_\lambda([j]^A) \doteq \Pr_\lambda([k]^A). \quad (5.27)$$

In particular, if all coefficients in (5.26) equal  $1/\sqrt{d}$ , all outcome probabilities equal  $1/d$  almost everywhere:

$$\forall i \in \mathbb{N}_d: \quad \Pr_\lambda([i]^A) \doteq \frac{1}{d} = \Pr([i]^A). \quad (5.28)$$

The proof is essentially the same as in the previous section: we focus on the two terms with equal coefficients and apply the same steps. Let  $c_j$  and  $c_k$  be two equal coefficients. Similarly to the definitions (4.17) and (4.18) in the previous section, we define

$$\begin{aligned} |\theta\rangle &:= \cos(\theta/2)|j\rangle + \sin(\theta/2)|k\rangle, \quad \theta \in [0, \pi]; \\ \hat{O}_\theta &:= -1 \cdot [\theta] + 1 \cdot [\theta + \pi] + \sum_{i \in \mathbb{N}_r \setminus \{j, k\}} (i+2) \cdot [i]; \\ \hat{A}_{N,a} &:= \hat{O}_{a\pi/2N}, \quad a \in \mathcal{A}_N; \\ \hat{B}_{N,b} &:= \hat{O}_{b\pi/2N}, \quad b \in \mathcal{B}_N; \\ \hat{A}_0 &:= \hat{A}_{N,0}; \\ I'_N &:= \sum_{|a-b|=1} \Pr^{|\phi_d\rangle_{AB}}(A_{N,a} \neq B_{N,b}). \end{aligned} \quad (5.29)$$

Note that the observables  $\hat{A}_{N,0}$  and  $\hat{A}_{N,2N}$  correspond to the measurement in the basis  $\{|i\rangle_A\}_{i=0}^{d-1}$ . Also, these observables are equal up to the exchange of the eigenvalues  $+1$  and  $-1$ , which means that  $A_{N,0} = \pm 1$  iff  $A_{N,2N} = \mp 1$ . The value of  $I'_N$  is similar to that of  $I_N$  (4.19) in the previous section. The difference

<sup>14</sup>Namely, the basis  $\{|0'\rangle_B, |1'\rangle_B\}$ , where  $|i'\rangle_B := U_{ij}^* |j\rangle_B$  and  $U_{ij}$  is the unitary matrix such that  $U_{ij} |j\rangle_A = |i'\rangle_A$ .

comes from the fact that the coefficients of the terms involving  $|j\rangle_A$  and  $|k\rangle_A$  equal  $c_j$  instead of  $1/\sqrt{2}$ :

$$I'_N = 4Nc_j^2 \sin^2\left(\frac{\pi}{4N}\right) \leq \frac{\pi^2 c_j^2}{4N}. \quad (5.30)$$

Since  $c_j$  is constant,  $I'_N$ , like  $I_N$ , has the property  $\lim_{N \rightarrow \infty} I'_N = 0$ . Repeating the steps (4.20)–(4.24) gives

$$\Pr_\lambda(A_0 = 1) \doteq \Pr_\lambda(A_0 = -1). \quad (5.31)$$

Using projectors, this can be expressed as

$$\Pr_\lambda([j]^A)_{A_0} \doteq \Pr_\lambda([k]^A)_{A_0}. \quad (5.32)$$

Now, using the perfect correlation property (D.147) proven in Appendix D, setting  $I = \{j\}$ ,

$$\Pr_\lambda([j]^A) \doteq \Pr_\lambda([j]^B), \quad (5.33)$$

we get

$$\Pr_\lambda([j]^A) \doteq \Pr_\lambda([j]^B) \doteq \Pr_\lambda([k]^B) \doteq \Pr_\lambda([k]^A). \quad (5.34)$$

Note that the perfect correlation also implies that these probabilities are independent of the observables. We now focus on the case with all coefficients equal to  $1/\sqrt{d}$ . Then, for any basis  $\{|i'\rangle_A\}_{i=0}^{d-1}$  there is a basis  $\{|i'\rangle_B\}_{i=0}^{d-1}$  such that  $|\phi_d\rangle_{AB} = \sum_{i=0}^{d-1} (1/\sqrt{d}) |i'\rangle_A |i'\rangle_B$ .<sup>15</sup> It follows that for measurements in any basis  $\{|i'\rangle_A\}_{i=0}^{d-1}$  or  $\{|i'\rangle_B\}_{i=0}^{d-1}$ :

$$\begin{aligned} \forall i \in \mathbb{N}_d: \Pr_\lambda([i']^A) &\doteq \frac{1}{d} = \Pr([i']^A), \\ \Pr_\lambda([i']^B) &\doteq \frac{1}{d} = \Pr([i']^B). \end{aligned} \quad (5.35)$$

In this case triviality also holds for measurements involving multidimensional projectors (or, equivalently, observables with degenerate eigenvalues). Consider a measurement on  $A$  involving the projector  $\sum_{i \in J} [i]^A$ , where  $J \subset \mathbb{N}_d$ , and a simultaneous measurement on  $B$  in the basis  $\{|i\rangle_B\}_{i=0}^{d-1}$ . Then,

$$\Pr_\lambda\left(\sum_{i \in J} [i]^A\right) \doteq \sum_{i \in J} \Pr_\lambda([i]^B) \doteq \sum_{i \in J} \frac{1}{d} = \frac{\#J}{d} = \Pr\left(\sum_{i \in J} [i]^A\right). \quad (5.36)$$

Here we have made use of the perfect correlation property

$$\Pr_\lambda^{|\phi_d\rangle_{AB}}\left(\sum_{i \in J} [i]^A\right) \doteq \sum_{i \in J} \Pr_\lambda^{|\phi_d\rangle_{AB}}([i]^B), \quad (5.37)$$

also proven in Appendix D. Remember that by ParInd, the outcome probabilities of a measurement performed on  $A$  are independent of whether any

<sup>15</sup>Namely, the basis  $\{|i'\rangle_B\}_{i=0}^{d-1}$ , where  $|i'\rangle_B := U_{ij}^* |j\rangle_B$  and  $U_{ij}$  is the unitary matrix such that  $U_{ij} |j\rangle_A = |i'\rangle_A$ .

measurement is performed on  $B$ . Therefore, while we did consider a specific measurement on  $B$  in (5.36), the result

$$\Pr_\lambda \left( \sum_{i \in J} [i]^A \right) \doteq \Pr \left( \sum_{i \in J} [i]^A \right) \quad (5.38)$$

holds in general.

## 6 Generalizing to coefficients $c_i^2 \in \mathbb{Q}$

In this section the result of the previous sections is generalized to states of which the Schmidt coefficients equal square roots of rational numbers:<sup>16</sup>

$$|\phi_d\rangle_{AB} := \sum_{i=0}^{d-1} c_i |i\rangle_A |i\rangle_B \in \mathcal{H}_A \otimes \mathcal{H}_B, \quad (6.39)$$

where  $c_i^2 \in \mathbb{Q}$  and  $c_i > 0$  for all  $i$ ,  $d > 2$ ,  $\mathcal{H}_A$  and  $\mathcal{H}_B$  are both at least of dimension  $d$ , and  $\{|i\rangle_A\}_{i=0}^{d-1}$  and  $\{|i\rangle_B\}_{i=0}^{d-1}$  are sets of orthonormal vectors.<sup>17</sup> We shall prove that also in this case the  $\lambda$ -probabilities equal the quantum probabilities:

$$\forall i \in \mathbb{N}_d: \Pr_\lambda^{|\phi_d\rangle_{AB}}([i]^A) \doteq \Pr^{|\phi_d\rangle_{AB}}([i]^A). \quad (6.40)$$

The proof in this section is based on the following idea. System  $AB$  can be coupled to another system  $A'B'$  such that the combined system  $AA'BB'$  is approximately in a maximally entangled state. More precisely, for each  $i$  the term  $|i\rangle_A |i\rangle_B$  is coupled to an approximate entangled state with a number of terms that is proportional to  $c_i^2$ . To find these numbers, we consider the common denominator  $r$  of all fractions  $\{c_i^2\}_{i=0}^{d-1}$ , such that we can write  $c_i^2 = m_i/r$  with  $m_i \in \mathbb{N}$ . The numerators  $m_i$  are then proportional to  $c_i^2$ . Then, for measurements on  $AA'BB'$ , which is approximately in a maximally entangled state, the result of the previous section can be applied. It turns out that this also puts constraints on the  $\lambda$ -probabilities for measurements on system  $AB$  alone.

To produce the maximally entangled state, we assume the presence of yet another bipartite system  $A''B''$  that is prepared in a special ‘embezzling state’. The family of embezzling states was first introduced in van Dam & Hayden (2003), and they have the special property that any bipartite state can be approximately extracted from it using local unitary operations (i.e. by applying separate unitary transformations on  $A''$  and  $B''$ ). The precision of this operation can be enhanced by using an embezzling state of higher dimension.

Because we have to use the notion of ‘approximate states’ in this section, we define a distance measure for quantum states, called the trace distance:

$$D(|\psi\rangle, |\phi\rangle) := (1/2) \text{Tr} \left| |\psi\rangle\langle\psi| - |\phi\rangle\langle\phi| \right|, \text{ where} \quad (6.41)$$

$$|A| := \sqrt{A^\dagger A}.$$

<sup>16</sup>The set of rational numbers is denoted by  $\mathbb{Q}$ .

<sup>17</sup>Since we allow for state spaces of dimension larger than  $d$ ,  $\{|i\rangle_A\}_{i=0}^{d-1}$  and  $\{|i\rangle_B\}_{i=0}^{d-1}$  do not necessarily span  $\mathcal{H}_A$  and  $\mathcal{H}_B$  and therefore might not be bases.

As shown in Nielsen & Chuang (2004), the trace distance is a metric, and it provides an upper bound for the difference in quantum probabilities for the outcome associated with any projector  $\hat{P}$ :

$$\left| \langle \Psi | \hat{P} | \Psi \rangle - \langle \phi | \hat{P} | \phi \rangle \right| \leq D(|\psi\rangle, |\phi\rangle). \quad (6.42)$$

Also, for pure states the following relation holds between the trace distance and another distance measure, the fidelity  $\mathcal{F}$ :

$$D(|\psi\rangle, |\phi\rangle) = \sqrt{1 - \mathcal{F}(|\psi\rangle, |\phi\rangle)^2},$$

where  $\mathcal{F}(|\psi\rangle, |\phi\rangle) := |\langle \psi | \phi \rangle|$ . (6.43)

Beginning with the proof, first define<sup>18</sup>

$$r := \text{LCD} \left( \{c_i^2\}_{i=0}^{d-1} \right);$$

$$m_i := r c_i^2. \quad (6.44)$$

The family of embezzling states  $\{|\tau_n\rangle\}_{n=1}^\infty$  is defined as

$$|\tau_n\rangle := \frac{1}{\sqrt{C_n}} \sum_{j=0}^{n-1} \frac{1}{\sqrt{j+1}} |j\rangle \otimes |j\rangle \in \mathbb{C}^n \otimes \mathbb{C}^n. \quad (6.45)$$

$C_n := \sum_{j=0}^{n-1} 1/(j+1)$  is a normalization constant, and  $\{|j\rangle\}_{j=0}^{n-1}$  an orthonormal basis of  $\mathbb{C}^n$ . System  $A''B''$  is prepared in one of these states, indexed by  $n$ . Another system  $A'B'$  is prepared in a simple product state

$$|0\rangle_{A'} |0\rangle_{B'} \in \mathcal{H}_{A'} \otimes \mathcal{H}_{B'}, \quad (6.46)$$

where  $\mathcal{H}_{A'}$  and  $\mathcal{H}_{B'}$  are at least of dimension  $\max_{i \in \mathbb{N}_d} m_i$ . Taking all systems together, the state before the measurement is

$$|\Psi_n\rangle := |\tau_n\rangle_{A''B''} \otimes |0\rangle_{A'} |0\rangle_{B'} \otimes |\phi_d\rangle_{AB}. \quad (6.47)$$

Now, for each  $i$ , a maximally entangled state with Schmidt number  $m_i$  is extracted from the embezzling state, and coupled to the term  $|i\rangle_A |i\rangle_B$ . In Appendix C it is shown that there are unitary operators  $U_n^{AA'A''}$ ,  $U_n^{BB'B''}$  which perform this task with a precision that increases as  $n \rightarrow \infty$ :

$$\lim_{n \rightarrow \infty} D \left( U_n^{AA'A''} \otimes U_n^{BB'B''} |\Psi_n\rangle, |\chi_n\rangle \right) = 0, \text{ where} \quad (6.48)$$

$$|\chi_n\rangle := |\tau_n\rangle_{A''B''} \otimes \sum_{i=0}^{d-1} \left( \sqrt{\frac{m_i}{r}} |i\rangle_A |i\rangle_B \sum_{j=0}^{m_i-1} \frac{1}{\sqrt{m_i}} |j\rangle_{A'} |j\rangle_{B'} \right)$$

$$= |\tau_n\rangle_{A''B''} \otimes \sum_{i=0}^{d-1} \sum_{j=0}^{m_i-1} \left( \frac{1}{\sqrt{r}} |i\rangle_A |i\rangle_B |j\rangle_{A'} |j\rangle_{B'} \right).$$

Here,  $\{|j\rangle_{A'}\}_{j=0}^{m-1}$  and  $\{|j\rangle_{B'}\}_{j=0}^{m-1}$  are sets of orthonormal vectors. Now consider a measurement that consists of first performing the above unitary transformation, resulting approximately in a maximally entangled state for the system  $AA'BB'$ , and then measuring this system in the basis  $\{|i, j\rangle_{AA'}\}$ , where

<sup>18</sup>“LCD” stands for “least common denominator”



$|i, j\rangle := |i\rangle|j\rangle$ . Note that, if the state  $U_n^{AA'A''} \otimes U_n^{BB'B''} |\Psi_n\rangle$  is close to a maximally entangled state, by (6.42) the outcome probabilities are also close to those of a maximally entangled state. Therefore, the result of the previous section can be applied. Actually, the unitary transformation can be included in the definition of the measurement projectors, because a unitary transformation applied to a complete set of orthogonal projectors gives another complete set of orthogonal projectors.

Define an index set containing all values of the pair  $(i, j)$  occurring in  $|\chi_n\rangle$ :

$$I^{\text{ind}} := \left\{ (0, 0), (0, 1), \dots, (0, m_0 - 1), \right. \\ (1, 0), \dots, (1, m_1 - 1), \\ \dots, \\ \left. (d-1, 0), \dots, (d-1, m_{d-1} - 1) \right\}. \quad (6.49)$$

For  $(i_1, j_1), (i_2, j_2) \in I^{\text{ind}}$  define, similarly to (5.29) in the previous section,<sup>19</sup>

$$\begin{aligned} |\theta\rangle &:= \cos(\theta/2)|i_1, j_1\rangle + \sin(\theta/2)|i_2, j_2\rangle; \\ \hat{O}_\theta &:= -1 \cdot [\theta] + 1 \cdot [\theta + \pi] + \sum_{(i,j) \in I^{\text{ind}} \setminus \{(i_1, j_1), (i_2, j_2)\}} (2^i 3^j + 2) \cdot [i, j]; \\ \hat{A}_{N,n,a} &:= \left( U_n^{AA'A''} \right)^{-1} \left( \mathbb{I}^{A''} \otimes \hat{O}_{a\pi/2N} \right) \left( U_n^{AA'A''} \right), \quad a \in \mathcal{A}_N; \\ \hat{B}_{N,n,b} &:= \left( U_n^{BB'B''} \right)^{-1} \left( \mathbb{I}^{B''} \otimes \hat{O}_{b\pi/2N} \right) \left( U_n^{BB'B''} \right), \quad b \in \mathcal{B}_N; \\ \hat{A}_{n,0} &:= \hat{A}_{N,n,0}. \end{aligned} \quad (6.50)$$

Note that, as in the previous section,  $A_{n,0} = \pm 1$  iff  $A_{N,n,2N} = \mp 1$ . The use of the expression  $(2^i 3^j + 2)$  in the definition of  $\hat{O}_\theta$  guarantees that distinct eigenvalues are assigned to every projector  $[i, j]$ , as in the previous section. Note that the observables  $\hat{A}$  and  $\hat{B}$  include the unitary operators  $U_n^{AA'A''}$  and  $U_n^{BB'B''}$ , and therefore depend on  $n$ , which corresponds to the precision of the embezzlement transformation.

Finally, a correlation measure is again defined, which now also depends on the precision  $n$ :

$$I_{N,n} := \sum_{|a-b|=1} \Pr^{|\Psi_n\rangle} (A_{N,n,a} \neq B_{N,n,b}). \quad (6.51)$$

This quantity has the property, similar to  $I_N$  in the previous section, that it vanishes when taking the limits  $n \rightarrow \infty$  and then  $N \rightarrow \infty$ :

$$\forall (i_1, j_1), (i_2, j_2) \in I^{\text{ind}}: \lim_{N \rightarrow \infty} \lim_{n \rightarrow \infty} I_{N,n} = 0. \quad (6.52)$$

The proof of (6.52) can be found in Appendix A.1. The correctness is already suggested by the fact that  $I_{N,n}$  consists of  $2N$  probabilities, and as mentioned above, those probabilities get closer to probabilities for a maximally entangled state as  $n$  increases.

<sup>19</sup>While the symbols defined in (6.50) and (6.51) depend on the pairs  $(i_1, j_1)$  and  $(i_2, j_2)$ , these are not included as indices for notational simplicity.

Analogously to the previous two sections, we have

$$\begin{aligned}
\left| \Pr_\lambda^{|\Psi_n\rangle} (A_{n,0} = 1) - \Pr_\lambda^{|\Psi_n\rangle} (A_{n,0} = -1) \right| &= \left| \Pr_\lambda^{|\Psi_n\rangle} (A_{N,n,0} = 1) - \Pr_\lambda^{|\Psi_n\rangle} (A_{N,n,2N} = 1) \right| \\
&\leq \sum_{|a-b|=1} \left| \Pr_\lambda^{|\Psi_n\rangle} (A_{N,n,a} = 1) - \Pr_\lambda^{|\Psi_n\rangle} (B_{N,n,b} = 1) \right| \\
&\leq \sum_{|a-b|=1} \Pr_\lambda^{|\Psi_n\rangle} (A_{N,n,a} \neq B_{N,n,b}).
\end{aligned} \tag{6.53}$$

Integrating with the measure  $\mu(\lambda)$  gives

$$\int_\Lambda d\mu(\lambda) \left| \Pr_\lambda^{|\Psi_n\rangle} (A_{n,0} = 1) - \Pr_\lambda^{|\Psi_n\rangle} (A_{n,0} = -1) \right| \leq I_{N,n}. \tag{6.54}$$

Let  $\epsilon > 0$ . By (6.52), we can choose  $n, N$  such that for all  $(i_1, j_1), (i_2, j_2) \in I^{\text{ind}}$ ,  $I_{N,n} < \epsilon$ , so that

$$\forall (i_1, j_1), (i_2, j_2) \in I^{\text{ind}}: \int_\Lambda d\mu(\lambda) \left| \Pr_\lambda^{|\Psi_n\rangle} (A_{n,0} = 1) - \Pr_\lambda^{|\Psi_n\rangle} (A_{n,0} = -1) \right| < \epsilon. \tag{6.55}$$

Defining the projector

$$\hat{E}_{(i,j),n} := \left( U_n^{AA'A''} \right)^{-1} \left( \mathbb{I}^{A''} \otimes [i, j]^{AA'} \right) \left( U_n^{AA'A''} \right), \tag{6.56}$$

and noting that

$$\hat{A}_{n,0} = -1 \cdot \hat{E}_{(i_1, j_1), n} + 1 \cdot \hat{E}_{(i_2, j_2), n} + \dots, \tag{6.57}$$

we get, switching to the notation with projectors,

$$\forall (i_1, j_1), (i_2, j_2) \in I^{\text{ind}}: \int_\Lambda d\mu(\lambda) \left| \Pr_\lambda^{|\Psi_n\rangle} \left( \hat{E}_{(i_2, j_2), n} \right)_{\hat{A}_{n,0}} - \Pr_\lambda^{|\Psi_n\rangle} \left( \hat{E}_{(i_1, j_1), n} \right)_{\hat{A}_{n,0}} \right| < \epsilon. \tag{6.58}$$

By CompQuant, we have

$$\sum_{(i', j') \in I^{\text{ind}}} \int_\Lambda d\mu(\lambda) \Pr_\lambda^{|\Psi_n\rangle} \left( \hat{E}_{(i', j'), n} \right) = 1 \quad \Rightarrow \quad \sum_{(i', j') \in I^{\text{ind}}} \Pr_\lambda^{|\Psi_n\rangle} \left( \hat{E}_{(i', j'), n} \right) \doteq 1. \tag{6.59}$$

Using this, and a triangle inequality, we get

$$\begin{aligned}
\forall (i, j) \in I^{\text{ind}}: &\int_\Lambda d\mu(\lambda) \left| \Pr_\lambda^{|\Psi_n\rangle} \left( \hat{E}_{(i,j), n} \right)_{\hat{A}_{n,0}} - \frac{1}{r} \right| \\
&= \frac{1}{r} \int_\Lambda d\mu(\lambda) \left| r \cdot \Pr_\lambda^{|\Psi_n\rangle} \left( \hat{E}_{(i,j), n} \right)_{\hat{A}_{n,0}} - \sum_{(i', j') \in I^{\text{ind}}} \Pr_\lambda^{|\Psi_n\rangle} \left( \hat{E}_{(i', j'), n} \right)_{\hat{A}_{n,0}} \right| \\
&\leq \frac{1}{r} \sum_{(i', j') \in I^{\text{ind}}} \int_\Lambda d\mu(\lambda) \left| \Pr_\lambda^{|\Psi_n\rangle} \left( \hat{E}_{(i,j), n} \right)_{\hat{A}_{n,0}} - \Pr_\lambda^{|\Psi_n\rangle} \left( \hat{E}_{(i', j'), n} \right)_{\hat{A}_{n,0}} \right| < \frac{\#I^{\text{ind}}}{r} \epsilon = \epsilon.
\end{aligned} \tag{6.60}$$

Using a triangle inequality once more,

$$\begin{aligned} \forall i \in \mathbb{N}_d: \int_{\Lambda} d\mu(\lambda) \left| \left( \sum_{j=0}^{m_i-1} \Pr_{\lambda}^{|\Psi_n\rangle} \left( \hat{E}_{(i,j),n} \right) \right)_{\hat{A}_{n,0}} - \frac{m_i}{r} \right| \\ \leq \sum_{j=0}^{m_i-1} \int_{\Lambda} d\mu(\lambda) \left| \Pr_{\lambda}^{|\Psi_n\rangle} \left( \hat{E}_{(i,j),n} \right)_{\hat{A}_{n,0}} - \frac{1}{r} \right| < m_i \epsilon. \end{aligned} \quad (6.61)$$

Using the correlation property

$$\sum_{j=0}^{m_i-1} \Pr_{\lambda}^{|\Psi_n\rangle} \left( \hat{E}_{(i,j),n} \right) \doteq \Pr_{\lambda}^{|\Psi_n\rangle} ([i]^B), \quad (6.62)$$

proven in Appendix D, we arrive at

$$\begin{aligned} \forall i \in \mathbb{N}_d: \int_{\Lambda} d\mu(\lambda) \left| \left( \sum_{j=0}^{m_i-1} \Pr_{\lambda}^{|\Psi_n\rangle} \left( \hat{E}_{(i,j),n} \right) \right)_{\hat{A}_{n,0}} - \frac{m_i}{r} \right| \\ = \int_{\Lambda} d\mu(\lambda) \left| \Pr_{\lambda}^{|\Psi_n\rangle} ([i]^B) - \frac{m_i}{r} \right| < m_i \epsilon. \end{aligned} \quad (6.63)$$

Since (6.63) holds for any  $\epsilon > 0$ , we get, also applying (3.14) by noting that  $|\Psi_n\rangle = |\phi_d\rangle_{AB} \otimes \dots$ ,

$$\begin{aligned} \int_{\Lambda} d\mu(\lambda) \left| \Pr_{\lambda}^{|\phi_d\rangle_{AB}} ([i]^B) - \frac{m_i}{r} \right| = 0 \\ \Rightarrow \Pr_{\lambda}^{|\phi_d\rangle_{AB}} ([i]^B) \doteq \frac{m_i}{r} = c_i^2 = \Pr^{|\phi_d\rangle_{AB}} ([i]^B), \end{aligned} \quad (6.64)$$

which is the desired result.

Note that because of ParInd, once we have derived results for the probabilities for measurements on system  $B$  using the correlation property and certain measurements performed on system  $AA'A''$ , these results still hold when other measurements, or no measurements at all, are being performed on system  $AA'A''$ .

Also note that, starting with the system  $AA'A''BB'B''$ , with  $A''B''$  prepared in an embezzling state, we have reached a conclusion about measurements on  $AB$ , prepared in the pure state (6.39). Above we have applied (3.14), which states that the probabilities  $\Pr_{\lambda}^{|\phi_d\rangle_{AB}} ([i]^B)$  are independent of the states of the systems  $A''B''$  and  $A'B'$ . Therefore, while these systems were used for the derivation of (6.64), it follows that the result also holds without considering the systems  $A''B''$  and  $A'B'$  in specially prepared states.

Like in the previous section, the result also holds with  $A$  and  $B$  interchanged, and also for degenerate measurements.

## 7 Generalizing to arbitrary coefficients

Now suppose we have a state of the form

$$|\phi_d\rangle_{AB} := \sum_{i=0}^{d-1} c_i |i\rangle_A |i\rangle_B \in \mathcal{H}_A \otimes \mathcal{H}_B, \quad (7.65)$$

where  $c_i \in \mathbb{R}^+$  for all  $i$ ,  $d > 2$ ,  $\mathcal{H}_A$  and  $\mathcal{H}_B$  are both at least of dimension  $d$ , and  $\{|i\rangle_A\}_{i=0}^{d-1}$  and  $\{|i\rangle_B\}_{i=0}^{d-1}$  are sets of orthonormal vectors. Note that any state can be written in this form by Schmidt's decomposition theorem. Since we cannot generally write  $c_i^2 = m_i/r$  with  $m_i, r \in \mathbb{N}$ , the strategy of the previous section can not be applied. However, because the set of rational numbers is dense in the set of real numbers, the numbers  $c_i^2$  can be approximated with rational numbers  $c_{i,l}^2$  by defining the sequences

$$(c_{i,l})_{l=1}^\infty, \text{ with } \forall i, l \in \mathbb{N}: c_{i,l}^2 \in \mathbb{Q}, \quad (7.66)$$

$$\forall i: \lim_{l \rightarrow \infty} c_{i,l} = c_i, \quad (7.67)$$

$$\forall l: \sum_{i=0}^{d-1} c_{i,l}^2 = 1. \quad (7.68)$$

As in the previous section, the idea is to apply a unitary transformation that results approximately in a maximally entangled state. But since the  $c_i^2$  are not rational numbers, we have to use the rational numbers  $c_{i,l}^2$  to determine the numerators and denominators for the unitary transformations. This time, to achieve any desired precision, not only  $n$  must be chosen large enough, but also  $l$ . Again, the unitary transformation will be included in the measurement operators.

Instead of focussing on two terms, like we did in the previous section, the terms of the approximately maximally entangled state are now partitioned into two sets of equal size. Then, a measurement is considered using a projector that projects on the space spanned by one such set of terms, and measurements where this projector is 'rotated' by an angle  $\theta$ . For a maximally entangled state, the quantum-mechanical statistics for such measurements are equal to those of the measurements on Bell states, considered in Section 4. Since we have a state that is close to a maximally entangled state, by (6.42) the quantum probabilities are also close to those of Bell states, i.e. close to  $1/2$ . Using a similar derivation as in the previous two sections, it can be shown that also the  $\lambda$ -probabilities are close to  $1/2$ . Then, using a lemma proved in Appendix B, we can also derive  $\lambda$ -probabilities for projectors  $[i]^A$ . In fact, those probabilities are again equal to the quantum-mechanical ones for almost every  $\lambda \in \Lambda$ .

For each  $l \in \mathbb{N}$ , define  $r_l$  as the least *even* common denominator of the fractions  $\{c_{i,l}^2\}_{i=0}^{d-1}$ ,<sup>20</sup> and let  $\{m_{i,l}\}_{i=0}^{d-1}$  be the corresponding numerators:

$$\begin{aligned} r_l &:= \text{LCD} \left( \{c_{i,l}^2\}_{i=0}^{d-1} \cup \{1/2\} \right); \\ m_{i,l} &:= r_l c_{i,l}^2. \end{aligned} \quad (7.69)$$

The reason that  $r_l$  needs to be even is that an approximately maximally entangled state with an even number of terms is needed in order to be able to divide the terms into two groups of equal size.

As in the previous section, we assume the presence of two extra bipartite systems, of which one is prepared in an embezzling state (6.45):

$$|\Psi_n\rangle := |\tau_n\rangle_{A''B''} \otimes |0\rangle_{A'} |0\rangle_{B'} \otimes |\phi_d\rangle_{AB}. \quad (7.70)$$

<sup>20</sup>This is achieved by including  $1/2$  in the set of fractions.

Note that, for given  $l$ ,  $\mathcal{H}_{A'}$  and  $\mathcal{H}_{B'}$  are both at least of dimension  $\max_{i \in \mathbb{N}_d} m_{i,l}$ .  
Again, define an index set:

$$\begin{aligned} J_l^{\text{ind}} := & \left\{ (0, 0), (0, 1), \dots, (0, m_{0,l} - 1), \right. \\ & (1, 0), \dots, (1, m_{1,l} - 1), \\ & \dots, \\ & \left. (d-1, 0), \dots, (d-1, m_{d-1,l} - 1) \right\}. \end{aligned} \quad (7.71)$$

Note that  $\sum_{i=0}^{d-1} m_{i,l} = \sum_{i=0}^{d-1} r_i c_{i,l}^2 = r_l$  by (7.69) and (7.68), and therefore  $\#J_l^{\text{ind}} = r_l$ .

Let  $J_l \subset J_l^{\text{ind}}$  with  $\#J_l = r_l/2$ ; and  $p_l : J_l \rightarrow J_l^{\text{ind}} \setminus J_l$  an arbitrary bijection. Then, define

$$\begin{aligned} \hat{O}_{J_l, \theta}^{AA'} := & +1 \cdot \sum_{(i,j) \in J_l} [\cos(\theta/2)|i, j\rangle_{AA'} + \sin(\theta/2)|p_l(i, j)\rangle_{AA'}] \\ & - 1 \cdot \left( \mathbb{I}^{AA'} - \sum_{(i,j) \in J_l} [\cos(\theta/2)|i, j\rangle_{AA'} + \sin(\theta/2)|p_l(i, j)\rangle_{AA'}] \right); \\ \hat{A}_{J_l, N, n, l, a} := & \left( U_{n,l}^{AA'A''} \right)^{-1} \left( \mathbb{I}^{A''} \otimes \hat{O}_{J_l, a\pi/2N}^{AA} \right) \left( U_{n,l}^{AA'A''} \right), \quad a \in \mathcal{A}_N \setminus \{2N\}, \end{aligned} \quad (7.72)$$

and  $\hat{O}_{J_l, \theta}^{BB'}$ ,  $\hat{B}_{J_l, N, n, l, b}$  analogously. Furthermore, define

$$\begin{aligned} \hat{A}_{J_l, n, l, 0} & := \hat{A}_{J_l, N, n, l, 0}; \\ \hat{A}_{J_l, N, n, l, 2N} & := -\hat{A}_{J_l, n, l, 0}. \end{aligned} \quad (7.73)$$

Now the correlation measure can be defined:

$$I_{J_l, N, n, l} := \sum_{|a-b|=1} \Pr^{|\Psi_n\rangle} (A_{J_l, N, n, l, a} \neq B_{J_l, N, n, l, b}). \quad (7.74)$$

As shown in Appendix A.2, this measure has the property

$$\lim_{N \rightarrow \infty} \lim_{l \rightarrow \infty} \lim_{n \rightarrow \infty} I_{J_l, N, n, l} = 0. \quad (7.75)$$

Now,

$$\begin{aligned} 2 \left| \Pr_{\lambda}^{|\Psi_n\rangle} (A_{J_l, n, l, 0} = +1) - 1/2 \right| & = \left| \Pr_{\lambda}^{|\Psi_n\rangle} (A_{J_l, n, l, 0} = +1) - \Pr_{\lambda}^{|\Psi_n\rangle} (A_{J_l, n, l, 0} = -1) \right| \\ & = \left| \Pr_{\lambda}^{|\Psi_n\rangle} (A_{J_l, N, n, l, 0} = +1) - \Pr_{\lambda}^{|\Psi_n\rangle} (A_{J_l, N, n, l, 2N} = +1) \right| \\ & \leq \sum_{|a-b|=1} \left| \Pr_{\lambda}^{|\Psi_n\rangle} (A_{J_l, N, n, l, a} = +1) - \Pr_{\lambda}^{|\Psi_n\rangle} (B_{J_l, N, n, l, b} = +1) \right| \\ & \leq \sum_{|a-b|=1} \Pr_{\lambda}^{|\Psi_n\rangle} (A_{J_l, N, n, l, a} \neq B_{J_l, N, n, l, b}) \end{aligned} \quad (7.76)$$

Integrating with the measure  $\mu(\lambda)$  gives

$$\int_{\Lambda} d\mu(\lambda) \left| \Pr_{\lambda}^{|\Psi_n\rangle} (A_{J_l, n, l, 0} = +1) - 1/2 \right| \leq (1/2) I_{J_l, N, n, l}. \quad (7.77)$$

Let  $\epsilon > 0$ . By (7.75) and (7.67) we can choose  $N, l, n \in \mathbb{N}$  such that for all  $J_l \subset J_l^{\text{ind}}$  with  $\#J_l = r_l/2$ ,

$$\int_{\Lambda} d\mu(\lambda) \left| \Pr_{\lambda}^{|\Psi_n\rangle} (A_{J_l, n, l, 0} = +1) - 1/2 \right| < \epsilon; \quad \forall i \in \mathbb{N}_d: |c_{i,l}^2 - c_i^2| < \epsilon. \quad (7.78)$$

Now, define the projectors

$$\begin{aligned} \hat{E}_{(i,j), n, l} &:= \left( U_{n,l}^{AA'A''} \right)^{-1} \left( [i, j]^{AA'} \otimes \mathbb{I}^{A''} \right) \left( U_{n,l}^{AA'A''} \right), \\ \hat{F}_{(i,j), n, l} &:= \left( U_{n,l}^{BB'B''} \right)^{-1} \left( [i, j]^{BB'} \otimes \mathbb{I}^{B''} \right) \left( U_{n,l}^{BB'B''} \right). \end{aligned} \quad (7.79)$$

Noting that

$$\hat{A}_{J_l, n, l, 0} = +1 \cdot \left( \sum_{(i,j) \in J_l} \hat{E}_{(i,j), n, l} \right) - 1 \cdot \left( \mathbb{I}^{AA'A''} - \sum_{(i,j) \in J_l} \hat{E}_{(i,j), n, l} \right), \quad (7.80)$$

we switch to the notation with projectors, and using the correlation property (D.148),

$$\Pr_{\lambda}^{|\Psi_n\rangle} \left( \sum_{(i,j) \in J_l} \hat{E}_{(i,j), n, l} \right) \doteq \sum_{(i,j) \in J_l} \Pr_{\lambda}^{|\Psi_n\rangle} \left( \hat{F}_{(i,j), n, l} \right), \quad (7.81)$$

we arrive at, for all  $J_l \subset J_l^{\text{ind}}$  with  $\#J_l = r_l/2$ ,

$$\begin{aligned} \int_{\Lambda} d\mu(\lambda) \left| \Pr_{\lambda}^{|\Psi_n\rangle} \left( \sum_{(i,j) \in J_l} \hat{E}_{(i,j), n, l} \right) - 1/2 \right| &< \epsilon \\ \Rightarrow \int_{\Lambda} d\mu(\lambda) \left| \sum_{(i,j) \in J_l} \Pr_{\lambda}^{|\Psi_n\rangle} \left( \hat{F}_{(i,j), n, l} \right) - 1/2 \right| &< \epsilon. \end{aligned} \quad (7.82)$$

Now, we need the following lemma, proven in Appendix B:

**Lemma 1.** *Let  $\Lambda$  be a measurable space with measure  $\mu$ , let  $r$  be an even, positive integer, and let*

$$\left\{ (p_i^\lambda)_{i=0}^{r-1} \mid \lambda \in \Lambda \right\} \quad (7.83)$$

*be some collection of sequences satisfying, for all  $I \subset I^{\text{ind}} := \mathbb{N}_r$  with  $\#I = r/2$ ,*

$$\int_{\Lambda} d\mu(\lambda) \left| \left( \sum_{i \in I} p_i^\lambda \right) - 1/2 \right| < \epsilon. \quad (7.84)$$

*Then for all  $J \subset I^{\text{ind}}$*

$$\int_{\Lambda} d\mu(\lambda) \left| \left( \sum_{i \in J} p_i^\lambda \right) - \#J/r \right| < 2\epsilon. \quad (7.85)$$

Applying this lemma to the right-hand side of (7.82) with

$$(p_i^\lambda)_{i=0}^{r-1} \equiv \left( \Pr_\lambda^{|\Psi_n\rangle} \left( \hat{F}_{(0,0),n,l} \right), \Pr_\lambda^{|\Psi_n\rangle} \left( \hat{F}_{(0,1),n,l} \right), \dots, \Pr_\lambda^{|\Psi_n\rangle} \left( \hat{F}_{(d-1,m_{d-1,l}),n,l} \right) \right), \quad (7.86)$$

we get, for all  $i \in \mathbb{N}_d$ ,

$$\int_\Lambda d\mu(\lambda) \left| \sum_{j=0}^{m_{i,l}-1} \Pr_\lambda^{|\Psi_n\rangle} \left( \hat{F}_{(i,j),n,l} \right) - \frac{m_{i,l}}{r_l} \right| < 2\epsilon. \quad (7.87)$$

Again, applying perfect correlation (D.150),

$$\sum_{j=0}^{m_{i,l}-1} \Pr_\lambda^{|\Psi_n\rangle} \left( \hat{F}_{(i,j),n,l} \right) \doteq \Pr_\lambda^{|\Psi_n\rangle} ([i]^A), \quad (7.88)$$

and noting that  $m_{i,l}/r_l = c_{i,l}^2$  by (7.69), we obtain, also using (3.14),

$$\int_\Lambda d\mu(\lambda) \left| \Pr_\lambda^{|\phi_d\rangle^{AB}} ([i]^A) - c_{i,l}^2 \right| < 2\epsilon. \quad (7.89)$$

Finally, using the triangle inequality  $|x - y| \leq |x - z| + |z - y|$

$$\begin{aligned} & \int_\Lambda d\mu(\lambda) \left| \Pr_\lambda^{|\phi_d\rangle^{AB}} ([i]^A) - c_i^2 \right| \\ & \leq \int_\Lambda d\mu(\lambda) \left| \Pr_\lambda^{|\phi_d\rangle^{AB}} ([i]^A) - c_{i,l}^2 \right| + \int_\Lambda d\mu(\lambda) |c_{i,l}^2 - c_i^2| \\ & < 3\epsilon, \end{aligned} \quad (7.90)$$

and since the inequality holds for all  $\epsilon > 0$ , we get

$$\begin{aligned} & \int_\Lambda d\mu(\lambda) \left| \Pr_\lambda^{|\phi_d\rangle^{AB}} ([i]^A) - c_i^2 \right| = 0 \\ \Rightarrow & \Pr_\lambda^{|\phi_d\rangle^{AB}} ([i]^A) \doteq c_i^2 = \Pr_\lambda^{|\phi_d\rangle^{AB}} ([i]^A), \end{aligned} \quad (7.91)$$

which is the desired result. Again, the result also holds with  $A$  and  $B$  interchanged, and also for degenerate measurements.

## 8 Generalizing to any measurement

In the above sections, we considered measurements on subsystems of bipartite systems in entangled states. However, since in fact every measurement involves such entanglement, something can also be said about general measurements. In this section, we show how this can be done.<sup>21</sup> This generalization is performed step-by-step, by considering three types of measurements in turn: measurements of the first kind, measurements of the second kind, and POVM<sup>22</sup> measurements.

<sup>21</sup>The results of this section are largely based on a suggestion made by Guido Bacciagaluppi (private communication).

<sup>22</sup>POVM stands for Positive Operator-Valued Measure. A detailed treatment of these different types of measurements can be found in Busch et al. (1996).

### 8.1 Measurements of the first kind

The measurement process can be divided into two steps. First, the system to be measured is coupled to the measurement apparatus, which results in entanglement between the system and the apparatus. Then, the measurement result is read off the measurement apparatus. In the case of a measurement of the first kind with projectors  $\{\hat{E}_j^A\}_{j=0}^{d-1}$ , the coupling has the following form:

$$|\psi\rangle_A |0\rangle_B = \sum_{j=0}^{d-1} \hat{E}_j^A |\psi\rangle_A |0\rangle_B \mapsto \sum_{j=0}^{d-1} \hat{E}_j^A |\psi\rangle_A |j\rangle_B = \sum_{j=0}^{d-1} c_j |j\rangle_A |j\rangle_B, \quad (8.92)$$

where

$$|j\rangle_A := \frac{\hat{E}_j^A |\psi\rangle_A}{\|\hat{E}_j^A |\psi\rangle_A\|}, \quad \text{and} \quad c_j := |\langle j|\psi\rangle_A| = \|\hat{E}_j^A |\psi\rangle_A\|, \quad (8.93)$$

$|i\rangle_B$  are (mutually orthogonal) ‘pointer states’, and  $|0\rangle_B$  is the so-called ‘ready state’ of the measurement apparatus before measuring. The reading out of the measurement apparatus can then be described as a measurement performed on the apparatus in the ‘pointer basis’:

$$\Pr_{\sum_{j=0}^{d-1} c_j |j\rangle_A |j\rangle_B} ([i]^B). \quad (8.94)$$

Note that these probabilities equal the probabilities of the original measurement on system  $A$ , as they should, because both measurements are actually two different descriptions of a single measurement:

$$\Pr^{|\psi\rangle_A} (\hat{E}_i^A) = \Pr_{\sum_{j=0}^{d-1} c_j |j\rangle_A |j\rangle_B} ([i]^B). \quad (8.95)$$

Since these probabilities are equal by definition, the same relation holds when considering  $\lambda$ -probabilities:

$$\Pr_{\lambda}^{|\psi\rangle_A} (\hat{E}_i^A)_{\hat{O}^A} = \Pr_{\lambda}^{\sum_{j=0}^{d-1} c_j |j\rangle_A |j\rangle_B} ([i]^B)_{\hat{O}^B}, \quad (8.96)$$

where  $\hat{O}^B$  is an observable that includes the pointer states as eigenstates. Now, to the probability on the right-hand side of the above equation the result of Section 7 can be applied. Therefore, for any measurement of the first kind, the  $\lambda$ -probabilities are trivial:<sup>23</sup>

$$\Pr_{\lambda}^{|\psi\rangle_A} (\hat{E}_i^A) \doteq \Pr^{|\psi\rangle_A} (\hat{E}_i^A) = c_i^2. \quad (8.97)$$

<sup>23</sup>More generally, the measurement on the apparatus may involve one or more multidimensional projectors on  $\mathcal{H}_B$  that each have multiple pointer states as eigenstates. If this is the case, this corresponds to a degenerate measurement on system  $A$  where, for each such projector on  $\mathcal{H}_B$ , the projectors  $\{\hat{E}_j^A\}$  corresponding to these pointer states are summed to obtain the projector on  $\mathcal{H}_A$  corresponding to the projector on  $\mathcal{H}_B$  (like  $\hat{E}_i^A$  corresponds to  $[i]^B$  in (8.95) and (8.96)). The result of Section 7, also holding for degenerate measurements, still applies. The same considerations apply to the measurements considered in the next two subsections, where in Section 8.3 not projectors, but positive operators  $\{\hat{F}_j^A\}$  are summed.



## 8.2 Measurements of the second kind

In measurements of the second kind, the coupling between system and measurement apparatus is more general. When a measurement of the first kind is performed, a system initially in an eigenstate of the observable remains in that same eigenstate after the measurement. In contrast, when a measurement of the second kind is performed, such a system may end up in a different state. Therefore, the coupling is as follows:

$$|\psi\rangle_A |0\rangle_B = \sum_{j=0}^{d-1} \hat{E}_j^A |\psi\rangle_A |0\rangle_B \mapsto \sum_{j=0}^{d-1} c_j |j\rangle_A |j\rangle_B \quad (8.98)$$

where  $c_j$  is as in (8.95), but now the states  $|j\rangle_A$  can be anything and are in general not mutually orthogonal. This prevents the direct application of the result of Section 7 to the right-hand side of (8.98). However, the measurement apparatus can generally be decomposed into subsystems which get entangled with each other during the measurement (for example, different particles of a pointer). Or, the measurement apparatus may interact with another system before any interaction with the experimenter takes place (for example, a computer recording the outcome). In this case, we can write  $\mathcal{H}_B \equiv \mathcal{H}_{B_1} \otimes \mathcal{H}_{B_2}$ , and we have the coupling

$$|\psi\rangle_A |0\rangle_{B_1} |0\rangle_{B_2} = \sum_{j=0}^{d-1} \hat{E}_j^A |\psi\rangle_A |0\rangle_{B_1} |0\rangle_{B_2} \mapsto \sum_{j=0}^{d-1} c_j |j\rangle_A |j\rangle_{B_1} |0\rangle_{B_2} \mapsto \sum_{j=0}^{d-1} c_j |j\rangle_A |j\rangle_{B_1} |j\rangle_{B_2}. \quad (8.99)$$

Here, the states  $|j\rangle_{B_1}$  are mutually orthogonal, as are the states  $|j\rangle_{B_2}$ . Writing  $|j\rangle_{AB_1} \equiv |j\rangle_A |j\rangle_{B_1}$ , the  $\lambda$ -probabilities of the measurement are now given by

$$\Pr_{\lambda}^{\sum_{j=0}^{d-1} c_j |j\rangle_{AB_1} |j\rangle_{B_2}} ([j]_{B_2})_{\hat{O}_{B_2}} \quad (8.100)$$

and, noting that the states  $|j\rangle_{AB_1}$  are mutually orthogonal, we see that again the result of Section 7 can be applied to it, resulting again in (8.97).

## 8.3 POVM measurements

The result can also be extended to POVM measurements. In the case of such a measurement, the coupling is even more general. It is characterized by a complete set of positive operators  $\{\hat{F}_j^A\}_{j=0}^{d-1}$ , which do not have to be projection operators. In this case the coupling to the measurement apparatus is as follows:

$$|\psi\rangle_A |0\rangle_B \mapsto \sum_{j=0}^{d-1} \hat{M}_j^A |\psi\rangle_A |j\rangle_B = \sum_{j=0}^{d-1} c_j |j\rangle_A |j\rangle_B, \quad (8.101)$$

where

$$|j\rangle_A := \frac{\hat{M}_j^A |\psi\rangle_A}{\left\| \hat{M}_j^A |\psi\rangle_A \right\|}, \quad \text{and} \quad c_j := |\langle j | \psi \rangle_A| = \left\| \hat{M}_j^A |\psi\rangle_A \right\|, \quad (8.102)$$

and the  $\hat{M}_j^A$  are operators such that  $(\hat{M}_j^A)^\dagger \hat{M}_j^A = \hat{F}_j^A$ .

As in the case of measurements of the second kind, the measurement apparatus can generally be decomposed into two subsystems which get entangled during the measurement process:

$$|\psi\rangle_A |0\rangle_{B_1} |0\rangle_{B_2} = \sum_{j=0}^{d-1} \hat{M}_j^A |\psi\rangle_A |0\rangle_{B_1} |0\rangle_{B_2} \mapsto \sum_{j=0}^{d-1} c_j |j\rangle_A |j\rangle_{B_1} |0\rangle_{B_2} \mapsto \sum_{j=0}^{d-1} c_i |j\rangle_A |j\rangle_{B_1} |j\rangle_{B_2}. \quad (8.103)$$

Again, by writing  $|j\rangle_{AB_1} \equiv |j\rangle_A |j\rangle_{B_1}$  we end up with  $\lambda$ -probabilities

$$\Pr_{\lambda}^{\sum_{j=0}^{d-1} c_j |j\rangle_{AB_1} |j\rangle_{B_2}} ([i]^{B_2})_{\hat{O}^{B_2}} \quad (8.104)$$

to which the result of Section 7 can be applied, so that we get

$$\Pr_{\lambda}^{|\psi\rangle_A} (\hat{F}_i^A) \doteq \Pr^{|\psi\rangle_A} (\hat{F}_i^A) = c_i^2. \quad (8.105)$$

## 9 Discussion

We summarize the main differences between the deduction presented in this article and C&R's derivation:

- As mentioned in the Introduction, C&R define their 'Freedom of Choice' assumption in such a way that they can deduce 'no-signalling', which is similar to ParInd, from it (Section VII.A).<sup>24</sup> Instead, we explicitly assume ParInd, whereby their entire issue of deducing ParInd from a freedom assumption becomes irrelevant.
- The proof in Section 4 is considerably simplified compared with that of C&R (Section VII.B).
- The result of Section 5 is not present in C&R's work. Instead, C&R presume that the result of Section 4 can be extended to maximally entangled states of Schmidt number  $2^n$ , by considering  $n$  copies of a system in a Bell state (Section VII.C). Note that results similar to that of Section 5 have been derived in Leifer (2014), Ghirardi & Romano (2012) and Barrett et al. (2006).
- Before generalizing to states with arbitrary Schmidt coefficients in Section 7, we consider states with Schmidt coefficients that are square roots of rational numbers in Section 6. C&R did not perform this intermediate step, and only considered transforming the state approximately to a maximally entangled state with Schmidt number  $2^n$ . Then, they applied the result for Bell states (Section 4), without justifying how this result can be applied to approximate states (Section VII.C). Actually, filling this gap is problematic, and therefore in this article a whole different approach, not present in C&R's work, is introduced to prove the result, including defining the correlation measures for approximate states  $I_{N,n}$  and  $I_{J_1, N, n, l}$  and proving Lemma 1 in Appendix B.

<sup>24</sup>Sections denoted using Roman numerals are sections in Colbeck & Renner (2012b). Sections denoted using Arabic numerals are sections in this article.

- In contrast to C&R (Section IV.B), we do not use random variables to represent measurement settings. In our opinion, a free choice is not something that is best described using random variables; any pattern can be chosen for the measurement settings, and a random variable with a well-defined probability distribution does not seem to be the right mathematical object to represent such a free choice (see also Butterfield, 1992).

Although Theorem 1 is presented as an impossibility theorem for hidden variable theories satisfying ParInd, it also applies to some theories that violate it. A hidden variable theory violating ParInd may satisfy ParInd after partly integrating over the hidden variables. If such a theory is not trivial at the level of the remaining hidden variables, i.e. measurement outcomes depend on them, then the theory is also shown to be incompatible with QM by the theorem. An example is the class of crypto-nonlocal theories considered by Leggett (2003). Leggett introduces hidden variables  $\mathbf{u}$  and  $\mathbf{v}$ , representing definite photon polarizations, and an additional hidden variable  $\lambda$ . When considering the level where all three hidden variables are included, ParInd is violated, so our theorem does not seem to apply. However, after averaging over  $\lambda$ , ParInd is satisfied, while the values of the remaining hidden variables  $\mathbf{u}$  and  $\mathbf{v}$  still give information about measurement outcomes: if a photon is measured in the direction  $\mathbf{u}$ , the outcome is always  $+1$ . Therefore, Theorem 1 can be applied at this level and thereby Leggett's class of theories is shown to be incompatible with QM. Actually, the result of Section 4 is sufficient to rule out Leggett's models, as shown in Branciard et al. (2008).

One might worry about the fact that from Section 6 onwards, we have made use of extra systems in special embezzling states, and complex measurements performed on those systems. The objection can be raised that not every unitary operation can be implemented in practice, and that not every self-adjoint operator corresponds to a real measurement. There does not seem to be a reason why the states and measurements used are impossible in principle.<sup>25</sup> Therefore, it seems contrived to block the derivation of the result by rejecting the use of certain states and measurements. However, when applying the argument of, for example, Section 4 to a bipartite system consisting of a measurement apparatus  $B$  and another system  $A$ , as is essentially done in Section 8, the derivation does include measurements with observables having eigenvectors like  $1/\sqrt{2}(|1\rangle_B + |2\rangle_B)$ , where  $|1\rangle_B$  and  $|2\rangle_B$  are distinct (macroscopic) pointer states of the measurement apparatus. Such measurements are special in the sense that they can transform an apparatus from the state  $|1\rangle_B$  into the state  $|2\rangle_B$ . These are also the kind of measurements that can detect whether the measurement apparatus has collapsed, and such measurements are practically very hard, if not impossible, to perform. They are similar to what Barrett (1999) calls *A-measurements*, which can empirically distinguish collapse from no-collapse versions of quantum mechanics. Such measurements have to date not been performed, and one might be skeptical about the use of such measurements in the derivation. This worry might be reduced by considering  $B$  not to be the whole measurement apparatus, but only a small part of it, while still getting entangled with system  $A$  during the measurement process.

<sup>25</sup>Note that C&R themselves are working in the field of quantum computing and information, where it is usually assumed that any unitary gate can in principle be implemented and any measurement can in principle be performed.

As mentioned in the Introduction, C&R have used their claim to argue that a system's wave function is in a one-to-one correspondence with its ontic state. Roughly, the argument is as follows. C&R distinguish ontic states on the basis of outcome probabilities: if two states predict the same outcome probabilities for all possible measurements, then the states are considered equal. Now, if two systems in equal quantum states were in different ontic states, a variable could be introduced which represents the ontic state of the system. Furthermore, outcome probabilities would depend on this variable, contradicting C&R's claim. Therefore, equal quantum states imply equal ontic states, and the outcome probabilities for a quantum state equal those for the corresponding ontic state.

Now suppose that two systems are in equal ontic states, but in different quantum states. Following the above, outcome probabilities for both quantum states would equal those for the ontic state. But this contradicts the fact that in quantum mechanics, different quantum states imply different outcome probabilities for some measurements. Therefore, equal quantum states imply equal ontic states. It follows that there is a one-to-one correspondence between quantum states and ontic states.

Of course, this result is based on C&R's strong claim that no non-trivial hidden variable theory is compatible with quantum mechanics. If the theorem as presented in this article is used, a weaker result follows: there is a one-to-one correspondence between quantum states and ontic states, *if, at the level of the ontic states, ParInd is satisfied.*

## 10 Acknowledgements

I would like to thank F.A. Muller, Klaas Landsman, Dennis Dieks, Roger Colbeck and Guido Bacciagaluppi for valuable discussions and corrections. This work is part of the research programme 'The Structure of Reality and the Reality of Structure', which is (partly) financed by the Netherlands Organisation for Scientific Research (NWO).

## A Calculation of correlation measures

In this appendix it is shown that, similar to  $\lim_{N \rightarrow \infty} I_N = 0$ , we also have  $\lim_{N \rightarrow \infty} \lim_{n \rightarrow \infty} I_{N,n} = 0$  and  $\lim_{N \rightarrow \infty} \lim_{l \rightarrow \infty} \lim_{n \rightarrow \infty} I_{J_l, N, n, l} = 0$ .

**A.1**  $I_{N,n}$ 

Recall the definitions from Section 6:

$$\begin{aligned}
|\chi_n\rangle &:= |\tau_n\rangle_{A''B''} \otimes \sum_{i=0}^{d-1} \left( \sqrt{\frac{m_i}{r}} |i\rangle_A |i\rangle_B \sum_{j=0}^{m_i-1} \frac{1}{\sqrt{m_i}} |j\rangle_{A'} |j\rangle_{B'} \right), \\
I_{N,n} &:= \sum_{|a-b|=1} \Pr^{|\Psi_n\rangle} (A_{N,n,a} \neq B_{N,n,b}), \text{ where} \\
|\Psi_n\rangle &:= |\tau_n\rangle_{A''B''} \otimes |0\rangle_{A'} |0\rangle_{B'} \otimes |\phi_d\rangle_{AB}, \\
\hat{A}_{N,n,a} &:= \left( U_n^{AA'A''} \right)^{-1} \left( \mathbb{I}^{A''} \otimes \hat{O}_{a\pi/2N}^{AA'} \right) \left( U_n^{AA'A''} \right), \quad a \in \mathcal{A}_N, \\
\hat{B}_{N,n,b} &:= \left( U_n^{BB'B''} \right)^{-1} \left( \mathbb{I}^{B''} \otimes \hat{O}_{b\pi/2N}^{BB'} \right) \left( U_n^{BB'B''} \right), \quad b \in \mathcal{B}_N. \quad (\text{A.106})
\end{aligned}$$

Instead of including the unitary operators  $U_n^{AA'A''}$ ,  $U_n^{BB'B''}$  in the definition of  $\hat{A}_{N,n,a}$  and  $\hat{B}_{N,n,b}$ , they can also be attached to the state  $|\Psi_n\rangle$ . In this case, the quantum probabilities remain the same. Define

$$\begin{aligned}
\hat{\bar{A}}_{N,a} &:= \mathbb{I}^{A''} \otimes \hat{O}_{a\pi/2N}^{AA'}, \quad a \in \mathcal{A}_N; \\
\hat{\bar{B}}_{N,b} &:= \mathbb{I}^{B''} \otimes \hat{O}_{b\pi/2N}^{BB'}, \quad b \in \mathcal{B}_N. \quad (\text{A.107})
\end{aligned}$$

Note that these are just  $\hat{A}_{N,n,a}$  and  $\hat{B}_{N,n,b}$  (6.50) without the unitary operators. Then,

$$I_{N,n} = \sum_{|a-b|=1} \Pr^{(U_n^{AA'A''} \otimes U_n^{BB'B''})|\Psi_n\rangle} (\bar{A}_{N,a} \neq \bar{B}_{N,b}). \quad (\text{A.108})$$

For the state  $|\chi_n\rangle$ , which is close to  $(U_n^{AA'A''} \otimes U_n^{BB'B''})|\Psi_n\rangle$  for large  $n$ , a calculation of quantum probabilities yields for any  $a, b$  with  $|a-b|=1$ :

$$\Pr^{|\chi_n\rangle} (\bar{A}_{N,a} \neq \bar{B}_{N,b}) = \frac{2}{d} \sin^2 \left( \frac{\pi}{4N} \right). \quad (\text{A.109})$$

From the property of the trace distance (6.42), we have

$$\begin{aligned}
& \left| \Pr^{(U_n^{AA'A''} \otimes U_n^{BB'B''})|\Psi_n\rangle} (\bar{A}_{N,a} \neq \bar{B}_{N,b}) - \Pr^{|\chi_n\rangle} (\bar{A}_{N,a} \neq \bar{B}_{N,b}) \right| \\
& \leq D \left( (U_n^{AA'A''} \otimes U_n^{BB'B''})|\Psi_n\rangle, |\chi_n\rangle \right). \quad (\text{A.110})
\end{aligned}$$

Taking the limit  $n \rightarrow \infty$  on the right-hand side yields 0 by (6.48), and therefore, also using (A.109), for  $|a-b|=1$ ,

$$\begin{aligned}
& \lim_{n \rightarrow \infty} \left| \Pr^{(U_n^{AA'A''} \otimes U_n^{BB'B''})|\Psi_n\rangle} (\bar{A}_{N,a} \neq \bar{B}_{N,b}) - \frac{2}{d} \sin^2 \left( \frac{\pi}{4N} \right) \right| = 0 \\
\Rightarrow & \lim_{n \rightarrow \infty} \Pr^{(U_n^{AA'A''} \otimes U_n^{BB'B''})|\Psi_n\rangle} (\bar{A}_{N,a} \neq \bar{B}_{N,b}) = \frac{2}{d} \sin^2 \left( \frac{\pi}{4N} \right). \quad (\text{A.111})
\end{aligned}$$

Finally, because  $I_{N,n}$  is the sum of  $2N$  such probabilities,

$$\lim_{n \rightarrow \infty} I_{N,n} = 2N \frac{2}{d} \sin^2 \left( \frac{\pi}{4N} \right) \leq \frac{\pi^2}{4Nd} \Rightarrow \lim_{N \rightarrow \infty} \lim_{n \rightarrow \infty} I_{N,n} = 0. \quad (\text{A.112})$$

## A.2 $I_{J_l, N, n, l}$

Recall the definitions from Section 7:

$$\begin{aligned}
 I_{J_l, N, n, l} &:= \sum_{|a-b|=1} \Pr^{|\Psi_n\rangle} (A_{J_l, N, n, l, a} \neq B_{J_l, N, n, l, b}), \text{ where} \\
 |\Psi_n\rangle &:= |\tau_n\rangle_{A''B''} \otimes |0\rangle_{A'} |0\rangle_{B'} \otimes \sum_{i=0}^{d-1} c_i |i\rangle_A |i\rangle_B, \\
 \hat{A}_{J_l, N, n, l, a} &:= \left( U_{n, l}^{AA'A''} \right)^{-1} \left( \mathbb{I}^{A''} \otimes \hat{O}_{J_l, a\pi/2N}^{AA} \right) \left( U_{n, l}^{AA'A''} \right), \quad a \in \mathcal{A}_N \setminus \{2N\}, \\
 \hat{B}_{J_l, N, n, l, b} &:= \left( U_{n, l}^{BB'B''} \right)^{-1} \left( \mathbb{I}^{B''} \otimes \hat{O}_{J_l, b\pi/2N}^{BB} \right) \left( U_{n, l}^{BB'B''} \right), \quad b \in \mathcal{B}_N \setminus \{2N\}.
 \end{aligned} \tag{A.113}$$

As in Appendix A.1, we move the unitary operators from the observables to the state. Define

$$\begin{aligned}
 |\chi_{n, l}\rangle &:= |\tau_n\rangle_{A''B''} \otimes \sum_{i=0}^{d-1} \sum_{j=0}^{m_{i, l}-1} \frac{c_i}{\sqrt{m_{i, l}}} |i, j\rangle_{AA'} |i, j\rangle_{BB'}; \\
 |\Phi_{n, l}\rangle &:= |\tau_n\rangle_{A''B''} \otimes \sum_{i=0}^{d-1} \sum_{j=0}^{m_{i, l}-1} \frac{1}{\sqrt{r_l}} |i, j\rangle_{AA'} |i, j\rangle_{BB'}; \\
 \hat{A}_{J_l, N, a} &:= \mathbb{I}^{A''} \otimes \hat{O}_{J_l, a\pi/2N}^{AA'}, \quad a \in \mathcal{A}_N \setminus \{2N\}; \\
 \hat{B}_{J_l, N, b} &:= \mathbb{I}^{B''} \otimes \hat{O}_{J_l, b\pi/2N}^{BB'}, \quad b \in \mathcal{B}_N; \\
 \hat{A}_{J_l, N, 2N} &:= -\hat{A}_{J_l, N, 0}.
 \end{aligned} \tag{A.114}$$

Then,

$$I_{J_l, N, n, l} = \sum_{|a-b|=1} \Pr \left( U_{n, l}^{AA'A''} \otimes U_{n, l}^{BB'B''} \right)^{|\Psi_n\rangle} (\bar{A}_{J_l, N, a} \neq \bar{B}_{J_l, N, b}). \tag{A.115}$$

Quantum mechanics tells us that for the state  $|\Phi_{n, l}\rangle$ , for any  $a, b$  with  $|a-b|=1$ ,

$$\Pr^{|\Phi_{n, l}\rangle} (\bar{A}_{J_l, N, a} \neq \bar{B}_{J_l, N, b}) = \sin^2 \left( \frac{\pi}{4N} \right). \tag{A.116}$$

Also, in the limit  $l \rightarrow \infty$ , the states  $|\chi_{n, l}\rangle$  and  $|\Phi_{n, l}\rangle$  approach each other:

$$\begin{aligned}
 1 \geq \mathcal{F}(|\chi_{n, l}\rangle, |\Phi_{n, l}\rangle) &= |\langle \chi_{n, l} | \Phi_{n, l} \rangle| = \sum_{i=0}^{d-1} \sum_{j=0}^{m_{i, l}-1} m_{i, l} \frac{c_i}{\sqrt{r_l \cdot m_{i, l}}} = \sum_{i=0}^{d-1} \sqrt{\frac{m_{i, l}}{r_l}} \cdot c_i \\
 &= \sum_{i=0}^{d-1} c_{i, l} \cdot c_i = 1 + \sum_{i=0}^{d-1} c_i (c_{i, l} - c_i) \geq 1 - \sum_{i=0}^{d-1} c_i |c_{i, l} - c_i| \\
 &\Rightarrow \lim_{l \rightarrow \infty} \mathcal{F}(|\chi_{n, l}\rangle, |\Phi_{n, l}\rangle) = 1 \text{ and } \lim_{l \rightarrow \infty} D(|\chi_{n, l}\rangle, |\Phi_{n, l}\rangle) = 0,
 \end{aligned} \tag{A.117}$$

where we used (7.67), as well as the relation between fidelity and trace distance (6.43). Using the triangle inequality  $\|x - y\| - \|y - z\| \leq \|x - z\|$  and (6.42), we

have, omitting  $(\bar{A}_{J_l, N, a} \neq \bar{B}_{J_l, N, b})$  after every  $\Pr$  to avoid cluttered notation,

$$\begin{aligned} & \left| \left| \Pr(U_{n,l}^{AA'A''} \otimes U_{n,l}^{BB'B''}) |\Psi_n\rangle - \Pr^{|\Phi_{n,l}\rangle} \right| - \left| \Pr^{|\Phi_{n,l}\rangle} - \Pr^{|\chi_{n,l}\rangle} \right| \right| \\ & \leq \left| \Pr(U_{n,l}^{AA'A''} \otimes U_{n,l}^{BB'B''}) |\Psi_n\rangle - \Pr^{|\chi_{n,l}\rangle} \right| \leq D \left( (U_{n,l}^{AA'A''} \otimes U_{n,l}^{BB'B''}) |\Psi_n\rangle, |\chi_{n,l}\rangle \right). \end{aligned} \quad (\text{A.118})$$

In the limit  $n \rightarrow \infty$ , the right-hand side vanishes by (C.145), and therefore

$$\lim_{n \rightarrow \infty} \left| \left| \Pr(U_{n,l}^{AA'A''} \otimes U_{n,l}^{BB'B''}) |\Psi_n\rangle - \Pr^{|\Phi_{n,l}\rangle} \right| - \left| \Pr^{|\Phi_{n,l}\rangle} - \Pr^{|\chi_{n,l}\rangle} \right| \right| = 0. \quad (\text{A.119})$$

Since  $\left| \Pr^{|\Phi_{n,l}\rangle} - \Pr^{|\chi_{n,l}\rangle} \right|$  is independent of  $n$  we have<sup>26</sup>

$$\lim_{n \rightarrow \infty} \left| \Pr(U_{n,l}^{AA'A''} \otimes U_{n,l}^{BB'B''}) |\Psi_n\rangle - \Pr^{|\Phi_{n,l}\rangle} \right| = \left| \Pr^{|\Phi_{n,l}\rangle} - \Pr^{|\chi_{n,l}\rangle} \right| \leq D(|\chi_{n,l}\rangle, |\Phi_{n,l}\rangle), \quad (\text{A.120})$$

where we also used (6.43) again. By (A.117), the right-hand side vanishes in the limit  $l \rightarrow \infty$ , and therefore we have, also applying (A.116), for any  $a, b$  with  $|a - b| = 1$ ,

$$\lim_{l \rightarrow \infty} \lim_{n \rightarrow \infty} \left| \Pr(U_{n,l}^{AA'A''} \otimes U_{n,l}^{BB'B''}) |\Psi_n\rangle (\bar{A}_{J_l, N, a} \neq \bar{B}_{J_l, N, b}) - \sin^2 \left( \frac{\pi}{4N} \right) \right| = 0. \quad (\text{A.121})$$

The result now follows by following steps similar to (A.110)-(A.112).

$$\begin{aligned} & \lim_{l \rightarrow \infty} \lim_{n \rightarrow \infty} \Pr(U_{n,l}^{AA'A''} \otimes U_{n,l}^{BB'B''}) |\Psi_n\rangle (\bar{A}_{J_l, N, a} \neq \bar{B}_{J_l, N, b}) = \sin^2 \left( \frac{\pi}{4N} \right) \\ & \Rightarrow \lim_{l \rightarrow \infty} \lim_{n \rightarrow \infty} I_{J_l, N, n, l} = 2N \sin^2 \left( \frac{\pi}{4N} \right) \\ & \Rightarrow \lim_{N \rightarrow \infty} \lim_{l \rightarrow \infty} \lim_{n \rightarrow \infty} I_{J_l, N, n, l} = 0. \end{aligned} \quad (\text{A.122})$$

## B Lemma 1

This lemma is based on the proposition that if we have  $r$  numbers, where  $r$  is even, and the sum of any half of these numbers is close to  $1/2$ , then the sum of  $j$  of these numbers is close to  $j/r$ . One might guess that this can be proven by first showing that any individual number is close to  $1/r$ , and then concluding that the sum of  $r/2$  of such numbers are close to  $1/2$ , but it turns out that in our case this does not work. Instead, we make use of the fact that any sum of numbers can be expressed as a linear combination of sums of  $r/2$  numbers. For example, suppose we have the sequence  $(p_i)_{i=0}^9$ , where the sum of any five

<sup>26</sup>The states  $|\Phi_{n,l}\rangle$  and  $|\chi_{n,l}\rangle$  only depend on  $n$  through the embezzling states  $|\tau_n\rangle$  of the subsystems  $A''$  and  $B''$ , but the measurements are performed on the subsystems  $AA'$  and  $BB'$ .

members of the sequence is within  $\epsilon$  from  $1/2$ . We can express any sum of  $p_i$ 's as a linear combination of such sums, for example:

$$\begin{aligned} \sum_{i \in \{0,1\}} p_i = \frac{1}{5} & \left( \sum_{i \in \{0,1,2,3,4\}} p_i + \sum_{i \in \{0,1,5,6,7\}} p_i + \sum_{i \in \{0,1,2,8,9,2\}} p_i + \sum_{i \in \{0,1,3,4,5\}} p_i + \sum_{i \in \{1,2,7,8,9\}} p_i \right. \\ & \left. - \sum_{i \in \{2,3,4,5,6\}} p_i - \sum_{i \in \{7,8,9,2,3\}} p_i - \sum_{i \in \{4,5,6,7,8\}} p_i \right). \end{aligned} \quad (\text{B.123})$$

Since we add or subtract eight sums that are within  $\epsilon$  close of  $1/2$ , and we divide by 5, we find that  $\sum_{i \in \{1,2\}} p_i$  is within  $8\epsilon/5$  of  $1/5 = 2/10$ . In the proof below, we generalize this result, adding an extra index  $\lambda$  because we consider not a single sequence but a collection of sequences over which we average using the measure  $\mu$ .

**Lemma 1.** *Let  $\Lambda$  be a measurable space with measure  $\mu$ , let  $r$  be an even, positive integer, and let*

$$\left\{ (p_i^\lambda)_{i=0}^{r-1} \mid \lambda \in \Lambda \right\} \quad (\text{B.124})$$

be some collection of sequences satisfying, for all  $I \subset I^{\text{ind}} := \mathbb{N}_r$  with  $\#I = r/2$ ,

$$\int_{\Lambda} d\mu(\lambda) \left| \left( \sum_{i \in I} p_i^\lambda \right) - 1/2 \right| < \epsilon. \quad (\text{B.125})$$

Then for all  $J \subset I^{\text{ind}}$

$$\int_{\Lambda} d\mu(\lambda) \left| \left( \sum_{i \in J} p_i^\lambda \right) - \#J/r \right| < 2\epsilon. \quad (\text{B.126})$$

First the case  $\#J \leq r/2$  is considered. Let  $J \subset I^{\text{ind}}$ ,  $\#J \leq r/2$ , and  $\epsilon > 0$ . By (B.125) we have, for all  $I \subset I^{\text{ind}}$  with  $\#I = r/2$ ,

$$\int_{\Lambda} d\mu(\lambda) \left| \left( \sum_{i \in I} p_i^\lambda \right) - 1/2 \right| < \epsilon. \quad (\text{B.127})$$

Define the sequence  $(f_i)_{i=0}^{r-\#J-1}$  containing the elements of  $I^{\text{ind}} \setminus J$  in arbitrary order. Also define

$$\begin{aligned} K_a &:= J \cup \left\{ \overline{f_{a\bar{x}}}, \overline{f_{ax+1}}, \dots, \overline{f_{ax+(x-1)}} \right\} \\ &\text{where } x := r/2 - \#J, \quad \bar{y} := y \bmod (r - \#J), \quad a \in \mathbb{N}_{r/2}; \\ L_b &= \left\{ \overline{f_{b(r/2)}}, \overline{f_{b(r/2)+1}}, \dots, \overline{f_{b(r/2)+(r/2-1)}} \right\}, \quad b \in \mathbb{N}_{r/2-\#J}. \end{aligned} \quad (\text{B.128})$$

Note that for all  $a$  and  $b$ ,  $\#K_a = \#L_b = r/2$ . The  $K_a$  are index sets for the sums of  $p$ 's to be *added*, like the index sets  $\{0, 1, 2, 3, 4\}, \{0, 1, 5, 6, 7\}, \dots$  in (B.123). The  $L_b$  are index sets for the sums of  $p$ 's to be *subtracted*, like the index sets  $\{2, 3, 4, 5, 6\}, \{7, 8, 9, 2, 3\}, \dots$  in (B.123).



Define

$$R_a^\lambda := \sum_{k \in K_a} p_k^\lambda \quad \text{and} \quad T_b^\lambda := \sum_{l \in L_b} p_l^\lambda. \quad (\text{B.129})$$

Then

$$\sum_{i \in J} p_i^\lambda = \frac{\sum_{a=0}^{r/2-1} R_a^\lambda - \sum_{b=0}^{r/2-\#J-1} T_b^\lambda}{r/2}, \quad (\text{B.130})$$

and

$$\begin{aligned} \int d\mu(\lambda) \left| \sum_{i \in J} p_i^\lambda - \#J/r \right| &= \frac{1}{r/2} \int_{\Lambda} d\mu(\lambda) \left| \sum_{a=0}^{r/2-1} R_a^\lambda - \sum_{b=1}^{r/2-\#J} T_b^\lambda - \#J/2 \right| \\ &= \int_{\Lambda} d\mu(\lambda) \frac{1}{r/2} \left| \sum_{a=0}^{r/2-1} (R_a^\lambda - 1/2) - \sum_{b=0}^{r/2-\#J-1} (T_b^\lambda - 1/2) \right| \\ &\leq \int_{\Lambda} d\mu(\lambda) \frac{1}{r/2} \left( \sum_{a=0}^{r/2-1} |R_a^\lambda - 1/2| + \sum_{b=0}^{r/2-\#J-1} |T_b^\lambda - 1/2| \right) \\ &= \frac{1}{r/2} \left( \sum_{a=0}^{r/2-1} \int_{\Lambda} d\mu(\lambda) |R_a^\lambda - 1/2| + \sum_{b=0}^{r/2-\#J-1} \int_{\Lambda} d\mu(\lambda) |T_b^\lambda - 1/2| \right) \\ &\leq \frac{r - \#J}{r/2} \epsilon < 2\epsilon, \end{aligned} \quad (\text{B.131})$$

where we used (B.130) and (B.127). For the case  $\#J > r/2$ , define  $K := I^{\text{ind}} \setminus J$ . Then  $\#K = r - \#J < r/2$  and

$$\begin{aligned} \int_{\Lambda} d\mu(\lambda) \left| \sum_{i \in J} p_i^\lambda - \#J/r \right| &= \int_{\Lambda} d\mu(\lambda) \left| \left( 1 - \sum_{i \in K} p_i^\lambda \right) - (1 - \#K/r) \right| \\ &= \int_{\Lambda} d\mu(\lambda) \left| \sum_{i \in K} p_i^\lambda - \#K/r \right| < 2\epsilon. \end{aligned} \quad (\text{B.132})$$

Concluding, for any  $J \subset I^{\text{ind}}$ ,

$$\int_{\Lambda} d\mu(\lambda) \left| \sum_{i \in J} p_i^\lambda - \#J/r \right| < 2\epsilon. \quad (\text{B.133})$$

## C Embezzlement

Recall the definitions:

$$\begin{aligned}
|\chi_{n,l}\rangle &:= |\tau_n\rangle_{A''B''} \otimes \sum_{i=0}^{d-1} \sum_{j=0}^{m_{i,l}-1} \frac{c_i}{\sqrt{m_{i,l}}} |i, j\rangle_{AA'} |i, j\rangle_{BB'}; \\
&= \sum_{i=0}^{d-1} \sum_{k=0}^{n-1} \sum_{j=0}^{m_{i,l}-1} \frac{c_i}{\sqrt{C_n m_{i,l} (k+1)}} |i\rangle_A |i\rangle_B |j\rangle_{A'} |j\rangle_{B'} |k\rangle_{A''} |k\rangle_{B''}. \\
|\Psi_n\rangle &:= |\tau_n\rangle_{A''B''} \otimes |0\rangle_{A'} |0\rangle_{B'} \otimes \sum_{i=0}^{d-1} c_i |i\rangle_A |i\rangle_B; \\
|\tau_n\rangle &:= \frac{1}{\sqrt{C_n}} \sum_{k=0}^{n-1} \frac{1}{\sqrt{k+1}} |k\rangle \otimes |k\rangle; \\
C_n &:= \sum_{k=0}^{n-1} \frac{1}{k+1}; \\
r_l &:= \text{LCD} \left( \{c_{i,l}^2\}_{i=0}^{d-1} \cup 1/2 \right) \\
m_{i,l} &:= r_l c_{i,l}^2
\end{aligned} \tag{C.134}$$

Note that the derivation in this section also holds without the subscripts  $l$ . In that case the definition of  $r_l$  is replaced by

$$r := \text{LCD} \left( \{c_i^2\}_{i=0}^{d-1} \right); \tag{C.135}$$

Define  $\lfloor x \rfloor$  as the *floor* of  $x$ : the greatest integer not greater than  $x$ , and  $\lceil x \rceil$  as the *ceiling* of  $x$ : the smallest integer not smaller than  $x$ . Define the unitary operator  $U_{n,l}^{AA'A''}$  with, for  $k \in \mathbb{N}_n$  and  $i \in \mathbb{N}_d$ :

$$U_{n,l}^{AA'A''} (|k\rangle_{A''} |0\rangle_{A'} |i\rangle_A) = |\lfloor k/m_{i,l} \rfloor\rangle_{A''} |k \bmod m_{i,l}\rangle_{A'} |i\rangle_A, \tag{C.136}$$

and an arbitrary action on the other basis vectors.  $U_{n,l}^{BB'B''}$  is defined similarly. Then,

$$\begin{aligned}
&\left( U_{n,l}^{AA'A''} \otimes U_{n,l}^{BB'B''} \right) |\Psi_n\rangle \\
&= \sum_{i=0}^{d-1} \sum_{k=0}^{n-1} \frac{c_i}{\sqrt{C_n (k+1)}} |i\rangle_A |i\rangle_B |k \bmod m_{i,l}\rangle_{A'} |k \bmod m_{i,l}\rangle_{B'} |\lfloor k/m_{i,l} \rfloor\rangle_{A''} |\lfloor k/m_{i,l} \rfloor\rangle_{B''}.
\end{aligned} \tag{C.137}$$

Calculating the fidelity:

$$\begin{aligned}
\mathcal{F}\left(|\chi_{n,l}\rangle, \left(U_{n,l}^{AA'A''} \otimes U_{n,l}^{BB'B''}\right)|\Psi_n\rangle\right) &= |\langle\chi_{n,l}|\left(U_{n,l}^{AA'A''} \otimes U_{n,l}^{BB'B''}\right)|\Psi_n\rangle| \\
&= \sum_{i=0}^{d-1} \sum_{k=0}^{n-1} \frac{c_i}{\sqrt{C_n(k+1)}} \frac{c_i}{\sqrt{C_n m_{i,l}(\lfloor k/m_{i,l}\rfloor + 1)}} \\
&= \sum_{i=0}^{d-1} \sum_{k=0}^{n-1} \frac{c_i^2}{C_n \sqrt{m_{i,l}(k+1)(\lfloor k/m_{i,l}\rfloor + 1)}}.
\end{aligned} \tag{C.138}$$

Now, using  $k+1 \leq m_{i,l}^2 \lceil (k+1)/m_{i,l} \rceil$  and  $\lfloor k/m_{i,l} \rfloor + 1 \leq \lceil (k+1)/m_{i,l} \rceil$ ,

$$\begin{aligned}
&\sum_{k=0}^{n-1} \frac{1}{\sqrt{m_{i,l}(k+1)(\lfloor k/m_{i,l}\rfloor + 1)}} \\
&\geq \sum_{k=0}^{n-1} \frac{1}{\sqrt{m_{i,l}^2 \lceil (k+1)/m_{i,l} \rceil (\lfloor k/m_{i,l}\rfloor + 1)}} \geq \sum_{k=0}^{n-1} \frac{1}{m_{i,l} \lceil (k+1)/m_{i,l} \rceil} \\
&= \sum_{k=0}^{m_{i,l}-1} \frac{1}{m_{i,l}} + \sum_{k=m_{i,l}}^{2m_{i,l}-1} \frac{1}{2m_{i,l}} + \cdots + \sum_{k=\lfloor n/m_{i,l} \rfloor m_{i,l}}^{n-1} \frac{1}{(\lfloor n/m_{i,l} \rfloor + 1)m_{i,l}} \\
&= \left( \sum_{k'=1}^{\lfloor n/m_{i,l} \rfloor} \frac{1}{k'} \right) + \left( \frac{n}{m_{i,l}} - \lfloor n/m_{i,l} \rfloor \right) \frac{1}{\lfloor n/m_{i,l} \rfloor + 1} \\
&= \mathcal{Z}\left(\frac{n}{m_{i,l}}\right), \text{ where } \mathcal{Z}(y) := \left( \sum_{k=1}^{\lfloor y \rfloor} \frac{1}{k} \right) + (y - \lfloor y \rfloor) \frac{1}{\lfloor y \rfloor + 1}
\end{aligned} \tag{C.139}$$

Note that for integer values of  $y$ ,  $\mathcal{Z}(y) = \sum_{k=1}^y 1/k = C_y$ . For non-integer values,  $\mathcal{Z}(y)$  equals the grey area in Figure 2 between  $x = 0$  and  $x = y$ . From this we can read off the inequalities:

$$\mathcal{Z}(y) \leq 1 + \int_1^y \frac{1}{x} dx = 1 + \ln y \tag{C.140}$$

$$\mathcal{Z}(y) \geq \int_0^y \frac{1}{x+1} dx = \ln(y+1) \tag{C.141}$$

$$\mathcal{Z}(y_2) - \mathcal{Z}(y_1) \leq \int_{y_1}^{y_2} \frac{1}{x} dx = \ln y_2 - \ln y_1. \tag{C.142}$$

Now, noting that

$$\mathcal{F}\left(|\chi_{n,l}\rangle, \left(U_{n,l}^{AA'A''} \otimes U_{n,l}^{BB'B''}\right)|\Psi_n\rangle\right) = \sum_{i=0}^{d-1} c_i^2 \frac{\mathcal{Z}\left(\frac{n}{m_{i,l}}\right)}{\mathcal{Z}(n)} \tag{C.143}$$

and

$$\frac{\mathcal{Z}\left(\frac{n}{m_{i,l}}\right)}{\mathcal{Z}(n)} = 1 - \frac{\mathcal{Z}(n) - \mathcal{Z}\left(\frac{n}{m_{i,l}}\right)}{\mathcal{Z}(n)} > 1 - \frac{\ln m_{i,l}}{\ln n}, \tag{C.144}$$

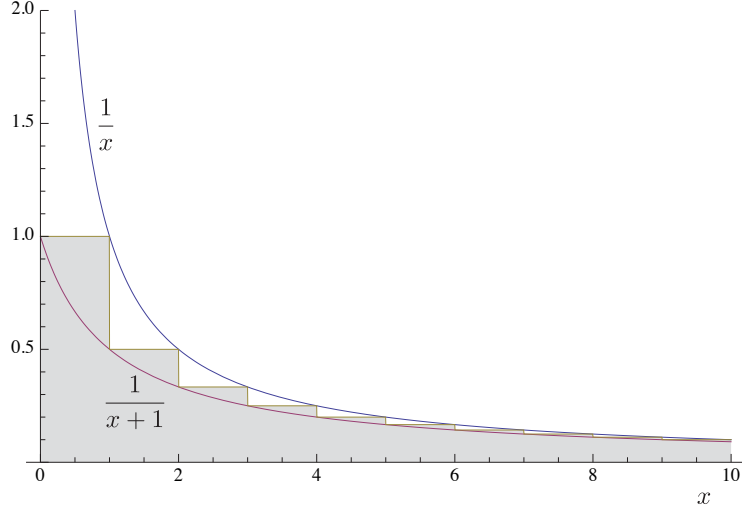


Figure 2: Using the two plotted graphs, bounds for the grey area can be derived.

it follows, using (6.43), that

$$\begin{aligned}
 D\left(|\chi_{n,l}\rangle, \left(U_{n,l}^{AA'A''} \otimes U_{n,l}^{BB'B''}\right) |\Psi_n\rangle\right) &< \sqrt{1 - \left(1 - \frac{\ln m_{i,l}}{\ln n}\right)^2} \\
 \Rightarrow \lim_{n \rightarrow \infty} D\left(|\chi_{n,l}\rangle, \left(U_{n,l}^{AA'A''} \otimes U_{n,l}^{BB'B''}\right) |\Psi_n\rangle\right) &= 0. \tag{C.145}
 \end{aligned}$$

## D Perfect Correlation

Quantum mechanics predicts that there is perfect correlation between the outcomes of certain measurements on  $A$  and  $B$ . For example, for a measurement on the state  $|\phi_d\rangle_{AB} = \sum_{i=0}^{d-1} c_i |i\rangle_A |i\rangle_B$ , when the outcome on one side corresponds to the projector  $[i]^A$ , the outcome on the other side corresponds to the projector  $[i]^B$ , and vice versa. Therefore, the probabilities of these outcomes must be equal. This property remains when considering  $\lambda$ -probabilities. Similar properties hold for more complex measurements on  $AA'A''$  and  $BB'B''$ . What follows is a formal proof of this fact. First, recall the definitions

$$\begin{aligned}
 |\Psi_n\rangle &:= |\tau_n\rangle_{A'B''} \otimes |0\rangle_{A'} |0\rangle_{B'} \otimes |\phi_d\rangle_{AB}, \\
 \hat{E}_{(i,j),n,l} &:= \left(U_{n,l}^{AA'A''}\right)^{-1} \left([i,j]^{AA'} \otimes \mathbb{I}^{A''}\right) \left(U_{n,l}^{AA'A''}\right), \\
 \hat{F}_{(i,j),n,l} &:= \left(U_{n,l}^{BB'B''}\right)^{-1} \left([i,j]^{BB'} \otimes \mathbb{I}^{B''}\right) \left(U_{n,l}^{BB'B''}\right). \tag{D.146}
 \end{aligned}$$

**Lemma 2** (Perfect correlation). *The following identities hold:*

$$\forall I \subset \mathbb{N}_d: \Pr_\lambda^{|\phi_d\rangle_{AB}} \left( \sum_{i \in I} [i]^A \right) \doteq \sum_{i \in I} \Pr_\lambda^{|\phi_d\rangle_{AB}} ([i]^B), \quad (\text{D.147})$$

$$\forall I_l \subset J_l^{\text{ind}}: \Pr_\lambda^{|\Psi_n\rangle} \left( \sum_{(i,j) \in I_l} \hat{E}_{(i,j),n,l} \right) \doteq \sum_{(i,j) \in I_l} \Pr_\lambda^{|\Psi_n\rangle} (\hat{F}_{(i,j),n,l}) \quad (\text{D.148})$$

$$\sum_{j=0}^{m_{i,l}-1} \Pr_\lambda^{|\Psi_n\rangle} (\hat{E}_{(i,j),n,l}) \doteq \Pr_\lambda^{|\Psi_n\rangle} ([i]^B), \text{ and} \quad (\text{D.149})$$

$$\sum_{j=0}^{m_{i,l}-1} \Pr_\lambda^{|\Psi_n\rangle} (\hat{F}_{(i,j),n,l}) \doteq \Pr_\lambda^{|\Psi_n\rangle} ([i]^A). \quad (\text{D.150})$$

Note that the lemma above and the proof below also hold with the subscripts  $l$  removed. In this proof, probabilities are considered for the event that an outcome does *not* correspond to a projector  $[i]$ . We denote such an event as  $\neg[i]$ . Also, probabilities are considered for the event that an outcome does not correspond to any projector from a set indexed by  $I$ . In that case, the event will be written as  $\neg \bigvee_{i \in I} [i]$ .

First, note that according to the Born rule,

$$\sum_{i' \in I} \Pr^{|\phi_d\rangle_{AB}} \left( \sum_{i \in I} [i]^A, [i']^B \right) = \sum_{i' \in I} \langle \phi_d | \sum_{i \in I} [i]^A \otimes [i']^B | \phi_d \rangle_{AB}, \quad (\text{D.151})$$

where  $|\phi_d\rangle_{AB}$  has the form

$$|\phi_d\rangle_{AB} = \sum_{i=0}^{d-1} c_i |i\rangle_A |i\rangle_B. \quad (\text{D.152})$$

Using this and applying CompQuant, it follows that

$$\begin{aligned} \Pr^{|\phi_d\rangle_{AB}} \left( \sum_{i \in I} [i]^A, \neg \bigvee_{i' \in I} [i']^B \right) = 0 &\Rightarrow \Pr_\lambda^{|\phi_d\rangle_{AB}} \left( \sum_{i \in I} [i]^A, \neg \bigvee_{i' \in I} [i']^B \right) \doteq 0, \\ \sum_{i' \in I} \Pr^{|\phi_d\rangle_{AB}} \left( \neg \sum_{i \in I} [i]^A, [i']^B \right) = 0 &\Rightarrow \sum_{i' \in I} \Pr_\lambda^{|\phi_d\rangle_{AB}} \left( \neg \sum_{i \in I} [i]^A, [i']^B \right) \doteq 0. \end{aligned} \quad (\text{D.153})$$

Then, using  $\Pr(X) = \sum_{i \in I} \Pr(X, Y_i) + \Pr(X, \neg \bigvee_{i \in I} Y_i)$  for a set of mutually exclusive events  $\{Y_i\}_{i \in I}$ , we have for all  $\hat{O}^A, \hat{O}^B$ ,

$$\begin{aligned} \Pr_\lambda^{|\phi_d\rangle_{AB}} \left( \sum_{i \in I} [i]^A \right)_{\hat{O}^A} &= \sum_{i' \in I} \Pr_\lambda^{|\phi_d\rangle_{AB}} \left( \sum_{i \in I} [i]^A, [i']^B \right)_{\hat{O}^A \otimes \hat{O}^B} + \Pr_\lambda^{|\phi_d\rangle_{AB}} \left( \sum_{i \in I} [i]^A, \neg \bigvee_{i' \in I} [i']^B \right)_{\hat{O}^A \otimes \hat{O}^B} \\ &\doteq \sum_{i' \in I} \Pr_\lambda^{|\phi_d\rangle_{AB}} \left( \sum_{i \in I} [i]^A, [i']^B \right)_{\hat{O}^A \otimes \hat{O}^B} + \sum_{i' \in I} \Pr_\lambda^{|\phi_d\rangle_{AB}} \left( \neg \sum_{i \in I} [i]^A, [i']^B \right)_{\hat{O}^A \otimes \hat{O}^B} \\ &= \sum_{i \in I} \Pr_\lambda^{|\phi_d\rangle_{AB}} ([i]^B)_{\hat{O}^B}. \end{aligned} \quad (\text{D.154})$$

By setting  $I = \{i\}$  we have, for all  $\hat{O}^A, \hat{O}^B$  and for all  $i$ ,

$$\Pr_\lambda^{|\phi_d\rangle^{AB}} ([i]^A)_{\hat{O}^A} = \Pr_\lambda^{|\phi_d\rangle^{AB}} ([i]^B)_{\hat{O}^B}, \quad (\text{D.155})$$

which means that both probabilities are context-independent, so the observables can be dropped. Therefore, from (D.154) we arrive at

$$\Pr_\lambda^{|\phi_d\rangle^{AB}} \left( \sum_{i \in I} [i]^A \right) \doteq \sum_{i \in I} \Pr_\lambda^{|\phi_d\rangle^{AB}} ([i]^B). \quad (\text{D.156})$$

To prove (D.148), note that

$$\begin{aligned} & \Pr^{|\Psi_n\rangle} \left( \sum_{(i,j) \in I_l} \hat{E}_{(i,j),n,l}, \hat{F}_{(i',j'),n,l} \right) = \\ & \langle \Psi_n | \left( U_{n,l}^{AA'A''} \otimes U_{n,l}^{BB'B''} \right)^{-1} \left( \sum_{(i,j) \in I_l} [i,j]^{AA'} \otimes [i',j']^{BB'} \otimes \mathbb{I}^{A''B''} \right) \left( U_{n,l}^{AA'A''} \otimes U_{n,l}^{BB'B''} \right) | \Psi_n \rangle, \end{aligned} \quad (\text{D.157})$$

where, by (C.136), writing  $[x]$  for the *floor* of  $x$ , which is the greatest integer not greater than  $x$ ,

$$\left( U_{n,l}^{AA'A''} \otimes U_{n,l}^{BB'B''} \right) | \Psi_n \rangle = \sum_{i=0}^{d-1} \sum_{k=0}^{n-1} c_{i,k} |i\rangle_A |i\rangle_B |k \bmod m_{i,l}\rangle_{A'} |k \bmod m_{i,l}\rangle_{B'} |[k/m_{i,l}]\rangle_{A''} |[k/m_{i,l}]\rangle_{B''}. \quad (\text{D.158})$$

Using a similar strategy as above, we get

$$\begin{aligned} & \Pr^{|\Psi_n\rangle} \left( \sum_{(i,j) \in I_l} \hat{E}_{(i,j),n,l}, \neg \bigvee_{(i',j') \in I_l} \hat{F}_{(i',j'),n,l} \right) = 0 \\ \Rightarrow & \Pr_\lambda^{|\Psi_n\rangle} \left( \sum_{(i,j) \in I_l} \hat{E}_{(i,j),n,l}, \neg \bigvee_{(i',j') \in I_l} \hat{F}_{(i',j'),n,l} \right) \doteq 0; \\ & \sum_{(i',j') \in I_l} \Pr^{|\Psi_n\rangle} \left( \neg \sum_{(i,j) \in I_l} \hat{E}_{(i,j),n,l}, \hat{F}_{(i',j'),n,l} \right) = 0 \\ \Rightarrow & \sum_{(i',j') \in I_l} \Pr_\lambda^{|\Psi_n\rangle} \left( \neg \sum_{(i,j) \in I_l} \hat{E}_{(i,j),n,l}, \hat{F}_{(i',j'),n,l} \right) \doteq 0. \end{aligned} \quad (\text{D.159})$$

Then, for all  $\hat{O}^{AA'A''}, \hat{O}^{BB'B''}$ ,

$$\begin{aligned}
& \Pr_{\lambda}^{|\Psi_n\rangle} \left( \sum_{(i,j) \in I_l} \hat{E}_{(i,j),n,l} \right)_{\hat{O}^{AA'A''}} \\
&= \sum_{(i',j') \in I_l} \Pr_{\lambda}^{|\Psi_n\rangle} \left( \sum_{(i,j) \in I_l} \hat{E}_{(i,j),n,l}, \hat{F}_{(i',j'),n,l} \right)_{\hat{O}^{AA'A''} \otimes \hat{O}^{BB'B''}} \\
&+ \Pr_{\lambda}^{|\Psi_n\rangle} \left( \sum_{(i,j) \in I_l} \hat{E}_{(i,j),n,l}, \neg \bigvee_{(i',j') \in I_l} \hat{F}_{(i',j'),n,l} \right)_{\hat{O}^{AA'A''} \otimes \hat{O}^{BB'B''}} \\
&\doteq \sum_{(i',j') \in I_l} \Pr_{\lambda}^{|\Psi_n\rangle} \left( \sum_{(i,j) \in I_l} \hat{E}_{(i,j),n,l}, \hat{F}_{(i',j'),n,l} \right)_{\hat{O}^{AA'A''} \otimes \hat{O}^{BB'B''}} \\
&+ \sum_{(i',j') \in I_l} \Pr_{\lambda}^{|\Psi_n\rangle} \left( \neg \sum_{(i,j) \in I_l} \hat{E}_{(i,j),n,l}, \hat{F}_{(i',j'),n,l} \right)_{\hat{O}^{AA'A''} \otimes \hat{O}^{BB'B''}} \\
&= \sum_{(i,j) \in I_l} \Pr_{\lambda}^{|\Psi_n\rangle} \left( \hat{F}_{(i,j),n,l} \right)_{\hat{O}^{BB'B''}}. \tag{D.160}
\end{aligned}$$

By setting  $I_l = (i, j)$ , we get, for all  $i, j$  and for all  $\hat{O}^{AA'A''}, \hat{O}^{BB'B''}$ ,

$$\Pr_{\lambda}^{|\Psi_n\rangle} \left( \hat{E}_{(i,j),n,l} \right)_{\hat{O}^{AA'A''}} \doteq \Pr_{\lambda}^{|\Psi_n\rangle} \left( \hat{F}_{(i,j),n,l} \right)_{\hat{O}^{BB'B''}}. \tag{D.161}$$

As in (D.155), this means the probabilities are context-independent, and (D.148) follows from (D.160).

Similarly, to prove (D.149), consider the probability

$$\begin{aligned}
& \Pr^{|\Psi_n\rangle} \left( \hat{E}_{(i,j),n,l}, [i']^B \right) = \\
& \langle \Psi_n | \left( U_{n,l}^{AA'A''} \otimes \mathbb{I}^{BB'B''} \right)^{-1} \left( [i, j]^{AA'} \otimes [i']^B \otimes \mathbb{I}^{A''B'B''} \right) \left( U_{n,l}^{AA'A''} \otimes \mathbb{I}^{BB'B''} \right) | \Psi_n \rangle, \tag{D.162}
\end{aligned}$$

where, by (C.136),

$$\left( U_{n,l}^{AA'A''} \otimes \mathbb{I}^{BB'B''} \right) | \Psi_n \rangle = \sum_{i=0}^{d-1} \sum_{k=0}^{n-1} c_{i,k} |i\rangle_A |k \bmod m_{i,l}\rangle_{A'} |[k/m_{i,l}]\rangle_{A''} \otimes |i\rangle_B |0\rangle_{B'} |k\rangle_{B''}. \tag{D.163}$$

Again, using CompQuant, it follows that

$$\begin{aligned}
& \sum_{j=0}^{m_{i,l}-1} \Pr^{|\Psi_n\rangle} \left( \hat{E}_{(i,j),n,l}, \neg [i]^B \right) = 0 \quad \Rightarrow \quad \sum_{j=0}^{m_{i,l}-1} \Pr_{\lambda}^{|\Psi_n\rangle} \left( \hat{E}_{(i,j),n,l}, \neg [i]^B \right) \doteq 0; \\
& \Pr^{|\Psi_n\rangle} \left( \neg \bigvee_{j=0}^{m_{i,l}-1} \hat{E}_{(i,j),n,l}, [i]^B \right) = 0 \quad \Rightarrow \quad \Pr_{\lambda}^{|\Psi_n\rangle} \left( \neg \bigvee_{j=0}^{m_{i,l}-1} \hat{E}_{(i,j),n,l}, [i]^B \right) \doteq 0. \tag{D.164}
\end{aligned}$$

Then, for all  $\hat{O}^{AA'A''}, \hat{O}^B$ ,

$$\begin{aligned}
& \sum_{j=0}^{m_{i,l}-1} \Pr_{\lambda}^{|\Psi_n\rangle} \left( \hat{E}_{(i,j),n,l} \right)_{\hat{O}^{AA'A''}} \\
&= \sum_{j=0}^{m_{i,l}-1} \Pr_{\lambda}^{|\Psi_n\rangle} \left( \hat{E}_{(i,j),n,l}, [i]^B \right)_{\hat{O}^{AA'A''} \otimes \hat{O}^B} + \sum_{j=0}^{m_{i,l}-1} \Pr_{\lambda}^{|\Psi_n\rangle} \left( \hat{E}_{(i,j),n,l}, \neg [i]^B \right)_{\hat{O}^{AA'A''} \otimes \hat{O}^B} \\
&\doteq \sum_{j=0}^{m_{i,l}-1} \Pr_{\lambda}^{|\Psi_n\rangle} \left( \hat{E}_{(i,j),n,l}, [i]^B \right)_{\hat{O}^{AA'A''} \otimes \hat{O}^B} + \Pr_{\lambda}^{|\Psi_n\rangle} \left( \neg \bigvee_{j=0}^{m_{i,l}-1} \hat{E}_{(i,j),n,l}, [i]^B \right)_{\hat{O}^{AA'A''} \otimes \hat{O}^B} \\
&= \Pr_{\lambda}^{|\Psi_n\rangle} ([i]^B)_{\hat{O}^B}. \tag{D.165}
\end{aligned}$$

Using the context-independence of the probability on the left-hand side of (D.161), we arrive at (D.149). Since the above derivation also holds when all  $A$  and  $B$  are interchanged, we immediately also get (D.150).

## References

- Barrett, J., Kent, A., & Pironio, S. (2006). Maximally nonlocal and monogamous quantum correlations. *Physical Review Letters*, *97*, 170409. doi:10.1103/PhysRevLett.97.170409.
- Barrett, J. A. (1999). *The quantum mechanics of minds and worlds*. Oxford: Oxford University Press.
- Bell, J. S. (1964). On the Einstein–Podolsky–Rosen paradox. *Physics*, *1*, 195–200.
- Bell, J. S. (2004). The theory of local beables. In *Speakable and Unspeakable in Quantum Mechanics*. Cambridge: Cambridge University Press.
- Bohm, D. (1952). A suggested interpretation of the quantum theory in terms of “hidden” variables. I and II. *Physical Review*, *85*, 166–179. doi:10.1103/PhysRev.85.166.
- Branciard, C., Brunner, N., Gisin, N., Kurtsiefer, C., Lamas-Linares, A., Ling, A., & Scarani, V. (2008). Testing quantum correlations versus single-particle properties within Leggett’s model and beyond. *Nature Physics*, *4*, 681–685. doi:10.1038/nphys1020.
- Braunstein, S. L., & Caves, C. M. (1989). Wringing out better Bell inequalities. *Nuclear Physics B - Proceedings Supplements*, *6*, 211 – 221. doi:10.1016/0920-5632(89)90441-6.
- Busch, P., Lahti, P., & Mittelstaedt, P. (1996). *The Quantum Theory of Measurement*. Lecture Notes in Physics Monographs. Springer Berlin Heidelberg.
- Butterfield, J. (1992). Bell’s theorem: What it takes. *The British Journal for the Philosophy of Science*, *43*, 41–83. doi:10.1093/bjps/43.1.41.



- Colbeck, R. (2010). Frequently asked (and answered) questions for “Quantum theory cannot be extended”. URL: <http://perimeterinstitute.ca/personal/rcolbeck/FAQ.html> (retrieved on March 24, 2015).
- Colbeck, R., & Renner, R. (2011). No extension of quantum theory can have improved predictive power. *Nature Communications*, *2*, 411. doi:10.1038/ncomms1416.
- Colbeck, R., & Renner, R. (2012a). Is a system’s wave function in one-to-one correspondence with its elements of reality? *Physical Review Letters*, *108*, 150402. doi:10.1103/PhysRevLett.108.150402.
- Colbeck, R., & Renner, R. (2012b). The completeness of quantum theory for predicting measurement outcomes. *ArXiv e-prints*, . arXiv:1208.4123.
- Colbeck, R., & Renner, R. (2013). A short note on the concept of free choice. *ArXiv e-prints*, . arXiv:1302.4446.
- van Dam, W., & Hayden, P. (2003). Universal entanglement transformations without communication. *Physical Review A*, *67*, 060302. doi:10.1103/PhysRevA.67.060302.
- Di Lorenzo, A. (2012). Determination of hidden-variable models reproducing the spin singlet. *Physical Review A*, *86*, 042119. doi:10.1103/PhysRevA.86.042119.
- Einstein, A., Podolsky, A., & Rosen, N. (1935). Can quantum-mechanical description of physical reality be considered complete? *Physical Review*, *47*, 777. doi:10.1103/PhysRev.47.777.
- Fine, A. (1982). Hidden variables, joint probability, and the Bell inequalities. *Physical Review Letters*, *48*, 291–295. doi:10.1103/PhysRevLett.48.291.
- Ghirardi, G., & Romano, R. (2012). Local parts of hidden variable models for maximally entangled bipartite qudits. *Physical Review A*, *86*, 022107. doi:10.1103/PhysRevA.86.022107.
- Ghirardi, G., & Romano, R. (2013a). About possible extensions of quantum theory. *Foundations of Physics*, *43*, 881–894. doi:10.1007/s10701-013-9724-6.
- Ghirardi, G., & Romano, R. (2013b). Comment on “Is a system’s wave function in one-to-one correspondence with its elements of reality?” [arXiv:1111.6597]. *ArXiv e-prints*, . arXiv:1302.1635.
- Gleason, A. (1957). Measures on the closed subspaces of a Hilbert space. *Indiana University Mathematics Journal*, *6*, 885–893.
- Kochen, S., & Specker, E. (1975). The problem of hidden variables in quantum mechanics. In *The Logico-Algebraic Approach to Quantum Mechanics* (pp. 293–328). Dordrecht: D. Reidel Publishing Company.
- Landsman, K. (2015). On the Colbeck–Renner theorem. *Journal of Mathematical Physics*, *56*. doi:10.1063/1.4936556.

- Leggett, A. (2003). Nonlocal hidden-variable theories and quantum mechanics: An incompatibility theorem. *Foundations of Physics*, 33, 1469–1493. doi:10.1023/A:1026096313729.
- Leifer, M. (2014). Is the quantum state real? an extended review of -ontology theorems. *Quanta*, 3, 67–155. doi:10.12743/quanta.v3i1.22.
- Nielsen, M. A., & Chuang, I. L. (2004). *Quantum Computation and Quantum Information*. Cambridge: Cambridge University Press.
- Pearle, P. M. (1970). Hidden-variable example based upon data rejection. *Physical Review D*, 2, 1418–1425. doi:10.1103/PhysRevD.2.1418.
- Pusey, M. F., Barrett, J., & Rudolph, T. (2012). On the reality of the quantum state. *Nat Phys*, 8, 475–478. doi:10.1038/nphys2309.
- Scarani, V. (2013). Physics and the bumper sticker. URL: <https://spreadquantum.wordpress.com/2013/06/26/physics-and-the-bumper-sticker/> (Retrieved on March 24, 2015).
- Shimony, A. (1984). Controllable and uncontrollable non-locality. *Foundations of Quantum Mechanics in the Light of New Technology, Tokyo. The Physical Society of Japan. Repr. in Shimony, A. Search for a Naturalistic World View*, 2, 130–139.