On No Signalling Condition and the Probability Distribution of All Outcomes in a Local Hidden Variable Model[1]

Soe Moe Thar
Nanyang Technological University

29 October 2014

[1]Undergraduate dissertation written under the supervision of Assistant Professor Tomasz Paterek

Abstract

I discuss the method to construct a complete probability distribution describing the likelihood of all possible outcomes in an experiment even those unmeasured in a single experimental run. The experiment consists of measuring a system at two different places which are space-like separated. The existence of all the outcomes in each run of the experiment and no signalling condition as well as freedom of the experimenters to choose a setting is assumed. In addition, different instances of no signalling condition are used to derive two different inequalities, one being the Bell's/CHSH inequality. A general method for the derivation of possibly new Bell's inequalities is briefly discussed.

Acknowledgement

I would like to thank Professor Tomasz Paterek whose contribution, commitment and supervision were phenomenal in this work.

Contents

1 Introduction

The notion of *reality* and *determinism* could be traced back to classical antiquity where ancient philosophers started to ponder and question on the existence of entities and causal relationships between them. Those ideas, of reality and determinism, were embedded in classical physics until the development of quantum mechanics in 20^{th} century. Such advancement, however, was responded with reappraisals, refutations and recognitions among physics community concerning the preceding classical ideas. Principles of quantum mechanics, especially of the ubiquitous Copenhagen interpretation, fundamentally challenge the classical ideas of determinism as well as that of reality. This nature of the new theory encouraged the proponents of determinism and classical realism to devise a way to reconcile the haphazard world of quantum mechanics with that of deterministic macroscopic world. Albert Einstein, a proponent of *causality* and *objective reality*, and other prominent figures such as John von Neumann and David Bohm, in their effort in this reconcilation, hypothesized a possible local deterministic interpretation of quantum mechanics where the seemingly random results from the measurements are, in fact, influenced by a variable unaccessible at least to the experimenters and not present in quantum formalism. The name *hidden variable* was given to those variables which preserve realism as well as maintain determinism. Earlier hypotheses of such hidden variable theories could be traced as far back as 1927 to pilot wave theory presented by Louis de Broglie. In 1935, Einstein with two coauthors published a paper [1] in which they pointed out that quantum mechanics is incomplete based on the idea of *objective reality*. Their argument is most widely known as EPR paradox. There were counter arguments and critiques on EPR's assumption of objective reality most notably by Niels Bohr [2]. Many in the physics community, for a fair reason, were not convinced of the reality assumption while there were other who believed that since the final observation could be arrived by adopting either of these viewpoints, it is a matter of preference which interpretation is taken [3].

But in 1964, a breakthrough occured. John Stewart Bell presented a paper [4] in which he discusses that assuming the *objective reality* as EPR did and hence accepting that quantum mechanics is incomplete, one could introduce the hypothetical hidden variables whose purpose is to restore the completeness of quantum theory. If this step is taken, he shows that the statistical predictions i.e. the average values of different combinations of observables is different from the predictions of quantum mechanics. In his proof, Bell made use of another interesting assumption of *locality*. It essentially states that the situation of one system is independent of the actions done on the other system which is spatially separated from the first system. Employing those two assumptions, he was able to show that the expectation value given by hidden variable theory is different from that of quantum mechanics. Thus if we are to assume existence of a hidden variable theory which yields the same statistical predictions as quantum mechianics, then in such theory actions on one system must have an immediate consequence on another system i.e. nonlocality must be true. Such nonlocality however is incompatible with special relativity. Bell later published another paper [5] where he proved a more general result now known as *Bell/CHSH inequalities*. Various experiments have been performed since then and are found to be violating the Bell inequalities [3] and are in good agreement with the predictions from quantum mechanics. One interesting aspect of Bell's inequalities is the different assumptions in deriving it. Over time, different assumptions were made to derive either Bell's inequality or joint probability distribution of all observables. Some of those assumptions such as locality, local causality, factorisibility of joint probability distribution and completeness condition [6] could be read in a meticulous review by Wiseman [7]. Proof that existence of a joint probability distribution of all observables is equivalent to the validity of Bell's inequality was presented in a paper by Arthur Fine [8].

The present paper also shows the validity of Bell's inequality by deriving a joint probability distribution based on three fair assumptions. But before we study the derivation, it is illuminating to read the mathematical derivations of Bell's inequalities and thus different assumptions behind it. Thus the second section of this paper is devoted to the derivations of Bell's inequality and instances of violations of such inequalities. Assumptions such as locality and local causality will be treated as mathematical properties in such derivations. In the third section, I present a new proof on the existence of joint probability distribution based on *no-signalling* principle and other two assumptions. The last section focuses on the possible further investigation from the current one.

2 A review on EPR paradox, derivations of Bell's inequalities and analysis on their violations

2.1 EPR paradox

Based on the hidden variable hypothesis, in 1935, Albert Einstein, Boris Podolsky and Nathan Rosen submitted a paper titled "Can Quantum-Mechanical

Description of Physical Reality Be Considered Complete?" to Physical Review journal and it was published in May within the same year. In that paper, the authors concluded that the description of *reality* given by quantum mechanics is not complete for its negation leads to a contradiction. In their argument, the authors employed a major premises and a minor premise. They also proposed a criterion for identifying reality. I shall list those premises and dissect on their argument. Before we delve into the premises behind Einstein et al's paper (from now on refered to as EPR's paper and the authors as EPR), I would like to point out an assumption taken by EPR, namely the assumption on *objective reality.* We shall not pursue the justification behind this assumption but rather we shall interpret it as the existence of properties of a system regardless of any measurement.

The major premise put forward by EPR is a *necessary* requirement for a complete theory.

> every element of the physical reality must have a counter-part in the physical theory

One method for determining *element of the physical reality* as suggested by EPR is:

> The elements of the physical reality cannot be determined by a *priori* philosophical considerations, but must be found by an appeal to results of experiments and measurements."

Here I believe EPR intended to mean all observable properties of a system whether they are either simultaneously or asynchronously observed. The criterion for identifying reality proposed by EPR is,

> If, *without in any way disturbing a system, we can predict with certainty (i.e., with probability equal to unity) the value of a physical quantity, then there exists an element of physical reality corresponding to this physical quantity.*

This criterion was utilized in the second part of EPR's paper. The minor premise implemented by EPR is that quantum mechanical predictions, so far, are correct and in quantum mechanics, if two operators corresponding to two physical quantites do not commute, then those quantities cannot be measured simultaneously. EPR's conclusion is as follow:

> (1) *the quantum mechanical description of reality given by the wave function is not complete* or (2) *when the operators corresponding to two physical quantities do not commute the two quantites cannot have simultaneous reality.*

Now the validity of EPR's argument will be discussed. EPR first assumed the second statement of the conclusion is false. They assumed it's the case that if two operators corresponding to two physical quantities do not commute, then the two quantities can have simultaneous reality. They also assumed the negation of the first statement of their conclusion: the quantum mechanical description of reality given by the wave function is complete. The latter assumption enables them to invoke the *necessary* requirement for a complete theory to conclude that those two quantities with simultaneous reality must have a counter-part in the physical theory which in this case is quantum mechanics. Moreover those counter-parts must exist simultaneously too. But this contradicts their minor premise and since it is improbable for the premises to be true and the conclusion false, their argument is deemed valid.

In the second part of EPR's paper, the authors demonstrated that the first statement of the conclusion must be true for its negation leads to the negation of the second statement of the conclusion and they have already shown the conclusion to be valid. Begin with the assumption that the quantum mechanical description of reality given by the wave function is complete. Let us consider a quantum system which splits into two subsystem, system I and system II. The system can be represented as

$$\Psi(x_1, x_2) = \sum_{n=1}^{\infty} \psi_n(x_2) u_n(x_1), \tag{2.1}$$

where $u_1(x_1), u_2(x_1), \ldots$ are eigenfunctions of some physical quantity A in system I and similarly $\psi_1(x_2), \psi_2(x_2), \ldots$ are eigenfunctions of quantity B in system II. Now it is possible for the system I and II to possess another physical quantities, say P and Q with eigenfunctions $v_1(x_1), v_2(x_1), \ldots$ and $\phi_1(x_2), \phi_2(x_2), \ldots$ respectively. In that case, the whole system can be represented as

$$\Psi(x_1, x_2) = \sum_{n=1}^{\infty} \phi_n(x_2) v_n(x_1), \tag{2.2}$$

We are assuming that quantities A and B are of the same nature (e.g. position) and so are quantities P and Q (e.g. momentum). Now it is possible, as EPR had demonstrated, that the quantities on system I, A and P are incompatible to each other—they can't be simultaneously measured. From the similar nature, quantities B and P are thus incompatible to each other. Now, if one were to make a measurement of A and obtains say a_m, the eigenvalue of the function $u_m(x_1)$, then one immediately knows that the system is left in the state,

$$\Phi(x_1, x_2) = \psi_m(x_2) u_m(x_1). \tag{2.3}$$

On the other hand, if one instead chooses to measure P and obtains p_k, the eigenvalue of the function $v_k(x_1)$, then immediately after the system is in the state,

$$\Phi(x_1, x_2) = \phi_k(x_2)v_k(x_1). \tag{2.4}$$

Now we can see that on system II, both the eigenfunction $\psi_m(x_2)$ and $\phi_k(x_2)$ belong to the same reality. But as we have mentioned, they are not compatible to each other. Thus, the negation of the first condition in the EPR's conclusion above leads to the negation of the other condition. Since the conclusion itself is proven to be valid and assuming the truth of the premises above, the conclusion must be true. Hence, it must be the case that the quantum mechanical description of reality given by the wave function is not complete.

Finally, EPR argued that why their criterion of reality is a reasonable assumption. They stated assuming it is the case that

> two or more physical quantities can be regarded as simultaneous elements of reality *only when they can be simultaneously measured or predicted.*

one could argue the physical quantities from two incompatible operators cannot have simultaneous element of reality since they cannot be measured simultaneously. But EPR debunked that in such case, the reality of those physical quantites then depends on the measurment on the other subsystem which is forbidden according to their no interaction assumption. In the conclusion of the paper, the authors stated that although their argument shows quantum mechanics is not a complete theory, the question of whether it could be completed or not is left open. However, they believed that a complete theory is possible.

2.2 Possiblility of hidden variable interpretation in a simple quantum system

Before we derive the Bell's inequality, we should examine whether it is possible to have a simple quantum system explained by a hidden variable model. The answer is positive for a two state quantum system. We present here a simple LHV model for a spin-$\frac{1}{2}$ particle. Assume that there exists $\lambda \in [0, 1]$ which imposes deterministic results on the measurements of all spin directions of the particle. Associate an observable $\hat{B}$ with a particular measurment, in this case spin direction of the particle. We choose the result of such measurement to be either $+1$ or -1 corresponding to spin up or spin down direction.

From quantum mechanics, we know that we could associate any mixed state ρ with a Bloch vector $\vec{s}$ and any observable $\hat{B}$ with a Bloch vector $\vec{m}$ so that,

$$\rho = \frac{1}{2}(1 + \vec{s} \cdot \vec{\sigma}) \tag{2.5}$$

$$\hat{B} = \frac{1}{2}(1 + \vec{m} \cdot \vec{\sigma}) \tag{2.6}$$

The average of the measurement $\hat{B}$ then is given by

$$^{QM}\langle \hat{B} \rangle_\rho = \vec{m} \cdot \vec{s}. \tag{2.7}$$

We define for the measurement result b as follows:

$$b = \begin{cases} +1 & \text{if } \lambda \in [0, \bar{\lambda}], \\ -1 & \text{if } \lambda \in [\bar{\lambda}, 1] \end{cases} \quad \text{where } \bar{\lambda} = \frac{1}{2}(1 + \vec{s} \cdot \vec{m}) \tag{2.8}$$

Then the probability of obtaining result +1 is

$$P(b = +1) = \int_0^{\bar{\lambda}} \mathrm{d}\lambda = \frac{1}{2}(1 + \vec{s} \cdot \vec{m}) \tag{2.9}$$

Similarly,

$$P(b = -1) = \int_{\bar{\lambda}}^{1} \mathrm{d}\lambda = \frac{1}{2}(1 - \vec{s} \cdot \vec{m}). \tag{2.10}$$

Average value of the measurement,

$$^{LHV}\langle \hat{B} \rangle_\rho = \sum_{i=+1,-1} P(b = i) = \vec{s} \cdot \vec{m} = {}^{QM}\langle \hat{B} \rangle_\rho. \tag{2.11}$$

We have shown here one LHV model that agrees with QM results on *average* for the measurement on a single spin- $\frac{1}{2}$ particle and which *completes* quantum mechanics in a sense that in every experimental run, the outcome is predetermined (just as in classical physics). Notice that in a particular run of the experiment, we have no knowledge about the value of λ hence the name "hidden variable" is given and we only need one variable in the model.

2.3 Bell's 1964 Paper

EPR's argument is sound provided that we accept the assumptions in it to be true. Indeed, there was at least one critique on those assumptions put forward by Niels Bohr. But in 1964, a paper titled "On The Einstein Podolsky Rosen Paradox" by John Stewart Bell was published in Physics and in

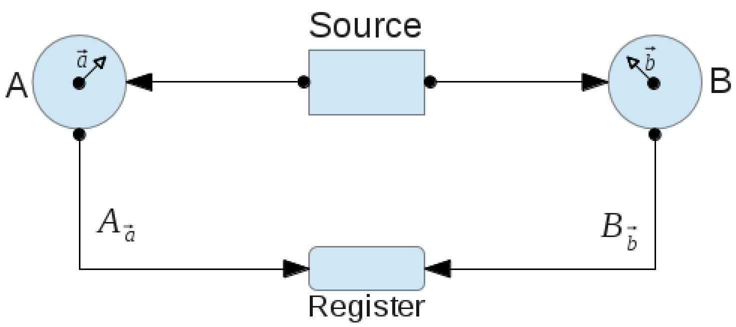


Figure 1: Schematic of the apparatus used in the derivation of Bell's inequality. The source sends out two entangled particles in the opposite directions towards the detectors. The observers can choose any setting at place A and B. After the settings $\vec{a}$ and $\vec{b}$ are chosen, the results $A_{\vec{a}}$ and $B_{\vec{b}}$ are sent out to the central register for the calculation of correlations.

this paper, the author showed that by assuming that quantum mechanics is not complete and consequently thus adopting quantities called *hidden variables* whose existence is to restore locality and determinism to the theory, the resultant statistical predictions by such theory could not be reconciled with that from quantum mechanis. Bell used the example by Bohm and Aharonov [9] on the explanation of EPR paradox using a pair of entangled spin-$\frac{1}{2}$ particles. First, he introduced hidden variable λ whose purpose is as mentioned to restore realism and determinism to the experimental outcomes. Furthermore, he declared that the results of an experiment say A and B are influenced by λ in the following way:

$$A(\vec{a}, \lambda) = \pm 1,\ B(\vec{b}, \lambda) = \pm 1. \tag{2.12}$$

where $\vec{a}$ and $\vec{b}$ are some unit vectors corresponding to the settings chosen by the experimenters in a spin measurements on a pair of entangled electrons. Here we are assuming the unit of spin to be $\frac{\hbar}{2}$. The *vital* assumption (2.12), he claimed, is that the result B on one side is not affected by setting $\vec{a}$ and similarly for A. This is the assumption which later is termed *locality* but it essentially encompasses the local casuality as envisaged by Einstein. But note that due to λ results A and B could still be related in some way and hence are not independent in general. Although the results of the experiment are determined by the hidden varaible λ and settings $\vec{a}$ and $\vec{b}$, the hidden variables themselves are considered to be probabilistic in the paper which can be seen by the introduction of probability distribution $\rho(\lambda)$ on λ. The expectation value of the product of the results is then

$$E(\vec{a}, \vec{b}) = \int d\lambda \rho(\lambda) A(\vec{a}, \lambda) B(\vec{b}, \lambda) \tag{2.13}$$

which should be equal to the quantum mechanical result

$$< \vec{\sigma} \cdot \vec{a}\, \vec{\sigma} \cdot \vec{b} >= -\vec{a} \cdot \vec{b}, \tag{2.14}$$

where we assumed the spins are described by the singlet state. But as the author showed, they are not equal.

2.3.1 The Bell's inequality

I shall present the first part of the Bell's Theorem where he showed that the expectation value in (2.13) is not equal to the quantum mechanical expectation value (2.14). Assuming a normalized probability distribution $\rho(\lambda)$ we have,

$$\int d\lambda\, \rho(\lambda) = 1. \tag{2.15}$$

Moreover, for a spin entangled pair of particles in the singlet state,

$$-A(\vec{b}, \lambda) = B(\vec{b}, \lambda), \tag{2.16}$$

i.e. their spins are anti-parallel to each other if measured along the same direction. In that case, equation (2.13) becomes

$$E(\vec{a}, \vec{b}) = -\int d\lambda\, \rho(\lambda) A(\vec{a}, \lambda) A(\vec{b}, \lambda). \tag{2.17}$$

If on one side, the experimenter chooses to measure the spin in another direction, say $\vec{c}$, then the difference in the expectation values between two scenarios is

$$\begin{aligned} E(\vec{a}, \vec{b}) - E(\vec{a}, \vec{c}) &= -\int d\lambda\, \rho(\lambda)[A(\vec{a}, \lambda)A(\vec{b}, \lambda) - A(\vec{a}, \lambda)A(\vec{c}, \lambda)] \\ &= \int d\lambda\, \rho(\lambda) A(\vec{a}, \lambda) A(\vec{b}, \lambda)[A(\vec{b}, \lambda)A(\vec{c}, \lambda) - 1] \end{aligned} \tag{2.18}$$

where we use the property that

$$A(\vec{b}, \lambda) \cdot A(\vec{b}, \lambda) = 1. \tag{2.19}$$

From equation (2.18), we can deduce that

$$|E(\vec{a}, \vec{b}) - E(\vec{a}, \vec{c})| \leq \int d\lambda \rho(\lambda)[1 - A(\vec{b}, \lambda)A(\vec{c}, \lambda)]. \tag{2.20}$$

Recognizing the second term in the sum on the right is $E(\vec{b}, \vec{c})$, we arrive at

$$1 + E(\vec{b}, \vec{c}) \geq |E(\vec{a}, \vec{b}) - E(\vec{a}, \vec{c})|. \tag{2.21}$$

The above is the Bell's inequality. In the original derivation, Bell used $P(\vec{a}, \vec{b})$ instead of $E(\vec{a}, \vec{b})$ to represent the expectation value but in this paper, we will use $E(\vec{a}, \vec{b})$ to differentiate it from the probability.

2.3.2 Violation of Bell's inequality

Now we proceed to show that there is a certain configuration of the settings on A and B which makes (2.21) invalid. In Bell's original derivation, he provided a counterexample which violates the above inequality. Here, we shall derive a relation between different settings which violates the Bell's inequality. First, assume that the expectation value given by the Bell's Theorem is the same as the quantum mechanical expectation value,

$$E(\hat{a}, \hat{b}) = \langle \vec{\sigma_1} \cdot \hat{a} \quad \vec{\sigma_2} \cdot \hat{b} \rangle = -\hat{a} \cdot \hat{b}. \tag{2.22}$$

Then using the Bell's inequality (2.21), we arrive at

$$1 - \hat{b} \cdot \hat{c} \geq |\hat{a} \cdot (\hat{b} - \hat{c})|. \tag{2.23}$$

Since

$$|\hat{a}| = |\hat{b}| = |\hat{c}| = 1, \tag{2.24}$$

we write

$$\hat{a} \cdot \hat{b} = \cos(\theta), \qquad \hat{a} \cdot \hat{c} = \cos(\phi), \qquad \hat{b} \cdot \hat{c} = \cos(\phi - \theta) \tag{2.25}$$

where we have assumed for simplicity that $\hat{a}$,$\hat{b}$ and $\hat{c}$ are coplanar. With this the inequality (2.23) is now

$$1 - \cos(\phi - \theta) \geq |\cos(\theta) - \cos(\phi)|. \tag{2.26}$$

In particular,

$$\cos(\theta) - \cos(\phi) \leq 1 - \cos(\phi - \theta). \tag{2.27}$$

Using the trigonometric identities,

$$-2\sin(\frac{\theta + \phi}{2})\sin(\frac{\theta - \phi}{2}) \leq 1 - (1 - 2\sin^2(\frac{\phi - \theta}{2})) \tag{2.28}$$

$$-2\sin(\frac{\theta + \phi}{2})\sin(\frac{\theta - \phi}{2}) \leq 2\sin^2(\frac{\phi - \theta}{2}) \tag{2.29}$$

$$\sin^2(\frac{\phi - \theta}{2}) + \sin(\frac{\theta + \phi}{2})\sin(\frac{\theta - \phi}{2}) \geq 0 \tag{2.30}$$

$$2\sin(\frac{\theta - \phi}{2})[\frac{\sin(\frac{\theta + \phi}{2}) + \sin(\frac{\theta - \phi}{2})}{2}] \geq 0 \tag{2.31}$$

finally we arrive at

$$\sin(\frac{\theta}{2})\cos(\frac{\phi}{2})\sin(\frac{\theta - \phi}{2}) \geq 0. \tag{2.32}$$

Let us interpret the above inequality and possible violations of it. For simplicity, let us agree upon a reasonable assumption that $0 \leq \phi, \theta \leq \pi$. We then have $1 \geq \sin(\frac{\theta}{2}) \geq 0$ and $1 \geq \cos(\frac{\phi}{2}) \geq 0$ and hence for the left hand side of the inequality to be negative, we need $sin(\frac{\theta-\phi}{2}) < 0$. Hence, it must be the case that $0 \leq \theta < \phi \leq \pi$ for any violation to occur. In fact, the above are exactly the cases put forward by Bell [4] and Clauser and Shimony [3] in their respective papers. In Bell's paper, $\hat{a}$ and $\hat{c}$ are perpendicular while $\hat{b}$ bisects the angle between them. Similarly, in Clauser et al. paper, $\hat{a}$ and $\hat{c}$ make an angle of $2\pi/3$ while $\hat{b}$ bisects it. Here we have arrived at a relation between θ and ϕ for the violation of Bell's inequality. Note that we restrict our analysis to the case where the vectors are coplanar. A generalized analysis I hope is possible.

2.4 Bell's inequality of 1971/CHSH inequality

The above Bell's inequality was used to theoretically predict the incompability between Quantum Mechanics and local hidden variable theory. To experimentally verify this violation, however, is difficult if we were to use the inequality (2.21) since it relies on a perfect correlation of spin between entangled particles i.e. the condition (2.16). In an actual experiment, such correlation could hardly happen. This difficulty leads to the generalization of Bell's inequality by Clauser, Horne, Shimony and Holt [10]. Bell later also derived a generalized inequality similar to that of CHSH [5]. We no longer assume a perfect correlation (2.16). Condition (2.13) remains the same. Preferring a more general derivation, we take into account the hidden variable contribution by the instruments used to measure the systems. Averaging over these instrument variables,

$$E(\vec{a}, \vec{b}) = \int d\lambda \, \rho(\lambda) \bar{A}(\vec{a}, \lambda) \bar{B}(\vec{b}, \lambda), \tag{2.33}$$

where $\bar{A}$ and $\bar{B}$ are averages over instrument variables and $\vec{a}$ and $\vec{b}$ represent setting chosen by observers. Instead of (2.12), we now have,

$$|\bar{A}| \leq 1, \quad |\bar{B}| \leq 1. \tag{2.34}$$

Also in the case where one or both of the measuring devices fails to measure the particles, we could simply assign value 0 to A and B for that run. Let $\vec{a'}$

and $\vec{b'}$ be alternative settings of instruments. Then

$$
\begin{aligned}
E(\vec{a},\vec{b}) - E(\vec{a},\vec{b'}) &= \int d\lambda\, \rho(\lambda)[\bar{A}(\vec{a},\lambda)\bar{B}(\vec{b},\lambda) - \bar{A}(\vec{a},\lambda)\bar{B}(\vec{b'},\lambda)] \\
&= \int d\lambda\, \rho(\lambda)[\bar{A}(\vec{a},\lambda)\bar{B}(\vec{b},\lambda)(1 \pm \bar{A}(\vec{a'},\lambda)\bar{B}(\vec{b'},\lambda))] - \\
&\quad - \int d\lambda\, \rho(\lambda)[\bar{A}(\vec{a},\lambda)\bar{B}(\vec{b'},\lambda)(1 \pm \bar{A}(\vec{a'},\lambda)\bar{B}(\vec{b},\lambda))].
\end{aligned}
\tag{2.35}
$$

Using two inequalities (2.34) above,

$$
\begin{aligned}
|E(\vec{a},\vec{b}) - E(\vec{a},\vec{b'})| \le & \int d\lambda\, \rho(\lambda)(1 \pm \bar{A}(\vec{a'},\lambda)\bar{B}(\vec{b'},\lambda)) \\
& + \int d\lambda\, \rho(\lambda)(1 \pm \bar{A}(\vec{a'},\lambda)\bar{B}(\vec{b},\lambda))
\end{aligned}
\tag{2.36}
$$

which leads to

$$
|E(\vec{a},\vec{b}) - E(\vec{a},\vec{b'})| \le 2 \pm (E(\vec{a'},\vec{b'}) + E(\vec{a'},\vec{b})), \tag{2.37}
$$

and finally to,

$$
|E(\vec{a},\vec{b}) - E(\vec{a},\vec{b'})| + |E(\vec{a'},\vec{b'}) + E(\vec{a'},\vec{b})| \le 2. \tag{2.38}
$$

By permuting the minus sign, we can obtain three more inequalities. All the inequalities obtained are collectively called CHSH inequalities. We shall briefly state the violation of CHSH inequality as the literature contains ample examples. The quantum mechanical maximum of the L.H.S of CHSH inequality can be calculated to be $2\sqrt{2}$ [11] in clear violation of the inequality.

2.5 Different assumptions in the derivation of Bell's inequalities

It is noteworthy that different assumptions may be used in derivation of the Bell's inequalities. For example, while the derivation of Bell's 1964 inequalities rests on the assumption of locality and determinism, Bell's 1976 inequalities are based on the assumption of local causality [7]. In order to avoid a possible confusion and recognize where our assumption in a subsequent proof lies, I will present a brief review on different assumptions behind the derivation of Bell's inequalities. To ensure a lucid interpretation of each assumption, we will utilize mathematical symbols in our translation. In the original derivations, Bell assigned $\rho(\lambda)$ to indicate that hidden variables λ are endowed with a probability distribution $\rho(\lambda)$. Hence it is acceptable to use $P(A, B|a, b, \lambda)$ to denote the probability of A and B having some values (usually ± 1) given the *settings* a and b take some values (usually $\{1, 2\}$) and given a λ from the set of all hidden variables, Λ. With this convention, determinism—which states that the outcomes of an experiment are predetermined by λ—could be translated to $P(A, B|a, b, \lambda)$ being either 1 or 0. This interpretation then can be reconciled with what is stated in (2.12). Next we discuss the assumption on locality. It states the outcome B is independent of what is being done at the site of outcome A. In our interpretation,

$$P(B|a, b, \lambda) = P(B|b, \lambda) \tag{2.39}$$

$$P(A|a, b, \lambda) = P(A|a, \lambda). \tag{2.40}$$

Another assumption we wish to address is that of local causality. It essentially states the statistical independence of outcome B from both the outcome A and the setting used to obtain outcome A, a. In other word,

$$P(B|A, a, b, \lambda) = P(B|b, \lambda). \tag{2.41}$$

Note that the same condition applies to A as well i.e. $P(A|B, a, b, \lambda) = P(A|a, \lambda)$. The assumption in our proof involves considering probabilities of the form $P(B_1, B_2|a)$ where B_1 and B_2 are outcomes at B's side. As we can see, these assumptions are inherently different from the previous ones.

3 The proof on the existence of complete joint probability distribution under no signalling assumption

Before I present the main proof, I would like to explain the *no signalling* condition that will be adopted in the proof. No signalling condition essentially states that the experimental setting of one party is *statistically independent* of the results at the other party. Although light is the fastest way of information transfer, we need not necessarily assume that signalling is limited to light transmission. The no signalling condition is fundamentally different from the local causality condition employed in CHSH inequality derivation. In the latter case, it is assumed that *every* result of observation is independent of the setting chosen on the other side. In contrast, it is assumed, in the current case, that the *statistics*—the probability or expectation value of obtaining certain sets of results—obtained on one side is independent of the setting chosen at other side. An instance of statiscal signalling is discussed in a paper by Valentini [12].

3.1 The proof

We consider a scenario where there are two observers A and B, spacelike separated, performing measurements on entangled particles emitted by a common source. We do not assume that all possible results of the experiment exist at *both* experimenters. Instead, it is sufficient to assume that such results exist only at one observer, say, at B (see Fig 2) side and it takes the form of *joint probability distribution* over all the results/outcomes. In other word, we assume that $P(\vec{b})$ exists where $\vec{b} \equiv (b_1, b_2, \ldots, b_n)$ and b_i with $i = 1, 2, \ldots, n$ is the result under the setting i. In contradistinction, no assumption is made about A's side. The setting there is denoted by x and the outcome in a particular experimental run by a_x. Altogether in a single experimental run, the LHV theory we consider specifies $P(a_x, b_1, b_2, \ldots, b_n)$. The results could take a set of m different values.

The second important assumption is that of no-signalling mentioned above. We assume that results at B and subsequently the probability distribution $P(\vec{b})$ is statistically independent of the settings at A. No signalling assumption, in our proof, takes the form,

$$P(\vec{b}|x) = P(\vec{b}) \tag{3.1}$$

where x is the setting on A's side i.e. $x = 1, 2, \ldots, n$. The last assumption which we describe the *freedom of choice* states that the actions of both observers, in this particular case the settings, are statistically independent of each other, and of the experimental probabilities. Now we proceed to the proof. By the last assumption,

$$P(a_x, \vec{b}) \equiv P(a, \vec{b}|x). \tag{3.2}$$

It follows that

$$\sum_{a_x} P(a_x, \vec{b}) = \sum_{a} P(a, \vec{b}|x) = P(\vec{b}|x) = P(\vec{b}). \tag{3.3}$$

The last equality is justified by our second assumption (3.1). We now have all the components necessary to show that the joint probability distribution for all the observers exists i.e. $P(\vec{a}, \vec{b})$ exists as a necessary consequence of

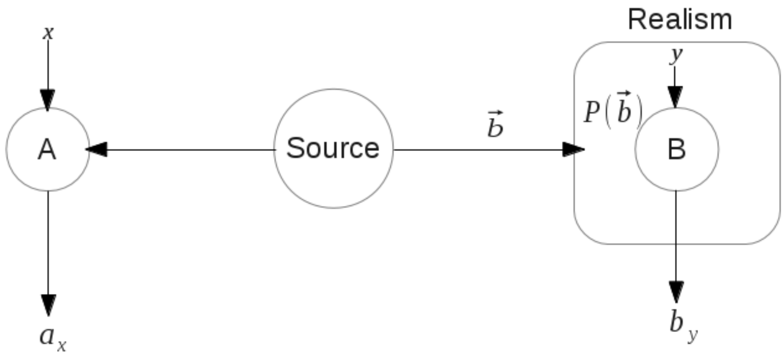


Figure 2: Diagram of the assumptions behind the derivation of a complete probability distribution. The source sends out predetermined results $\vec{b} \equiv b_1, b_2, \ldots, b_n$ towards place B. The choice of a particular $\vec{b}$ is based on the probability distribution $P(\vec{b})$. Observer at B chooses a setting y and obtains the result b_y. Similarly, observer at A chooses a setting x and obtains the result a_x. Note that no realism of the outcomes is assumed on A's side thus creating a semi-realistic system.

our assumptions. One way to construct the required probability distribution is to define

$$P(\vec{a}, \vec{b}) \equiv \frac{\prod_{x=1}^{n} P(a_x, \vec{b})}{P(\vec{b})^{n-1}}. \tag{3.4}$$

Then it is straightforward to show that this probability distribution returns appropriate experimental probabilities—the probabilities accessible by the experimenters at the end of the experiment. Clearly, this distribution is non-negative for all its constituents are non-negative. Next, we check whether it is possible to obtain experimental probabilities from this distribution. For example, if we want to obtain the experimental probability for a_k and b_l where $k, l \in \{1, 2, \ldots, n\}$, we sum (3.4) over all *other* outcomes (excluding

a_k and b_l) denoted by $\vec{a}_{\bar{k}}, \vec{b}_{\bar{l}}$. This gives

$$\begin{aligned}
\sum_{\vec{a}_{\bar{k}},\vec{b}_{\bar{l}}} \frac{\prod_{x=1}^{n} P(a_x,\vec{b})}{P(\vec{b})^{n-1}} &= \sum_{\vec{b}_{\bar{l}}} \frac{\prod_{x=1}^{n} \sum_{a_x, x\neq k} P(a_x,\vec{b})}{P(\vec{b})^{n-1}} \\
&= \sum_{\vec{b}_{\bar{l}}} P(a_k,\vec{b}) \\
&= \sum_{\vec{b}_{\bar{l}}} P(a,\vec{b}|k) \\
&= P(a,b_l|k) = P(a,b|k,l).
\end{aligned} \tag{3.5}$$

From the above, it is also straightforward to show that this distribution sums up to unity:

$$\begin{aligned}
\sum_{\vec{a},\vec{b}} \frac{\prod_{x=1}^{n} P(a_x,\vec{b})}{P(\vec{b})^{n-1}} &= \sum_{\vec{a}_k,\vec{b}_l} \sum_{\vec{a}_{\bar{k}},\vec{b}_{\bar{l}}} \frac{\prod_{x=1}^{n} P(a_x,\vec{b})}{P(\vec{b})^{n-1}} \\
&= \sum_{\vec{a}_k,\vec{b}_l} P(a,b|k,l) \\
&= \sum_{\vec{a},\vec{b}} P(a,b|k,l) = 1.
\end{aligned} \tag{3.6}$$

Hence, the distribution so obtained is indeed a normalized probability distribution which yields appropriate experimental probabilities as marginals.

Note that in this proof, we need only to assume that the results exist on one side and we could show the existence of probability distribution for all results. This is different from the usual assumptions in derivation of Bell's inequality where existence of results is assumed on both sides. One might ask the connection between this probability distribution constructed and the Bell/CHSH inequalities. The equivalent relation between existence of a complete probability distribution and validity of CHSH inequalities was proved in a paper by Arthur Fine [8]. A subset of the paper shows the equivalence realtion between the existences mentioned previously and that of a deterministic hidden-variable model. Moreover, it provides a way to derive the CHSH inequalities given the complete probability distribution. Let us now analyze the results from our proof. The first thing we may conclude is on the pervasiveness of realism. We start our proof with the assumption that out of two parties, realism, hence predefined results of the experiment, exists only on one side. But this eventually leads us to the deterministic hidden variable model where realism is assumed on both sides. In fact, this leads us to the determinism of the results. The second discussion might be on our assumption of joint probability distribution on all the results at one side i.e. $P(\vec{b})$. While the *uniqueness*—whether summation over different x in (3.3) would result in different $P(\vec{b})$—of such distribution is questioned [13], it is clear that as soon as we define (3.2), the uniqueness follows from (3.3).

3.2 CHSH inequality

In the previous section, we showed the existence of joint probability distribution assuming the no-signalling condition $P(\vec{b}|x) = P(\vec{b})$ and stated that it is related to the validity of CHSH inequality throught Arthur Fine's paper. We may ask ourselves whether we could derive CHSH inequality from the no signalling condition. And the answer is promising. First, we restrict ourselves to the analysis of two-outcomes case. Then we choose the no signalling condition from which we could derive CHSH inequality. The relevant no signalling condition turns out to be

$$P(b_1 = b_2|x = 1) - P(b_1 = b_2|x = 2) = 0. \tag{3.7}$$

Assume the other observer, A, has already made measurement so that the outcome is fixed and we can write $P(b_1 = b_2|x = i \in \{1, 2\}) = P(a_i b_1 = a_i b_2)$. All the outcomes are assumed dichotomic (± 1) hence their products are also dichotomic. We first define $v_i \equiv a_i b_1$ and $w_i \equiv a_i b_2$ and then define the *nested* correlation:

$$\begin{aligned} \varepsilon_i = \langle v_i \cdot w_i \rangle &= P(v_i = w_i) - P(v_i \neq w_i) \\ &= 2\,P(v_i = w_i) - 1. \end{aligned} \tag{3.8}$$

Hence,

$$P(v = w) = \frac{1}{2}(1 + \varepsilon_i). \tag{3.9}$$

Substituting (3.9) in (3.7),

$$\varepsilon_2 - \varepsilon_1 = 0. \tag{3.10}$$

Notice that $v, w \in \{+1, -1\}$ and hence we can devise the relation

$$v_i \cdot w_i = 1 - |v_i - w_i| = -1 + |v_i + w_i|. \tag{3.11}$$

The *nested* correlation thus becomes

$$\begin{aligned} \langle v_i \cdot w_i \rangle &= 1 - |\langle v_i - w_i \rangle| \\ &\leq 1 - |\langle v_i \rangle - \langle w_i \rangle|. \\ \langle v_i \cdot w_i \rangle &= -1 + |\langle v_i + w_i \rangle| \\ &\geq -1 + |\langle v_i \rangle + \langle w_i \rangle|. \end{aligned} \tag{3.12}$$

Concluding,

$$-1 + |\langle v_i \rangle + \langle w_i \rangle| \le \langle v_i \cdot w_i \rangle \le 1 - |\langle v_i \rangle - \langle w_i \rangle|. \tag{3.13}$$

Recognizing ε_i to be the same as $\langle v_i \cdot w_i \rangle$, correlation E_{i1} with $\langle v_i \rangle$ and E_{i2} with $\langle w_i \rangle$:

$$-1 + |E_{i1} + E_{i2}| \le \varepsilon_i \le 1 - |E_{i1} - E_{i2}|. \tag{3.14}$$

We then obtain the upper bounds,

$$\varepsilon_2 \le 1 - |E_{21} - E_{22}|, \tag{3.15}$$

$$-\varepsilon_1 \le 1 - |E_{11} + E_{12}|. \tag{3.16}$$

We proceed as:

$$2 - |E_{11} + E_{12}| - |E_{21} - E_{22}| \ge \varepsilon_2 - \varepsilon_1, \tag{3.17}$$

$$|E_{11} + E_{12}| - |E_{21} - E_{22}| \le 2. \qquad \text{by (3.10)} \tag{3.18}$$

The last inequality is the CHSH inequality. So, we could indeed obtain CHSH inequalities from assumption of a specific no-signalling condition.

3.3 Zohren-Gill inequality

The previous finding encourages us to derive more inequalities assuming different no signalling condition. We now consider the probability distribution of either $b_1 > b_2$ or $b_2 > b_1$. Again assuming bipartite model but with observables of d outcomes. No signalling condition then reads

$$P(b_2 < b_1 | x = 1) - P(b_2 < b_1 | x = 2) = 0. \tag{3.19}$$

We then find the upper bound of $P(b_2 < b_1|x = 1)$ using the inclusion relation (see Fig 3a) $\{b_2 \ge a_1\} \cap \{a_1 \ge b_1\} \subset \{b_2 \ge b_1\}$.

By identifying the complement of the inclusion relation, we obtain the probability relation:

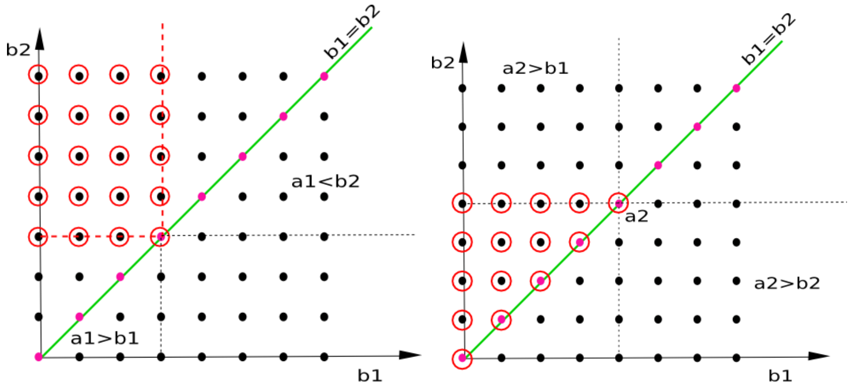
$$P(b_2 < a_1) + P(a_1 < b_1) > P(b_2 < b_1) \tag{3.20}$$

$$P(b_2 < a_1) + P(a_1 < b_1) > P(b_2 < b_1 | x = 1) \tag{3.21}$$

where we obtain (3.21) from dividing (3.20) by $P(x = 1)$ and then using the freedom of choice assumption $P(b_k < a_1) = P(b_k < a_1|x = 1)$. The lower bound for $P(b_2 < b_1|x = 2)$ is obtained in a similar manner but using the inclusion $\{a_2 \ge b_2\} \cap \{b_2 \ge b_1\} \subset \{a_2 \ge b_1\}$ instead (Fig 3b). The result is:

$$P(a_2 < b_2) + P(b_2 < b_1) > P(a_2 < b_1) \tag{3.22}$$

$$P(a_2 < b_2) + P(b_2 < b_1 | x = 2) > P(a_2 < b_1) \tag{3.23}$$



(a) $\{b_2 \geq a_1\} \cap \{a_1 \geq b_1\}$ (b) $\{a_2 \geq b_2\} \cap \{b_2 \geq b_1\}$

Figure 3: Inclusion relations between b_2, b_1, a_1 and b_2, b_1, a_2. Appropriate set relations between said variables are formed thereafter a probability measure is imposed on those relations. Resulting two inequalities are then combined using the no signalling condition to obtain the required inequality.

where we proceed as in the previous case. Lastly, using (3.19),

$$P(b_2 < a_1) + P(a_1 < b_1) + P(a_2 < b_2) - P(a_2 < b_1) > 0 \tag{3.24}$$

This inequality is known as Zohren-Gill inequality.

3.4 Relation between $P(\vec{b})$ and $P(f(\vec{b}))$

With the derivations of two previous inequalities based on different no signalling assumptions, it is intriguing to find out the relations between the probability distribution $P(\vec{b})$ and $P(f(\vec{b}))$ where $f(\vec{b})$ is some function on $\vec{b}$. The answer seems to be non-trivial. To simplify our analysis, we restrict ourselves to two-outcomes case. Then from the fact that $b_1, b_2 = \pm 1$, we could infer that there are only sixteen unique relations between b_1 and b_2.

In mathematical form, we wish to find whether,

$$P(b_1, b_2|x = 1) = P(b_1, b_2|x = 2) \iff P(f_k(b_1, b_2)|x = 1) = P(f_k(b_1, b_2)|x = 2), \tag{3.25}$$

where b_1 and b_2 are the outcomes and f_k with $k = 1, 2, ..., 16$ is a binary function. We show that the above equivalence indeed is the case. First, for the "if" part, we consult some results from information theory. It is known that, given the mutual information $I(A : B)$ between two variables A and B,

1. $I(A : B) \geq 0$
2. Data processing inequality: $I(f(A) : B) \leq I(A : B)$ and $I(A : f(B)) \leq I(A : B)$
3. $I(A : B) = 0 \iff P(A, B) = P(A)P(B)$.

Using the last result of the theory together with $P(A, B) = P(A|B)P(B)$, we deduce that

$$I(A : B) = 0 \iff P(A|B) = P(A). \tag{3.26}$$

Now using (3.26), we could transform the L.H.S of (3.25) into,

$$P(b_1, b_2|x) = P(b_1, b_2) \iff I(b_1, b_2 : x) = 0. \tag{3.27}$$

From the second result of the theory, $I(f_k(b_1, b_2) : x) \leq I(b_1, b_2 : x)$. This, together with the first result from the theory, $I(f_k(b_1, b_2) : x) \geq 0$, implies that $I(f_k(b_1, b_2) : x) = 0$. Hence we conclude that $P(f_k(b_1, b_2)|x) = P(f_k(b_1, b_2))$ and thus $P(f_k(b_1, b_2)|x = 1) = P(f_k(b_1, b_2)|x = 2)$.
Next, we prove the "only if" part. We start by assuming that $P(b_1, b_2|x = 1) \neq P(b_1, b_2|x = 2)$. This means there exist $\tilde{b_1}$ and $\tilde{b_2}$ such that $p_i \equiv P(\tilde{b_1}, \tilde{b_2}|x = 1) \neq P(\tilde{b_1}, \tilde{b_2}|x = 2) \equiv q_i$. Let us choose a function $f_k(b_1, b_2)$ that *marks* the inputs where $p_i \neq q_i$:

$$f_k(b_1, b_2) = \begin{cases} +1 & \text{if } b_1 = \tilde{b_1} \text{ and } b_2 = \tilde{b_2}, \\ -1 & \text{else.} \end{cases} \tag{3.28}$$

Then,

$$P(f_k(\tilde{b_1}, \tilde{b_2})|x = 1) = P(f_k = +1|x = 1) = p_i, \tag{3.29}$$

where the last equality follows from the fact that $f_k = +1$ implies $b_1 = \tilde{b_1}$ and $b_2 = \tilde{b_2}$. Likewise,

$$P(f_k(\tilde{b_1}, \tilde{b_2})|x = 2) = P(f_k = +1|x = 2) = q_i. \tag{3.30}$$

But we know that $p_i \neq q_i$ hence $P(f_k(b_1, b_2)|x = 1) \neq P(f_k(b_1, b_2)|x = 2)$. We have shown that

$$P(b_1, b_2|x = 1) \neq P(b_1, b_2|x = 2) \Rightarrow \exists f_k : P(f_k(b_1, b_2)|x = 1) \neq P(f_k(b_1, b_2)|x = 2). \tag{3.31}$$

whose contrapositive is the required "only if" part hence completing the proof. In words, no signalling condition holds for the probability distribution $P(b_1, b_2)$ if and only if the same condition holds for the probability of the functions of its arguments i.e. $P(f(b_1, b_2))$. Thus different f's could be linked to different Bell's inequalities.

3.5 General derivation of Bell's inequalities

The reader might notice that while we could derive different inequalities using different no signalling condtions, we employed different approaches in such derivations. It is, however, desirable to formulate a standardized method where all possible Bell's inequalities could be derived by a single general derivation supplied with different no signalling conditions. As in the case of the derivation of CHSH inequality, we simplify our analysis by assuming two settings, two outcomes scenario. Since b_1 and b_2 take only two values (± 1), the number of *binary* functions on those variables amounts to 16. By figuring out those functions, we may gain insight into the general derivation of Bell's inequalities. Referring to Table 1, function f_1 seems trivial. Function f_7, on the other hand, is interesting because its outputs are the same (-1) in the case where both the inputs are the same. Also when the inputs are different, the outputs are the same ($+1$). This, in a sense, is closely related to our no signalling assumption (3.7). The next interesting function is f_3 where we have the output $+1$ iff $b_1 > b_2$. Similarly f_5 represents $b_2 > b_1$. Out of the other functions, f_4 and f_6 turn out to be the same as b_1 and b_2 respectively. The only remaining function is f_2 whose form we suspect, when substituted in a similar no signalling condition as (3.7), will turn out to be identical to the latter. Nonetheless, investigations are ongoing.

b_1	b_2	f_1	f_2	f_3	f_4	f_5	f_6	f_7	f_8	...	f_{16}
-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	...	1
-1	1	-1	-1	-1	-1	1	1	1	1	...	1
1	-1	-1	-1	1	1	-1	-1	1	1	...	1
1	1	-1	1	-1	1	-1	1	-1	1	...	1

Table 1: Table of outcomes b_1, b_2 and the binary functions on the outcomes. There are 16 unique binary functions with f_9 onwards are simply negation of the previous functions.

4 Conclusion

We have provided a way to construct a complete probability distribution given the condtions of no-signalling, freedom of choice and realism on one side. Specific no signalling methods are found to be the conditions sufficient for derivation of a range of Bell's inequalities. When deriving said inequalities, however, we employ different techniques to achieve it. A much preferred general program for derivations of inequalities based on no signalling condition is still required and being investigated.

References

[1] Albert Einstein, Boris Podolsky, and Nathan Rosen. Can quantum-mechanical description of physical reality be considered complete? *Physical Review*, 47:777–780, 1935.

[2] Niels Bohr. Can quantum-mechanical description of physical reality be considered complete? *Physical Review*, 48:696–702, 1935.

[3] John F Clauser and Abner Shimony. Bell's theorem: experimental tests and implications. *Reports on Progress in Physics*, 41:1882–1927, 1978.

[4] John Stewart Bell. On the einstein podolsky rosen paradox. *Physics*, 1(3):195–200, 1964.

[5] M. Bell, K. Gottfried, and M. Veltman, editors. *John S. Bell on The Foundations of Quantum Mechanics.* World Scientific Publishing Co. Pte. Ltd., 2001.

[6] John P. Jarrett. On the physical significance of the locality conditions in the bell arguments. *Noûs*, 18:569–589, 1984.

[7] H. M. Wiseman. The two bell's theorem of john bell. *Journal of Physics A:Mathematical and Theoretical*, 47, 2014.

[8] Arthur Fine. Hidden variables, joint probabilitiy and the bell inequalities. *Physical Review Letters*, 48(5):291–295, 1982.

[9] D. Bohm and Y. Aharonov. Discussion of experimental proof for the paradox of einstein, rosen, and podolsky. *Physical Review*, 108(4):1070–1076, 1957.

[10] John F. Clauser, Michael A. Horne, Abner Shimony, and Richard A. Holt. Proposed experiment to test local hidden variable theories. *Physical Review Letters*, 23(15):880–884, 1969.

[11] Tomasz Paterek. *Quantum Communication: Non-Classical correlations and their applications.* PhD thesis, Uniwersytet Gdański, 2007.

[12] Antony Valentini. Signal-locality in hidden-variables theories. *Physics Letters A*, 297:273–278, 2002.

[13] C. Budroni and G. Morchio. Bell inequalities as constraints on unmeasurable correlations. *Foundations of Physics*, 42(4):544–554, 2011.