Eötvös-HPS 99-19

# Critical reflections on quantum probability theory*†

László E. Szabó‡
*Theoretical Physics Research Group of HAS*
*Department of History and Philosophy of Science*
*Eötvös University of Budapest*

**Abstract**

It is proved that quantum probabilities have no, in general, relative frequency interpretation. Consequently, quantum probability theory, as a probability theory, is inadequate for the description of quantum phenomena. Moreover, there is no reason to pursue a non-classical generalization of probability theory because all the phenomena of quantum physics can be well understood within the framework of the classical probability theory.

## 1 Introduction

The story of *quantum probability theory* and *quantum logic* begins with von Neumann's recognition[1], that quantum mechanics *can be regarded* as a kind of "probability theory", if the subspace lattice $L(H)$ of the system's Hilbert space $H$ plays the role of event algebra and the '$tr(WE)$'-s play the role of probability distributions over these events. This idea had been completed in the *Gleason theorem*[2]:

*Published in M. Rédei, M. Stoeltzner (eds.): *John von Neumann and the Foundations of Physics*, Kluwer, Dordrecht, 2001.
†Supported by OTKA Foundation (T 25841).
‡E-mail: leszabo@hps.elte.hu
[1]This idea appeared in J. von Neumann, *Mathematische Grundlagen der Quantenmechanik*, (Berlin: Springer, 1932). One can find it in a more explicit and somewhat different form in G. Birkhoff and J. von Neumann, The logic of quantum mechanics, *Ann. Math.* **37** (1936), 823-843.
[2]A. M. Gleason, Measures on the closed subspaces of a Hilbert space, *J. math. Phys* **6** (1957), 8855-8893.

**Definition 1** *A non negative real function $\mu$ on $L(H)$ is called (quantum) probability measure if $\mu(H) = 1$ and if whenever $E_1, E_2, \ldots$ are pairwise orthogonal subspaces, and $E = \bigvee_{i=1}^{\infty} E_i$, then $\mu(E) = \sum_{i=1}^{\infty} \mu(E_i)$.*

**Theorem 1 (Gleason 1957)** *If $H$ is a real or complex Hilbert space of dimension greater than 2, and $\mu$ is a probability measure on $L(H)$, then there exists a density operator $W$ on $H$, such that $(\forall E \in L(H))\, [\mu(E) = tr(WE)]$.*[3]

Formally, on the basis of mathematical analogy, the intersection and the (closed) linear union of subspaces are called 'conjunction' and 'disjunction' in the underlying 'event' lattice of *quantum probability theory*.

So far so good, but many think that we can go beyond this simple mathematical analogy, and regard quantum probability theory as a real probability theory replacing the classical one in its role in describing our world. As if the 'change of meaning' were such an easy matter[4], we are suggested to use the quantum logical connectives completely incompatible with the logical connectives of the metalanguage, at least when we are talking about the microphysical reality.

The quantum probability/quantum logic approach is based on the conviction that there are phenomena of quantum physics which cannot be accommodated in a world describable by the classical Kolmogorov theory of probability alone. The majority of experts share this conviction and, due to Feynmann[5], this opinion is also quite common among physicists whose field of interest is not foundations of physics.

There has been serious criticism of this approach, too. The first paper pointing out the contradictions which may appear if we assume that the event algebra is isomorphic with the subspace lattice of a Hilbert space was published by Strauss[6] a year after the famous Birkhoff and Neumann paper. It seems, however, that quantum probabilists and quantum logicians completely ignore the serious pitfalls pointed out by these authors. Bell expressed quite a similar disappointment[7]:

> *Why did such serious people take so seriously axioms which now seem so arbitrary? I suspect that they were misled by the pernicious misuse of the word 'measurement' in contemporary theory. This word very strongly suggests the ascertaining of some preexisting property of some thing, any instrument involved playing a purely passive role. Quantum experiments are just not like that, as we learned especially from*

[3]The subspaces, the corresponding projectors and the corresponding events are denoted by the same letter.

[4]Cf. Dummett, M., *The logical basis of metaphysics*, Duckworth, London 1995.

[5]R. Feynmann and A. Hibbs, *Quantum Mechanics and Path Integrals*, (New York: McGraw-Hill, 1965).

[6]M. Strauss, Mathematics as logical syntax — A method to formalize the language of a physical theory, *Erkenntnis*, **7** (1937), 147-153.

[7]J. S. Bell, *Speakable and unspeakable in quantum mechanics*, (Cambridge: Cambridge University Press, 1987), p. 166.

*Bohr. The results have to be regarded as the joint product of 'system' and 'apparatus,' the complete experimental set-up. But the misuse of the word 'measurement' makes it easy to forget this and then to expect that the 'results of measurements' should obey some simple logic in which the apparatus is not mentioned. The resulting difficulties soon show that any such logic is not ordinary logic. It is my impression that the whole vast subject of 'Quantum Logic' has arisen in this way from the misuse of a word. I am convinced that the word 'measurement' has now been so abused that the field would be significantly advanced by banning its use altogether, in favor for example of the word 'experiment.'*

My aim is to show in this paper that, beyond its counterintuitiveness, quantum probability theory is *inadequate* and *unnecessary*. It is inadequate because there cannot exist events in reality the *relative frequencies* of which would be equal with quantum probabilities. And it is unnecessary too, because there is no need in quantum mechanics to supersede the Kolmogorov theory of probability; we will see how quantum phenomena can, in general, be accommodated in the classical Kolmogorov theory of probability.

# 2 No frequency interpretation for quantum probabilities

## 2.1 Nonsensical probabilities for non-commuting elements

The following theorem illustrates that for non-commuting elements of $L(H)$ quantum probability theory predicts "probabilities" which are not interpretable as relative frequencies.

**Theorem 1** *Let $E_1$ and $E_2$ be two non-commuting elements of $L(H)$. There exists a pure state $\Psi$ for which the probabilities violate inequality*

$$p(E_1) + p(E_2) - p(E_1 \wedge E_2) \leq 1 \tag{1}$$

**Proof** Arbitrary $E_1$ and $E_2$ can be written in the following form:

$$\begin{aligned} E_1 &= (E_1 \wedge E_2) \vee A, \\ E_2 &= (E_1 \wedge E_2) \vee B, \end{aligned} \tag{2}$$

such that $A \perp E_1 \wedge E_2$ and $B \perp E_1 \wedge E_2$.

First we prove the following statements:

(a) $A \neq \emptyset$ and $B \neq \emptyset$ and $A \neq B$.

(b) $A \not\perp B$.

Indeed, if $A = \emptyset$ or $B = \emptyset$ or $A = B$ would hold then either $E_1 < E_2$ or $E_2 < E_1$, that would contradict to the assumed non-commutativity of $E_1$ and $E_2$. For proving (b) we show that from $A \perp B$ also the commutativity of $E_1$ and $E_2$ would follow. Commutativity is equivalent with $E_1 = (E_1 \wedge E_2) \vee (E_1 \wedge E_2^\perp)$. Using (2) we have

$$\begin{aligned} &(E_1 \wedge E_2) \vee \left[((E_1 \wedge E_2) \vee A) \wedge ((E_1 \wedge E_2) \vee B)^\perp\right] \\ &= (E_1 \wedge E_2) \vee \left[((E_1 \wedge E_2) \vee A) \wedge \left((E_1 \wedge E_2)^\perp \wedge B^\perp\right)\right] \\ &= (E_1 \wedge E_2) \vee \left[((E_1 \wedge E_2) \vee A) \wedge (E_1 \wedge E_2)^\perp\right] \wedge B^\perp \end{aligned} \tag{3}$$

Since $A \perp (E_1 \wedge E_2)$, the distributivity holds in the square brackets. Therefore we can continue (3) as follows:

$$\begin{aligned} &= (E_1 \wedge E_2) \vee (A \wedge B^\perp) \\ &= (E_1 \wedge E_2) \vee A = E_1 \end{aligned}$$

which proves (b).

Now, from (a) and (b) it follows that there exists at least one normalized vector $\Psi \in A$ such that $\Psi \notin B$. Such a $\Psi$ is a state vector for which the inequality

$$\underbrace{\langle \Psi, E_1 \Psi \rangle}_{1} + \underbrace{\langle \Psi, E_2 \Psi \rangle}_{>0} - \underbrace{\langle \Psi, (E_1 \wedge E_2 \Psi) \rangle}_{0} > 1 \tag{4}$$

holds, which is a violation of (1). □

The strange meaning of (4) is obvious! If $E_1$ happens with certainty, how can $E_2$ occur without $E_1$? It is remarkable that equation

$$p(E_1 \vee E_2) = p(E_1) + p(E_2) - p(E_1 \wedge E_2) \tag{5}$$

implies inequality (1). Von Neumann regarded (5) as a fundamental property of a probability measure, on which we should insist in the quantum case, too.[8]

[8]Cf. M. Rédei, Why John von Neumann did not like the Hilbert space formalism of quantum mechanics (and what he liked instead), *Studies in the History and Philosophy of Modern Physics*, **27** (1996) 493-510.

## 2.2 The Laboratory Record Argument

While I am completely convinced that Theorem 1 indicates a serious problem, a quantum probabilist might argue that this problem is a fictitious one because $E_1$ and $E_2$ do not commute and therefore they are not measurable simultaneously. Consequently, there is no experimental situation which would give rise to the nonsensical probabilities (4).

My second argument does not, however, appeal to conjunctions of non-commuting elements of $L(H)$. Neither does this argument appeal to a clear concept of 'event' in quantum mechanics. The only thing it appeals to is that quantum mechanics must be applicable to the everyday laboratory situations.

**Example** I don't know what "quantum event" is, the probability of which is a number like $tr(WE)$, but anyone who knows should be able to tell a laboratory assistant when does such an "event" occur. According to the instruction he makes a record like this:

| Run | $A_1$ | $A_2$ | $A_3$ | $A_4$ | $A_1A_3$ | $A_1A_4$ | $A_2A_3$ | $A_2A_4$ |
|---|---|---|---|---|---|---|---|---|
| 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 2 | 1 | 0 | 1 | 0 | 1 | 0 | 0 | 0 |
| 3 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 4 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 1 |
| 5 | 1 | 0 | 0 | 1 | 0 | 1 | 0 | 0 |
| 6 | 1 | 0 | 1 | 0 | 1 | 0 | 0 | 0 |
| 7 | 1 | 0 | 1 | 0 | 1 | 0 | 0 | 0 |
| 8 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |
| 99998 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 99999 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 100000 | 1 | 0 | 0 | 1 | 0 | 1 | 0 | 0 |
| $N = 100000$ | $N_1$ | $N_2$ | $N_3$ | $N_4$ | $N_{13}$ | $N_{14}$ | $N_{23}$ | $N_{24}$ |

where, assume, $(A_1, A_3)$, $(A_1, A_4)$, $(A_2, A_3)$ and $(A_2, A_4)$ are pairs of "quantum events" corresponding to commuting projectors. "0" stands for the case if an event does not happen and "1" if it does. The relative frequencies can be computed from this table:

$$\nu_1 = \frac{N_1}{N}, \; \nu_2 = \frac{N_2}{N}, \; \ldots \nu_{24} = \frac{N_{24}}{N}$$

Notice that each row of this table corresponds to one of the $2^4$ possible classical truth-value functions over propositions $A_1$ *happened*, $A_2$ *happened*, $A_3$ *happened*, and $A_4$ *happened*. In Pitowsky's language (See Appendix I.) we could say that each row corresponds to a vertex, $\mathbf{u}^\varepsilon$, of the corresponding classical polytope $\mathcal{C}(4, S)$, where $S = \{\{i, j\} \,|\, i = 1, 2;\, j = 3, 4\}$.

We can sum up our observation in the following stipulation:

**Stipulation** *If the components of a (finite) correlation vector* $\mathbf{p} = (\nu_i, \nu_{ij})$ *can be interpreted as relative frequencies of events, computed from a laboratory report, then*

$$(\forall i)\,(\forall j)\left[\begin{array}{c} \nu_i = \sum_{\varepsilon\in\{0,1\}^n} \lambda_\varepsilon u_i^\varepsilon \\ \nu_{ij} = \sum_{\varepsilon\in\{0,1\}^n} \lambda_\varepsilon u_{ij}^\varepsilon \end{array} \middle| \lambda_\varepsilon \geq 0;\ \sum_{\varepsilon\in\{0,1\}^n} \lambda_\varepsilon = 1\right] \tag{6}$$

$\lambda_\varepsilon = \frac{N_\varepsilon}{N}$, *where* $N_\varepsilon$ *is the number of type-*$\mathbf{u}^\varepsilon$ *rows in the record.*

In other words, relative frequencies are weighted averages of the classical truth-values. By virtue of Pitowsky theorem (see Appendix I.) (6) implies that relative frequencies must satisfy condition $\mathbf{p} \in \mathcal{C}(n, S)$ and, consequently, the corresponding Bell-type inequalities.

Now, consider the well known EPR-Aspect experiment (Appendix II.). Just as in the above example, we have four events and four conjunctions. Each of the conjunctions belongs to *commuting* projectors! In case of a particular choice of directions along which the spin components are measured, the quantum probabilities form the following correlation vector:

$$\mathbf{p} = (q\,(A_1)\,,\ q\,(A_2)\,, q\,(A_3)\ \ldots\ q\,(A_2 A_4)) = \left(\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{3}{8}, \frac{3}{8}, 0, \frac{3}{8}\right)$$

But these numbers do not satisfy the Clauser-Horne inequalities! That is, *quantum probabilities cannot be, in general, interpreted as relative frequencies of events.*

In order to clarify the importance of this result, I need to make a few remarks:

1. What we have actually recognized here is that "quantum probability" must not be interpreted as probability, if we insist on the frequentists' understanding of the term. The quantum mechanical $tr\,(WE)$ is not the (absolute) probability of a real event, but it is a conditional probability $p(A\,|a)$, which means the probability of the outcome-event $A$, given that the measurement-preparation $a$ has happened. $tr\,(WE)$ no doubt does have such a meaning, I believe, in accordance with the everyday laboratory practice. The controversial question is, whether it means something more: whether there exists an event $\widetilde{E}$ in reality, such that $tr\,(WE) = p\left(\widetilde{E}\right)$ holds. As we could see, if probability means relative frequency, then the probability function on an event algebra can be nothing else but the weighted average of the classical truth-value functions, therefore, according to Pitowsky's theorem, it must be Kolmogorovian, consequently, it satisfies the Bell inequalities. Since the $tr\,(WE)$-type quantities violate Bell inequalities, they are not interpretable as relative frequencies. In other

words, *there are no events, in general, that would happen with probabilities like* $tr\,(WE)$.

2. And accordingly, as I pointed out in an earlier paper[9], such a question, for example, whether there exists a common cause explanation for the EPR correlations, in its original form, is meaningless, because it is about the existence of common cause for correlations among *non-existing* events. *What the violation of the Bell inequalities indicates is not that the EPR correlations do not have common cause, but rather that there are no events which would have such correlations.*

3. The same holds for the question whether the classical Kolmogorov theory of probability can or cannot describe the probabilities of events observed in the quantum world. If we insist on the frequency interpretation of probability, the answer is clear: yes, it can, because relative frequencies *are* Kolmogorovian. If "quantum probabilities" do not satisfy the Kolmogorov axioms, then they are not interpretable as relative frequencies, or, if they are thought as relative frequencies, then *there are no events in reality which would happen with these frequencies.*

4. The fact that "non-Kolmogorovian probabilities" are not interpretable as relative frequencies explain why von Neumann's program to create a "non-commuting version of probability theory" necessarily failed[10]. In his new mathematical model he wanted to *reproduce* the same *values* of "quantum probabilities"!

   $$\text{something new} \stackrel{1}{=} tr(WE) \stackrel{2}{=} \text{relative frequency of event}$$

   Solving the problem of equantion 1 does not resolve the contradiction at equation 2.

5. Some people claim that "a quantum probabilist is, of course, not a frequentist". Then what is he? What is a tenable interpretation of probability, different from the frequency interpretation, applicable to quantum mechanics? In addition to the problem, that even the subjective probabilities must, in final analysis, satisfy the Kolmogorov axioms, we have no freedom in interpreting quantum mechanical probabilities, at all. It is because the laboratory practice singles out the frequency interpretation. If quantum probabilist's quantum mechanics is identical with the experimental physicist's one, then there is no other meaning of probability than

[9] E. Szabó, L., On an attempt to resolve the EPR-Bell paradox via Reichenbachian concept of common cause, forthcoming in *Int. J. of Theor. Phys.*

[10] Cf., Rédei M., Von Neumann's concept of quantum logic and quantum probability, in this volume.

relative frequency. Only in this sense we can say that quantum mechanics is an experimentally confirmed physical theory.

# 3 Do we really need quantum probability theory?

## 3.1 The double slit experiment

The conclusion we can draw from the previous section is that $L(H)$ can hardly play the role of an "algebra of events" for a probability theory and the numbers $tr(WE)$ cannot be interpreted as relative frequencies of events. Sometimes it is claimed that in spite of the above "difficulties" — which is of course an understatement — we must figure out something, we must solve these interpretational problems of "quantum probability theory" because there are phenomena in quantum physics which are not describable with the classical theory of probability. In this section I want to show that this is not true, we don't need quantum probability theory.

Our next example is the double slit experiment which is often quoted in order to justify why we need quantum probability theory (Fig. 1).

Denote $p(A)$ the probability of that "the particle arrives at a given point of the screen, Q, when only slit 1 is open". $p(B)$ denotes the similar probability for slit 2. In the experiment one finds that

$$p(A) + p(B) \neq p(A \vee B) \tag{7}$$

where $p(A \vee B)$ stands for the probability of "the particle arrives at point Q, when both slits are open'. According to the usual interpretation the double slit experiment shows that "*the method of computing probabilities involving subatomic particles is different from that of classical probability theory*"[11]. Therefore we must, as the usual conclusion says, 1) change probabilities for complex amplitudes (Feynmann) or 2) give up the Boolean event lattice and classical probability theory (quantum probability theory). There is, however, a bad mistake in this standard claim, namely a misinterpretation of the superposition principle. While it is true, approximately, that the solution of the linear Schrödinger equation with boundary condition 3 (both slits are open) is a superposition of the solutions belonging to boundary conditions 1 and 2 (see Fig. 1), that is, $\Psi_3 = \Psi_1 + \Psi_2$, it doesn't mean, on the other hand, that the relation of the three situations could be described as a logical disjunction.

Let me show at least two different ways in which we can, contrary to the standard view, correctly describe the double slit experiment within the framework of classical probability theory. In both cases, the precise usage of notions "event"

[11]Gudder, S., *Op.cit.*, p. 57.

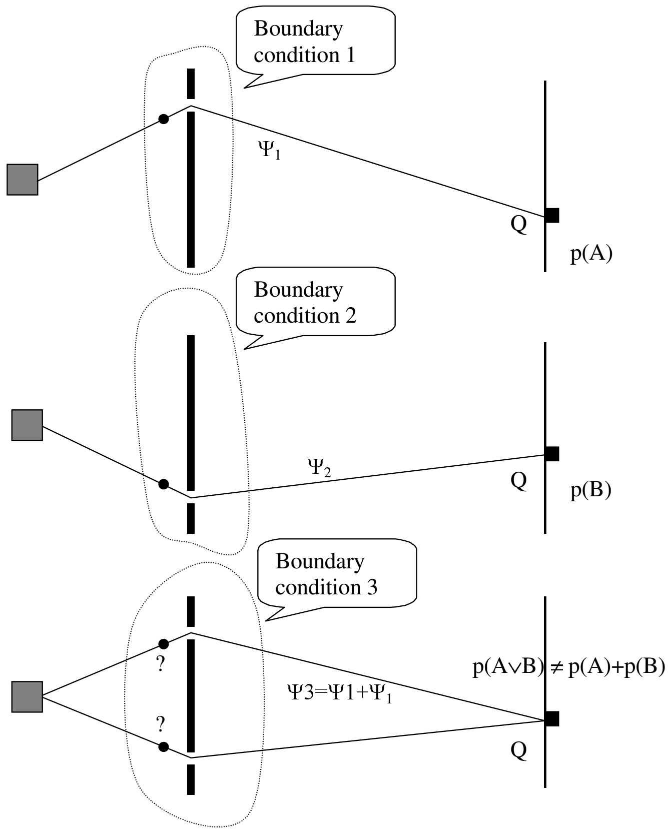


Figure 1: *According to the usual interpretation the double slit experiment indicates that the rules of computing probabilities in quantum mechanics must be different from that of classical probability theory*

and "disjunction" is what makes the classical probability theory satisfactory, while the formula (7) is, as we will see soon, based on the misuse of these notions.

**Version I**

We must precisely distinguish the following events:

$A$: "Slit 1 is open and slit 2 is closed and the particle is detected at $Q$"

$B$: "Slit 1 is closed and slit 2 is open and the particle is detected at $Q$"

$C$: "Slit 1 is open and slit 2 is open and the particle is detected at $Q$"

Obviously,

$$A \vee B \neq C$$

Consequently, we are not surprised that

$$p(A \vee B) = p(A) + p(B) \neq p(C)$$

That is, formula (7) is incorrect, consequently there is no violation of classical rules of probability calculation.

**Version II**

There is only one event:

$D$: “The particle is detected at $Q$”

There are, however, different *conditions* under which the probabilities are understood. But the Kolmogorov axioms are meant to apply to probabilities belonging to *one common* system of conditions! Consequently, it does not mean a violation of the Kolmogorov axioms if

$$p_{1\text{ is open; }2\text{ is closed}}(D) + p_{1\text{ is closed; }2\text{ is open}}(D) \neq p_{1\text{ is open; }2\text{ is open}}(D) \quad (8)$$

It is to be mentioned here that the conditions written in the indexes of probabilities in (8) are sometimes meant as conditioning events:

$\alpha$ : ”1 is open; 2 is closed”
$\beta$ : ”1 is closed; 2 is open”
$\gamma$ : “1 is open; 2 is open”

If, falsely, $\gamma$ is taken as the disjunction of $\alpha$ and $\beta$

$$\gamma = \alpha \vee \beta \quad (9)$$

then the conditional probabilities should satisfy the following inequality:

$$min\,(p(D\,|\alpha), p(D\,|\beta)) \leq p(D\,|\alpha \vee \beta) \leq max\,(p(D\,|\alpha), p(D\,|\beta)) \quad (10)$$

In the experiment, the interference pattern shows many violations of this condition. And again, we arrive at an “argument” for rejecting the probability theory and/or logic involved in the deduction of (10). The other reaction is, as van Fraassen rightly remarks[12], to reject (9). However I disagree with his explanation: “*To put it in other words: we reject the idea that the electron must have a definite position (in slit 1 or in slit 2) at time of its passing the barrier.*” Nothing forces us to jump to such a conclusion. (9) fails simply because sentence $\gamma$ is not the disjunction of $\alpha$ and $\beta$.

[12] B. Van Fraassen, *Quantum Mechanics – An Empiricist View*, (Oxford: Clarendon Press, 1991), p. 111.

## 3.2 The EPR experiment

We have seen that the double slit experiment does not prove the nonapplicability of Kolmogorov's classical theory of probability. It is true, however, that this example is not regarded as a serious one: it is rather used in quantum mechanics text books only. But now we are going to analyze the Einstein-Podolsky-Rosen experiment which is regarded as a crucial, empirically tested situation providing probabilities which do not conform with the Kolmogorovian theory. You can find the description of the experiment in Appendix II.

The question we would like to answer is whether the probabilities (17), measured in the Aspect experiment, can be accommodated in a Kolmogorovian probability model, or not. Let me first recall the standard argumentation which yields to a negative answer. Consider

$$p_1 = tr(WA_1),\ p_2 = tr(WA_2),\ p_3 = tr(WA_3),\ p_4 = tr(WA_4)$$

$$p_{13} = tr(WA_1A_3),\ p_{14} = tr(WA_1A_4),\ p_{23} = tr(WA_2A_3),\ p_{24} = tr(WA_2A_4)$$

the values of which are given in (18). Substituting these values into the last Clauser-Horne inequality of (16) we find that

$$\mathbf{p} = \left(\frac{1}{2},\frac{1}{2},\frac{1}{2},\frac{1}{2},\frac{3}{8},\frac{3}{8},0,\frac{3}{8}\right) \notin \mathcal{C}(n,S).$$

According to the Pitowsky theorem, as the standard argument goes on, the probabilities observed in the Aspect experiment have no Kolmogorovian representation.

We must recognize, however, that this is just the same problem we discussed in section 2.2. What the violation of the Clauser-Horne inequalities indicates is not that probabilities (17) cannot be represented in a Kolmogorov probability space, but *there are no events in reality which would happen with these frequencies.* The probabilities of the real physical events observed in the experiment, $p(A_1), p(A_2), p(A_3), p(A_4), p(a_1), p(a_2), p(a_3), p(a_4)$, do not violate Kolmogorovity, as it must be the case, since they are relative frequencies counting down from a laboratory record of a real experiment:

$$\begin{aligned}
\mathbf{p} &= (p(A_1), p(A_2), p(A_3), p(A_4), p(a_1), p(a_2), p(a_3), p(a_4), \\
&\quad p(A_1 \wedge A_2), \ldots p(A_4 \wedge a_4), \ldots p(a_3 \wedge a_4)) \\
&= \left(\frac{1}{4},\frac{1}{4},\frac{1}{4},\frac{1}{4},\frac{1}{2},\frac{1}{2},\frac{1}{2},\frac{1}{2},0,\frac{3}{32},\frac{3}{32},\frac{1}{4},0,\frac{1}{8},\frac{1}{8},0,\frac{3}{32},0,\frac{1}{4},\frac{1}{8},\right. \qquad (11)\\
&\quad \left.\frac{1}{8},0,\frac{1}{8},\frac{1}{8},\frac{1}{4},0,\frac{1}{8},\frac{1}{8},0,\frac{1}{4},0,\frac{1}{4},\frac{1}{4},\frac{1}{4},\frac{1}{4},0\right) \\
&\in \mathcal{C}(8, S_{\max})
\end{aligned}$$

Of course, there are no derived Bell-type inequalities for 36-dimensional correlation vectors, therefore condition (11) is tested numerically, by computer[13]. However, it is quite plausible if you take into attention that the correlation vector made of the measured relative frequencies of the outcome events only,

$$\begin{aligned}\mathbf{p} &= (p(A_1), p(A_2), p(A_3), p(A_4), p(A_1 \wedge A_3), p(A_1 \wedge A_4), p(A_2 \wedge A_3), p(A_2 \wedge A_4)) \\ &= \left(\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{3}{32}, \frac{3}{32}, 0, \frac{3}{32}\right)\end{aligned}$$

satisfies the Clauser-Horne inequalities.

What we can observe here is nothing else but what I formulated in my "Kolmogorovian Censorship" hypothesis[14]. We never encounter "naked" quantum probabilities in reality. A correlation vector consisting of empirically observed probabilities is always of the following form:

$$\mathbf{p} = (q_1\widetilde{p}_1, q_2\widetilde{p}_2, \ldots \widetilde{p}_1, \widetilde{p}_2, \ldots q_{ij}\widetilde{p}_{ij}, \ldots \widetilde{p}_{kl}, \ldots)$$

where $(q_1 \ldots q_n \ldots q_{ij} \ldots)$ are quantum probabilities and $(\widetilde{p}_1 \ldots \widetilde{p}_n \ldots \widetilde{p}_{ij} \ldots)$ are classical probabilities with which the corresponding measurement preparations occur. The hypothesis says that such a $\mathbf{p}$ is always classical. (In general, a product of quantum and classical probabilities does not necessarily form a classical correlation vector.) Bana and Durt proved such a theorem for finite number of measurements[15]. In the next section I shall give a more simple proof which is also valid for the infinite (but countable) case .

# 4 The meaning of quantum probability

We have seen that quantum probability is not probability, because it cannot be, in general, the relative frequency of an event. What is then the correct interpretation of the $tr(WE)$-type quantitites? As I have mentioned already, it is a *conditional* probability, $tr(WE) = p(E \,|\, e)$, that is, the probabilitiy of the meausurement oucome $E$, given that the measurement preparation $e$ happened. There is nothing new in this interpretation; whenever a statistical prediction of quantum mechanics is experimentally tested, $tr(WE)$ is identified with $p(E \,|\, e)$.

In order to make this picture complete, we want to see a big Kolmogorovian probability space where all these conditional probabilities are, jointly, represented. First, let me give an example, how can we solve a similar problem in the classical theory of probability.

[13] E. Szabó, E., Is quantum mechanics compatible with a deterministic universe? Two interpretations of quantum probabilities, *Foundations of Physics Letters*, **8** (1995), 421-440

[14] E. Szabó, L., *Op. cit.*

[15] Bana, G. and Durt, T., Proof of Kolmogorovian Censorship, *Foundations of Physics*, **27**, 1355-1373. (1997)

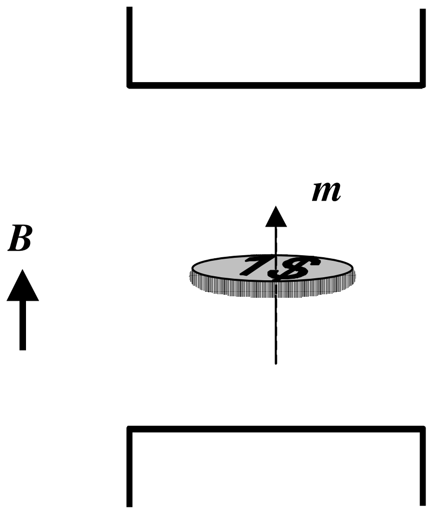


Figure 2: *We are tossing a coin which has a little magnetic moment. The probabilities of Heads (H) and Tails (T) are modified if the magnetic field is on*

We are tossing a coin which has a little magnetic momentum (Fig. 2). If the magnetic field is off, the probabilities are

$$\begin{aligned} p_{\text{off}}(\text{H}) &= 0.5 \\ p_{\text{off}}(\text{T}) &= 0.5 \end{aligned}$$

If the magnetic field is on, the probabilities are different:

$$\begin{aligned} p_{\text{on}}(\text{H}) &= 0.2 \\ p_{\text{on}}(\text{T}) &= 0.8 \end{aligned}$$

The event algebra $\mathcal{A}$ is shown in Figure 3. For the two different physical conditions we have two separate probability models: $(\mathcal{A}, p_{\text{off}})$ and $(\mathcal{A}, p_{\text{on}})$,which are, separately, Kolmogorovian. For example, they satisfy the simplest Bell-type inequality (1):

$$p_{\text{off}}(\text{H}) + p_{\text{off}}(\text{T}) - p_{\text{off}}(\text{H} \wedge \text{T}) \leq 1$$

and separately,

$$p_{\text{on}}(H) + p_{\text{on}}(T) - p_{\text{on}}(H \wedge T) \leq 1$$

Now, if we "forgot" that these probabilities belong to different physical conditions, and put them together into one formula prescribed for a Kolmogorovian

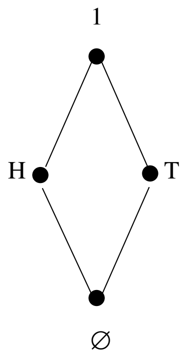

Figure 3: *The original event algebra $\mathcal{A}$*

probability theory, then we would find the "violation of the rules of classical probability theory":

$$p_{\text{off}}(\text{H}) + p_{\text{on}}(\text{T})) - p_{\text{off}}(\text{H} \wedge \text{T}) = 0.5 + 0.8 > 1$$

or

$$p_{\text{on}}(\text{H}) + p_{\text{off}}(\text{T}) = 0.2 + 0.5 \neq 1 = p_{\text{off}}(1) = p_{\text{off}}(\text{H} \vee \text{T})$$

If we wish to see these probabilities together in one probability model, then it is necessary 1) to extend the original event algebra (Fig. 4) in order to include new events corresponding to the different conditions, and, of course, 2) we must be able to tell the probabilities of the conditioning events. In the example of question assume that $p(\text{OFF}) = 0.5$ and $p(\text{ON}) = 0.5$. So, the unified probability model is $(\mathcal{A}', p)$, where

$$\begin{aligned} p(1) = p(2) = 1 - p(9) = 1 - p(10) &= 0.25 \\ p(3) = 1 - p(8) &= 0.1 \\ p(4) = 1 - p(7) &= 0.4 \\ p(\text{OFF}) = p(\text{ON}) &= 0.5 \\ p(\text{H}) = p(6) &= 0.35 \\ p(\text{T}) = p(5) &= 0.65 \end{aligned} \tag{12}$$

The original probabilities are represented as *conditional* probabilities (defined by the Bayes law):

$$p_{\text{on}}(\text{H}) = \frac{p(\text{H} \wedge \text{ON})}{p(\text{ON})} = \frac{p(3)}{p(\text{ON})} = \frac{0.1}{0.5} = 0.2$$

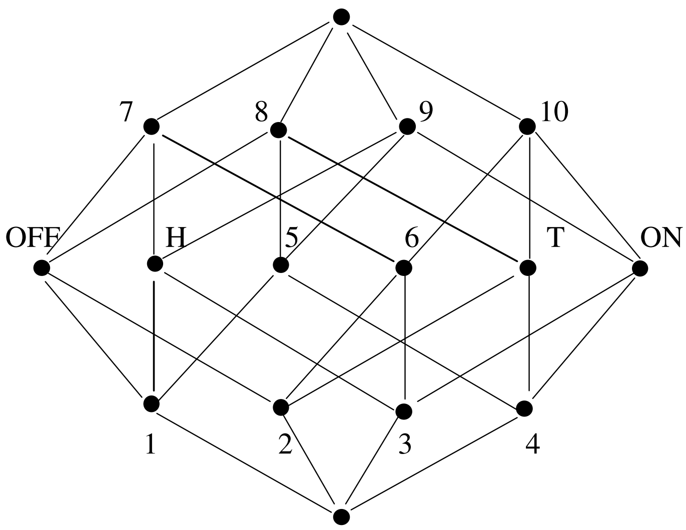


Figure 4: *The extended event algebra* $\mathcal{A}'$

$$\begin{aligned} p_{\rm on}(\mathrm{T}) &= \frac{p(\mathrm{T}\wedge\mathrm{ON})}{p(\mathrm{ON})} = \frac{p(4)}{p(\mathrm{ON})} = \frac{0.4}{0.5} = 0.8 \\ p_{\rm off}(\mathrm{H}) &= \frac{p(\mathrm{H}\wedge\mathrm{OFF})}{p(\mathrm{OFF})} = \frac{p(1)}{p(\mathrm{OFF})} = \frac{0.25}{0.5} = 0.5 \\ p_{\rm off}(\mathrm{T}) &= \frac{p(\mathrm{T}\wedge\mathrm{OFF})}{p(\mathrm{OFF})} = \frac{p(2)}{p(\mathrm{OFF})} = \frac{0.25}{0.5} = 0.5 \end{aligned} \tag{13}$$

To come back to the quantum case, the method can be the same as in the above classical example. For sake of simplicity, assume we consider a countable set of physically different measurement setups, denoted by $\mathcal{M}$. A Boolean $\sigma$-algebra of outcomes $\mathcal{A}_m$ belongs to each $m \in \mathcal{M}$. A good picture to have in mind is that each $\mathcal{A}_m$ is isomorphic with one of the maximal Boolean sublattices of $L(H)$. It is a fact of quantum mechanics that quantum probability $q_W$, belonging to a given state operator $W$, forms a Kolmogorovian probability measure on $\mathcal{A}_m$. Let us denote it by $p_m$.

Again, if we wish to create a joint Kolmogorovian representation for all these $p_m$-s, we must embed algebras $\mathcal{A}_m$ into one larger algebra $\mathcal{A}$ and specify the probability distribution $\varrho$ over the different measurement setups. To specify such a probability distribution is, of course, a delicate question, but not less than to specify how frequently is the magnetic field on and off in the above example. The only thing we assume about this probability distribution is that $(\mathcal{M},\varrho)$ is a discrete Kolmogorov probability space, i.e., $\varrho : \mathcal{M} \to [0,1]$ and $\sum_{m\in\mathcal{M}} \varrho(m) = 1$.

Let the extended algebra $\mathcal{A}$ consist of the sections of bundle $\bigcup_{m\in\mathcal{M}}\mathcal{A}_m$ :

$$\mathcal{A}=\left\{\alpha \,\middle|\, \alpha:\mathcal{M}\to\bigcup_{m\in\mathcal{M}}\mathcal{A}_m \text{ and } (\forall m)\ [\alpha(m)\in\mathcal{A}_m]\right\}$$

$\mathcal{A}$ is a Boolean $\sigma$-algebra with the following operations:

$$\begin{aligned}(\alpha\wedge\beta)(m) &= \alpha(m)\wedge\beta(m)\\ (\alpha\vee\beta)(m) &= \alpha(m)\vee\beta(m)\\ \left(\alpha^{\perp}\right)(m) &= \alpha(m)^{\perp}\end{aligned}$$

The minimal and maximal elements of $\mathcal{A}$ are

$$\begin{aligned}\emptyset,\mathbb{I} &: \mathcal{M}\to\bigcup_{m\in\mathcal{M}}\mathcal{A}_m\\ \emptyset(m) &= \emptyset_m\\ \mathbb{I}(m) &= \mathbb{I}_m\end{aligned}$$

It is easy to check that the following map defines a probability measure on $\mathcal{A}$:

$$\begin{aligned}p &: \mathcal{A}\to[0,1]\\ p(\alpha) &= \sum_{m\in\mathcal{M}}p_m(\alpha(m))\,\varrho(m)\end{aligned}$$

Elements $A_m\in\mathcal{A}_m$ and $m\in\mathcal{M}$ can be represented as follows:

$$\begin{aligned}A_m &\leftrightarrow \gamma_{A_m}\in\mathcal{A}\\ \gamma_{A_m}(i) &= \begin{cases}\emptyset_i\in\mathcal{A}_i & \text{if } i\neq m\\ A_m & \text{if } i=m\end{cases}\end{aligned}$$

and

$$\begin{aligned}m &\leftrightarrow \gamma_m\in\mathcal{A}\\ \gamma_m(i) &= \begin{cases}\emptyset_i\in\mathcal{A}_i & \text{if } i\neq m\\ \mathbb{I}_m\ \in\mathcal{A}_m & \text{if } i=m\end{cases}\end{aligned}$$

Moreover,

$$p(\gamma_m) = \sum_{i\in\mathcal{M}}p_i(\gamma_m(i))\,\varrho(i)=\varrho(m) \quad (14)$$

$$p(\gamma_{A_m}) = \sum_{i\in\mathcal{M}}p_i(\gamma_{A_m}(i))\,\varrho(i)=p_m(A_m)\,\varrho(m)=tr(WA_m)\,\varrho(m) \quad (15)$$

Therefore, "quantum probability" $tr(WA_m)$ obtains its interpretation in conditional probability of getting outcome $A_m$ given that measurement $m$ is performed:

$$p(\gamma_{A_m}|\,\gamma_m)=\frac{p(\gamma_{A_m}\wedge\gamma_m)}{p(\gamma_m)}=\frac{p(\gamma_{A_m})}{p(\gamma_m)}=tr(WA_m)$$

The existence of representations (14) and (15) proves the Kolmogorovian Censorship Hypothesis for countable number of measurements.

**Remarks**

- Since it was supposed in the above construction that $\mathcal{M}$ is countable, projectors $\{A_m\}_{m\in\mathcal{M}}$ do not cover the whole continuous $L(H)$. To cover the whole lattice with a continuous $\mathcal{M}$ needs further technical elaborations, which I want to publish somewhere else. It is true, however, that any countable subset of $L(H)$ can be covered in this way.

- The above representation of quantum probabilities is consistent in the sense that if a projector $P \in L(H)$ appears in more that one maximal Boolean sublattices in question, for example, $A_m \in \mathcal{A}_m$ and $A'_{m'} \in \mathcal{A}_{m'}$ correspond to the same $P$, then

$$p\left(\gamma_{A_m} | \gamma_m\right) = p\left(\gamma_{A'_{m'}} | \gamma_{m'}\right) = tr(WP)$$

  It is also consistent in the sense that the corresponding conditional probabilities reproduce quantum probabilities independently of how $(\mathcal{M}, \varrho)$ is chosen.

# 5 Appendix I.

Pitowsky elaborated a convenient geometric language for the discussion of the problem whether empirically given probabilities are Kolmogorovian or not[16].

Let $S$ be a set of pairs of integers $S \subseteq \{\{i,j\} \mid 1 \leq i < j \leq n\}$. Denote by $R(n,S)$ the linear space of real vectors having a form like $(f_1, f_2, ...f_{ij}, ...)$. For each $\varepsilon \in \{0,1\}^n$, let $u^\varepsilon$ be the following vector in $R(n,S)$:

**Definition 2** *The* classical correlation polytope $\mathcal{C}(n,S)$ *is the closed convex hull in* $R(n,S)$ *of vectors* $\{u^\varepsilon\}_{\varepsilon\in\{0,1\}^n}$ *:*

$$\mathcal{C}(n,S) := \left\{ a \left| a \in R(n,S) \text{ and } a = \sum_{\varepsilon\in\{0,1\}^n} \lambda_\varepsilon u^\varepsilon, \text{ where } \lambda_\varepsilon \geq 0 \text{ and } \sum_{\varepsilon\in\{0,1\}^n} \lambda_\varepsilon = 1 \right. \right\}$$

Consider now events $A_1, A_2, \ldots A_n$ and some of their conjunctions $A_i \wedge A_j$ : $(\{i,j\} \in S)$. Assume that we know their probabilities from which we can form a so called *correlation vector*:

$$\begin{aligned} \mathbf{p} &= (p_1, p_2, \ldots p_n, \ldots p_{ij}, \ldots) \\ &= (p(A_1), p(A_2), \ldots, p(A_n), \ldots p(A_i \wedge A_j), \ldots) \in R(n,S) \end{aligned}$$

[16] *Op. cit.*

**Definition 3** *We will then say that* **p** has a Kolmogorovian representation *if there exist a Kolmogorovian probability space* $(\Omega, \Sigma, \mu)$ *and measurable subsets*

$$X_{A_1}, X_{A_2}, \ldots X_{A_n} \in \Sigma$$

*such that*

*Pitowsky's theorem tells us the necessary and sufficient condition a correlation vector must satisfy in order to be Kolmogorovian.*

**Theorem 2 (Pitowsky, 1989)** *A correlation vector*

$$\mathbf{p} = (p_1, p_2, \ldots p_n, \ldots p_{ij}, \ldots)$$

*has a Kolmogorovian representation if and only if* $\mathbf{p} \in \mathcal{C}(n, S)$.

In case $n = 4$ and $S = S_4 = \{1,3\}, \{1,4\}, \{2,3\}, \{2,4\}\}$ the condition $\mathbf{p} \in \mathcal{C}(n, S)$ is equivalent with the following inequalities:

$$\begin{aligned}
&0 \le p_{ij} \le p_i \le 1, \\
&0 \le p_{ij} \le p_j \le 1, \qquad\qquad i = 1,2 \;\; j = 3,4 \\
&p_i + p_j - p_{ij} \le 1, \\
&-1 \le p_{13} + p_{14} + p_{24} - p_{23} - p_1 - p_4 \le 0, \\
&-1 \le p_{23} + p_{24} + p_{14} - p_{13} - p_2 - p_4 \le 0, \\
&-1 \le p_{14} + p_{13} + p_{23} - p_{24} - p_1 - p_3 \le 0, \\
&-1 \le p_{24} + p_{23} + p_{13} - p_{14} - p_2 - p_3 \le 0.
\end{aligned} \tag{16}$$

(16) reminds us the well known *Clauser-Horne inequalities*[17]

# 6 Appendix II.

Consider an Aspect-type EPR experiment with spin-$\frac{1}{2}$ particles (Fig. 5). The four detectors detect the spin-up events. The two switches are making choice from sending the particles to the Stern-Gerlach magnets directed into different directions. The *observed* events are the followings:

$A_1$ : The "left particle has spin 'up' along direction **a**" detector beeps
$A_2$ : The "left particle has spin 'up' along direction **a′**" detector beeps
$A_3$ : The "right particle has spin 'up' along direction **b**" detector beeps
$A_4$ : The "right particle has spin 'up' along direction **b′**" detector beeps
$a_1$ : The left switch selects direction **a**
$a_2$ : The left switch selects direction **a′**
$a_3$ : The right switch selects direction **b**
$a_4$ : The right switch selects direction **b′**

[17]See J. F. Clauser and A. Shimony, Bell's theorem: experimental tests and implications, *Rep. Prog. Phys.* **41** (1978), 1881-1927. There is, however, an important conceptual disagreement between (16) and the original Clauser-Horne inequalities, see L. E. Szabó, Quantum mechanics in an entirely deterministic universe, *Int. J. Theor. Phys.*, **34** (1995), 1751-1766.

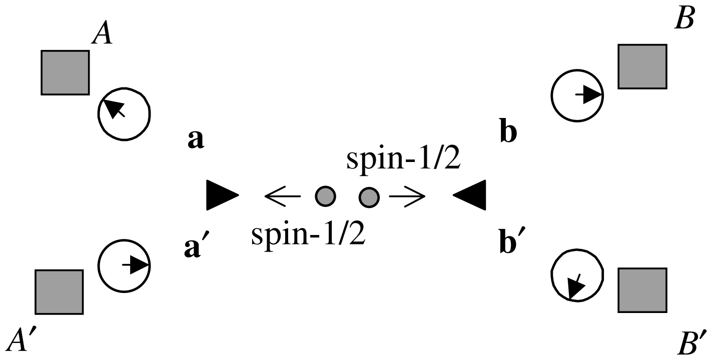


Figure 5: *The Aspect experiment with spin-$\frac{1}{2}$ particles*

For the probabilities of these events, in case of $\angle(\mathbf{a},\mathbf{a}') = \angle(\mathbf{a}',\mathbf{b}) = \angle(\mathbf{a},\mathbf{b}') = 120°$ and $\angle(\mathbf{b},\mathbf{a}') = 0$, we have

$$
\begin{aligned}
p(A_1) = p(A_2) = p(A_3) = p(A_4) &= \frac{1}{4} \\
p(a_1) = p(a_2) = p(a_3) = p(a_4) &= \frac{1}{2} \\
p(A_1 \wedge a_1) = p(A_1) &= \frac{1}{4} \\
p(A_2 \wedge a_2) = p(A_2) &= \frac{1}{4} \\
p(A_3 \wedge a_3) = p(A_3) &= \frac{1}{4} \\
p(A_4 \wedge a_4) = p(A_4) &= \frac{1}{4} \\
p(A_1 \wedge a_2) = p(A_2 \wedge a_1) = p(A_3 \wedge a_4) = p(A_4 \wedge a_3) &= 0 \qquad (17) \\
p(A_1 \wedge A_3) = p(A_1 \wedge A_4) = p(A_2 \wedge A_4) &= \frac{3}{32} \\
p(A_2 \wedge A_3) &= 0 \\
p(a_1 \wedge a_2) = p(a_3 \wedge a_4) &= 0 \\
p(a_1 \wedge a_3) = p(a_1 \wedge a_4) = p(a_2 \wedge a_3) = p(a_2 \wedge a_4) &= \frac{1}{4} \\
p(A_1 \wedge a_3) = p(A_1 \wedge a_4) = p(A_2 \wedge a_3) = p(A_2 \wedge a_4) & \\
= p(A_3 \wedge a_1) = p(A_3 \wedge a_2) = p(A_4 \wedge a_1) = p(A_4 \wedge a_2) &= \frac{1}{8}
\end{aligned}
$$

These statistical data *agree* with quantum mechanical results, in the following sense:

$$\begin{aligned}
\frac{p(A_1 \wedge a_1)}{p(a_1)} = tr(WA_1) = \frac{p(A_2 \wedge a_2)}{p(a_2)} = tr(WA_2) & \\
= \frac{p(A_3 \wedge a_3)}{p(a_3)} = tr(WA_3) = \frac{p(A_4 \wedge a_4)}{p(a_4)} = tr(WA_4) &= \frac{1}{2} \\
\frac{p(A_1 \wedge A_3 \wedge a_1 \wedge a_3)}{p(a_1 \wedge a_3)} = \frac{p(A_1 \wedge A_3)}{p(a_1 \wedge a_3)} = tr(WA_1A_3) = \frac{1}{2}\sin^2 \angle(\mathbf{a}, \mathbf{b}) &= \frac{3}{8} \\
\frac{p(A_1 \wedge A_4 \wedge a_1 \wedge a_4)}{p(a_1 \wedge a_4)} = \frac{p(A_1 \wedge A_4)}{p(a_1 \wedge a_4)} = tr(WA_1A_4) = \frac{1}{2}\sin^2 \angle(\mathbf{a}, \mathbf{b}') &= \frac{3}{8} \qquad (18) \\
\frac{p(A_2 \wedge A_3 \wedge a_2 \wedge a_3)}{p(a_2 \wedge a_3)} = \frac{p(A_2 \wedge A_3)}{p(a_2 \wedge a_3)} = tr(WA_2A_3) = \frac{1}{2}\sin^2 \angle(\mathbf{a}', \mathbf{b}) &= 0, \\
\frac{p(A_2 \wedge A_4 \wedge a_2 \wedge a_4)}{p(a_2 \wedge a_4)} = \frac{p(A_2 \wedge A_4)}{p(a_2 \wedge a_4)} = tr(WA_2A_4) = \frac{1}{2}\sin^2 \angle(\mathbf{a}', \mathbf{b}') &= \frac{3}{8}
\end{aligned}$$

where the outcomes are identified with the following projectors

$$\begin{aligned}
A_1 &= P_{span\left\{\psi_{+\mathbf{a}}\otimes\psi_{+\mathbf{a}},\psi_{+\mathbf{a}}\otimes\psi_{-\mathbf{a}}\right\}} \\
A_2 &= P_{span\left\{\psi_{+\mathbf{a}'}\otimes\psi_{+\mathbf{a}'},\psi_{+\mathbf{a}'}\otimes\psi_{-\mathbf{a}'}\right\}} \\
A_3 &= P_{span\left\{\psi_{-\mathbf{b}}\otimes\psi_{+\mathbf{b}},\psi_{+\mathbf{b}}\otimes\psi_{+\mathbf{b}}\right\}} \\
A_4 &= P_{span\left\{\psi_{-\mathbf{b}'}\otimes\psi_{+\mathbf{b}'},\psi_{+\mathbf{b}'}\otimes\psi_{+\mathbf{b}'}\right\}}
\end{aligned}$$

of the Hilbert space $H^2 \otimes H^2$. The state of the system is assumed to be represented by $W = P_{\Psi_s}$, where $\Psi_s = \frac{1}{\sqrt{2}}\left(\psi_{+\mathbf{a}} \otimes \psi_{-\mathbf{a}} - \psi_{-\mathbf{a}} \otimes \psi_{+\mathbf{a}}\right)$.