Comment on "The quantum pigeonhole principle and the nature of quantum correlations"

by Y. Aharonov, F. Colombe, S. Popescu, I. Sabadini, D. C. Struppa, and J. Tollaksen

Analysis by Stephen Parrott
Version 2

Changes in Version 2: There are no substantive changes. The exposition has been expanded. A bad repeated typo has been corrected. In the original, some projectors Π_j^A were defined, and thereafter called P_j^A! So annoying! If any reader is puzzled by something similar, please let me know. I welcome comments.

1 Introduction

The first paragraph of the paper, which it set in boldface type, reads:

> **"The pigeonhole principle: 'If you put three pigeons in two pigeonholes, at least two of the pigeons end up in the same hole' is an obvious yet fundamental priniciple of Nature as it captures the very essence of counting. Here however, we show that in quantum mechanics, this is not true! We find instances when three quantum particles are put in two boxes, yet no two particles are in the same box."**

When I saw this a few weeks ago, I looked forward to an amusing puzzle, in view of previous experience with writings of some of the authors. I did not expect to find a convincing quantum violation of the classical pigeonhole principle, and after reading the paper's short and simple example, I did not find one. Indeed, I wondered if anyone would take seriously the authors' conclusion:

> "In conclusion, we presented a new quantum effect that requires us to revisit some of the most basic notions of quantum physics – the notions of separability, of correlations and interactions. It is still very early to say what the implications of this revision are, but we feel one should expect them to be major since we are dealing with such fundamental concepts."

Today I found in my inbox a free sample of some articles in Physics World, the member magazine of the Institute of Physics. One of them is entitled "Paradoxical pigeons are the latest quantum conundrum", by science writer Marcus Woo. It quotes one of the authors of "Pigeonhole principle", Jeff Tollaksen, as follows:

> "It [the quantum pigeonhole principle] really has immense implications."

> "This is at least as equally profound, if not more profound [than the Einstein-Rosen-Podolsky paradox]."

So, I guess that at least the editors of Physics World must take seriously the authors' conclusion that their discovery will require "major" revisions of "some of the most basic notions of quantum physics".

2 The authors' argument

Following is a summary of the authors' argument. The reader is urged to have on hand the original from http://arxiv.org/abs/1407.3194. The original argument occupies only about a page, and is simple.

First the authors consider a system with one quantum particle and two boxes denoted L (left) and R (right). The system is in state

$$|+\rangle = \frac{1}{\sqrt{2}}(|L\rangle + |R\rangle) \quad . \tag{1}$$

Then the authors consider a tripartite system describing three particles (indexed $1, 2, 3$), in an initial state

$$\Psi := |+\rangle_1 |+\rangle_2 |+\rangle_3 \quad .$$

They state:

> "Now, it is obvious that in this state any two particles have non-zero probability to be found in the same box. We want, however, to show that there are instances in which we can guarantee that no two particles are together; we cannot arrange that to happen in every instance, but, crucially, there are instances like that. To find those instances we subject each particle to a *second measurement* [emphasis mine] . . .

WHOA!

At this point, I did a double take. The authors are talking about a *second* measurement, but they haven't said anything about a *first measurement*!

Reading further into the argument, it turns out that the first measurement is an *imaginary* measurement — it isn't actually performed, but we *imagine* how it could have come out *if* it had been performed.

In a usual jargon, this first measurement is a *counterfactual* measurement, and the authors' argument is a counterfactual argument. It is well known that counterfactual arguments are problematic in quantum mechanics, but let's soldier on and see what it leads to in this case.

What the authors call the "second measurement" is a genuine measurement which the authors define. In the context of the authors' argument, it can be a considered as a so-called "postselection" in which the initial state Ψ is "postselected" onto a final state Φ defined by

$$\Phi := |+i\rangle_1 \, |+i\rangle_2 \, |+i\rangle_3 \quad \text{with } |+i\rangle := (|L\rangle + i|R\rangle)/\sqrt{2} \quad .$$

This is not precisely how the authors characterize it, but the authors' more complicated measurement (with respect to 8 orthogonal projectors) yields the postselection as one of its results because the projector onto Φ is one of the 8 orthogonal projectors defining the measurement.

The authors are interested only in the case in which the postselection is successful — i.e., after the "second" measurement (the first measurement being unperformed), the initial state Ψ has been transformed into the final state Φ.

Next, they return to discuss the imaginary first measurement, which so far has not even been defined. The authors describe it intuitively as follows:

> "Let us now check whether two of the particles are in the same box. Since the state [Ψ] is symmetric, we couold focus on particles 1 and 2 without any loss of generality – any result obtained for this pair applies to any other pair."

They then define a measurement with respect to two orthogonal projectors $\{\Pi_{1,2}^{same}, \Pi_{1,2}^{diff}\}$ which intuitively yields one result if particles 1 and 2 "are in the same box", and the other result if they are not.

They argue that if the imaginary first measurement *had* resulted in the two particles being in different boxes, and that imaginary measurement had been followed by the genuine second measurement, then Φ could not have been obtained as the postselected state. Hence (according to the authors), particles 1 and 2 were in different boxes all along, and the same for the other two pairings of particles 2 and 3 and particles 1 and 3.

The implicit assumption seems to be that when the quantum system is in state Ψ, it is meaningful to speak of particles 1 and 2 (say) being in different boxes. That is, the authors' argument seems to rest on an unstated assumption that there is some underlying classical reality in which each particle is in some definite box. (If this were not the case, then how could one speak of particles 1 and 2 being in different boxes?)

Bell's Theorem[1] as well as many other quantum "paradoxes"[2] have taught us that an assumption that quantum measurements reveal some underlying reality existing before the measurement often leads to apparent paradoxes. I would not hesitate to characterize this as "profound". However, I can't see how the authors' "quantum pigeonhole" effect is of a different nature, a new profundity.

The main difference that I see is that in Bell's and many "paradoxes" of the same nature, "reality" assumptions are often mixed up with "locality" assumptions.[3] Sometimes one can save "reality" by denying "locality", or vice versa. But the "quantum pigeonhole" effect seems to have nothing to do with locality. I am not sufficiently familiar with quantum paradoxes to know if there are other "paradoxes" which do not involve "locality"

[1]For references, see any good, modern quantum mechanics text beyond the elementary level.

[2]My favorite for its straightforward and totally surprising nature is [2].

[3]"Locality" assumptions posit that physical objects cannot travel faster than the speed of light, that information cannot be transmitted faster than the speed of light, that spacelike separated observers cannot communicate, and other assumptions of that nature.

The authors seem to view their argument as showing that three objects can occupy two boxes with no two of them in the same box. But I think it can equally well be viewed as denying that it is permissible to speak of a particle as being in a definite box before it is observed to be in that box.

And I would be surprised if many will view that last observation as profound. To illustrate, consider the one-particle state (1):

$$|+\rangle = \frac{1}{\sqrt{2}}(|L\rangle + |R\rangle) \quad .$$

Is the particle in the left box or the right box at this moment? The reader is challenged to mentally formulate an answer.

I can think of only three possibly sensible answers:

1. The particle is in both boxes because if measured, it will sometimes appear in one box and sometimes in the other.

2. The particle is in neither box because there is no box in which it will surely appear if measured.

3. The question is meaningless.

My personal choice is answer 3. Given the reader's choice, how can a later postselection (tomorrow, say) possibly change his or her answer unless one goes so far as to claim that the future can influence the past?[4]

Of course, a later postselection can give more information. For example, if we postselect on L tomorrow, then after the postselection, we can be sure that the particle is in the left box. But how can that change whether *today* the particle is in both boxes or neither? In the context of their more complicated example, that is what the authors appear to be claiming.

It is instructive to ask the same questions for a particle in a mixed state

$$\rho = \frac{1}{2}P^L + \frac{1}{2}P^R \quad ,$$

where for $A = L$ or R, P^A denotes the projector on $|A\rangle$. This is analogous to a classical situation in which the particle can be said to be definitely in one box or the other, but we don't know which.

Suppose that tomorrow, we postselect on $|L\rangle$. This means that we measure with respect to the orthonormal basis $|L\rangle$, $|R\rangle$, and only consider the case in which the measurement yielded $|L\rangle$. Suppose that tomorrow the measurer claims that the particle had to have been in the left box all along because if it had been in the right box, we could not have obtained the result $|L\rangle$ for the postselection.

In a classical situation in which we agree that it makes sense to say today that the particle is definitely in one box or the other (but we don't know which), this would seem a correct conclusion. It might be a justifiable conclusion for the quantum mixed state ρ.

[4]At least two of the authors do appear to have gone this far in previous publications, but in the present paper they do not.

But it is at least questionable (I imagine that most would say it is wrong) for the pure quantum state $(|L\rangle + |R\rangle)/\sqrt{2}$. This is because interference effects (as in a Mach-Zehnder interferometer) make untenable the classical view that the particle is definitely in one box or the other before we measure it.

It seems to me that the authors' argument is just a more complicated way of presenting the above conundrum. It seems to rest on a classical view that each of its three particles can be said to be definitely in one box or the other. My personal view is that there is indeed something strange and profound to think about here, but that the profundity is already present in the simpler example just analyzed.

3 A simple alternative view

The authors' argument involves an imaginary measurement on particles 1 and 2 with respect to two orthogonal projectors $\Pi^{same}_{1,2}$ and $\Pi^{diff}_{1,2}$. In a two-particle space these constitute a resolution of the identity (i.e., a collection of orthogonal projectors which sum to the identity). In the three-particle space, tensoring these with the identity I_3 for the Hilbert space of the third particle yields a resolution of the identity $\{\Pi^{same}_{1,2} \otimes I_3,\, \Pi^{diff}_{1,2} \otimes I_3\,\}$ for the three-particle space. The authors state correctly that the argument given for particles 1 and 2 also applies by symmetry to particles 1 and 3 and to particles 2 and 2.

However, this is only part of the story. If we collect together (using a fairly obvious notation)[5] all of the six operators corresponding to the three choices of two-particle spaces,

$$\begin{aligned} \{ \quad & \Pi^{same}_{1,2} \otimes I_3,\ \Pi^{diff}_{1,2} \otimes I_3, \\ & \Pi^{same}_{1,3} \otimes I_2,\ \Pi^{diff}_{1,3} \otimes I_2, \\ & \Pi^{same}_{2,3} \otimes I_1,\ \Pi^{diff}_{2,3} \otimes I_1\ \}, \end{aligned} \tag{2}$$

we do not obtain a resolution of the identity for the three-particle space because the ranges of these six projectors are not orthogonal, and they do not sum to the identity. Therefore, *they do not define a measurement.* However, this nonexistent "measurement" is precisely the "intermediate measurement" which the authors discuss around their equation (4).

Now the magician's sleight of hand begins to reveal itself.[6] For each of the

[5] However, the notation does have to be interpreted sympathetically. The problem is that while the notation for the tensor product of an operator on the two-particle space corresponding to particles 1 and 2 and an operator on the space of particle 3 is obvious, e.g., $\Pi^{same}_{1,2} \otimes I_3$, how do we denote a corresponding operator on the two-particle space corresponding to particles 1 and 3 like $\Pi^{same}_{1,3}$, tensored with an operator on the space of particle 2, like I_2. We use $\Pi^{same}_{1,3} \otimes I_2$ even though technically, the I_2 should go in the "middle". Inventing a new notation to handle this technical problem (such as defining an interchange operator which interchanges the spaces of particles 2 and 3 and writing the various expressions in terms of the interchange operator) seems more trouble than it is worth.

[6] Of course, this metaphor is not intended to imply that the authors deliberately misled us.

three two-element subsets of $\{1, 2, 3\}$, the corresponding operators of our (2) do constitute a resolution of the identity and so define a genuine measurement. For example, for the subset $\{1, 2\}$,

$$\mathcal{M}_{1,2} := \{\Pi_{1,2}^{same} \otimes I_3,\, \Pi_{1,2}^{diff} \otimes I_3\},$$

defines a bona fide measurement. So do the corresponding $\mathcal{M}_{1,3}$ and $\mathcal{M}_{2,3}$ defined by the second and third lines of (2), respectively.

But how do we perform all three measurements together? The reader is urged to think about this before reading the subsequent discussion.

Before proceeding, we make one more observation. Since the intermediate measurement is only an imaginary measurement, one could argue that we could simply *imagine* making three two-projector measurements $\mathcal{M}_{i,j}$ where i, j range over all two element subsets of $\{1, 2, 3\}$. Perhaps (though still questionable), if we admit the validity of discussing imaginary measurements as if they were real.[7] But then we should also admit the validity of basing conclusions on any *other* imaginary measurements which we might dream up. We shall see that there is such an imaginary measurement which yields the opposite conclusion from the paper's.

Though the union of the six operators comprising the three measurements $\mathcal{M}_1, \mathcal{M}_2$, and $\mathcal{M}_3$ do not comprise a measurement they do commute, so we can obtain a resolution of the identity by replacing them by the following obvious refinement.[8]

For $A = L$ or R and $j = 1, 2,$ or 3, Let Π_j^A denote the projector on $|A\rangle$ in the Hilbert space of particle j. For example, Π_3^R is the projector on $|R\rangle$ in the Hilbert space of particle 3.

Then the eight possible products $\Pi_1^{A_1} \otimes \Pi_2^{A_2} \otimes \Pi_3^{A_3}$ with each $A_j = L$ or $R, j = 1, 2, 3$, form a resolution of the identity for the three-particle space which refines the projectors of equation (2). For example,

$$\begin{aligned}
&\Pi_{1,2}^{same} \otimes I_3 = \\
&\qquad \Pi_1^L \otimes \Pi_2^L \otimes I_3 + \Pi_1^R \otimes \Pi_2^R \otimes I_3 \\
&\quad = \quad \Pi_1^L \otimes \Pi_2^L \otimes \Pi_3^L + \Pi_1^L \otimes \Pi_2^L \otimes \Pi_3^R + \Pi_1^R \otimes \Pi_2^R \otimes \Pi_3^L + \Pi_1^R \otimes \Pi_2^L \otimes \Pi_3^R,
\end{aligned}$$

and similarly for the other particle pairs. Although technically, we cannot measure with respect to the projectors of (2), we can measure with respect to this refinement, and the latter measurement gives the information that the paper's argument requires.

Actually, it gives more information. For example, if the final result corresponds to the projector $\Pi_1^1 \otimes \Pi_2^1 \otimes \Pi_3^1$ it gives the information that after the

[7] If in real life we performed measurement $\mathcal{M}_{1,2}$, that would alter the pre-measurement state Ψ, and also in general the result of a subsequent measurement $\mathcal{M}_{1,3}$, so perhaps we should not allow imagining performing $\mathcal{M}_{1,2}$ and $\mathcal{M}_{1,3}$ simultaneously. The rules supposed to be obeyed by such imaginary measurements are not clear to me!

[8] A *refinement* of a collection of commuting projectors $\{P_j\}$ is a collection of mutually orthogonal projectors $\{Q_k\}$ such that each P_j is a sum of some collection of Q_k.

measurement, all three particles are in Box 1, information that could not be obtained from knowing the answer to the three questions "Are particles i and j in the same or different boxes" for the three two-element subsets $\{i, j\}$ of $\{1, 2, 3\}$. This will be discussed more fully below.[9]

No matter what the result of the measurement with the refinement, we always find two or more particles in the same box! And all measurement results are compatible with success of the postselection on Φ. If we want, we can choose to view measurement with respect to the refinement as a counterfactual one which we only imagine performing, like the author's intermediate "measurement" (which technically is not a measurement, counterfactual or not, because (2) is not a resolution of the identity).

This new counterfactual argument parallels the paper's, but gives a different result. If the paper's reasoning is valid, then this new reasoning should be valid, too. The big difference between the two, is that counterfacuality is an extraneous element of the new argument, inserted only for comparison with the authors'. The new argument does not require conterfactuality, unlike the authors'.

How can two valid arguments give different results? I don't think they can. My interpretation is that the new argument is valid because it can be formulated in ordinary, noncounterfactual terms, and that the authors' argument should be regarded as either invalid or meaningless, according to taste.

Before leaving the topic, I should acknowledge that what we have been illustrating with our refined measurement is discussed in the paper's section entitled "THE NATURE OF QUANTUM CORRELATIONS", which I found difficult on first reading, and still cannot follow in complete detail. As I interpret it, this section observes that our refined measurement actually gives *more* information than their argument requires. The authors view this as a drawback because it disturbs the initial state Ψ more than would their (seemingly) proposed "measurement" with respect to the six projectors $\Pi^{same}_{i,j}$, $\Pi^{diff}_{i,j}$ where i, j ranges over all two-element subsets of $\{1, 2, 3\}$. The paper does not seem to recognize that its six-projector "measurement" is not actually a measurement.[10] To make it into a measurement, it must be refined to the eight-projector measurement. The refinement inevitably gives more information. Apparently for this reason, the authors consider the measurement of the refinement as being different from their proposed six-projector "intermediate measurement". But if it is regarded as different, then the intermediate measurement is not a "measurement" in the standard usage of the term.

The authors have not told us how their "intermediate measurement" can

[9]Some readers might view as splitting hairs the difference between the (improper) "measurement" with the six projectors of (2) and the genuine measurement with the eight projectors of the refinement of (2). A reader could be pardoned for considering them as effectively the "same". Yet this distinction is made by the authors and is critical to their argument. Because the distinction is important to the paper's argument, we must make it too.

[10]The paper repeatedly refers vaguely to its intermediate procedure as a "measurement" or a "global measurement" without precisely defining it, but never actually states that it is comprised of the six projectors just named.

be implemented within standard quantum mechanics without implementing the refinement. It may not make any difference because the intermediate measurement is only an imaginary measurement anyway. It seems to me that as soon as we open up the Pandora's Box of allowing the discussion of imaginary measurements which are never made, we enter some kind of Never-Never Land.

On page 3, the authors express the opinion that their "global measurement" (identical to the "inermediate measurement" so far as I can see)

> "is in some sense, better than the detailed meaurement [presumably our refinement] as it delivers the information about correlations while minimizing the disturbance that it produces to the state."

As an amusing final observation, note that actually, their intermediate "measurement" does not disturb the state Ψ at all because it is an imaginary, counterfactual measurement! The same is true for our eight-projector refinement if we agree not to actually perform it, but only to *imagine* performing it!

References

[1] Y. Aharonov, F. Colombe, S. Popescu, I. Sabadini, D. C. Struppa, and J. Tollaksen, "The quantum pigeonhole principle and the nature of quantum correlations", arXiv:1407.3194v1

[2] P. G. Kwiat and L. Hardy, "The mystery of the quantum cakes", Am. J. Phys. **68** (2000), 33-36