A Proof of Specker's Principle

Guido Bacciagaluppi*

4 March 2023

Abstract

Specker's principle, the condition that pairwise orthogonal propositions must be jointly orthogonal, has been much investigated recently within the programme of finding physical principles to characterise quantum mechanics. It largely appears, however, to lack a transparent justification. In this paper, I provide a derivation of Specker's principle from three assumptions (made suitably precise): the existence of maximal entanglement, the existence of non-maximal measurements, and no-signalling. I discuss these three assumptions and describe canonical examples of non-Specker sets of propositions satisfying any two of them. These examples display analogies with various approaches in the interpretation of quantum mechanics, notably ones based on retrocausation. I also discuss connections with the work of Popescu and Rohrlich. The core of the proof (and the main example violating no-signalling) is illustrated by a variant of Specker's tale of the seer of Nineveh, with which I open the paper.

*Descartes Centre for the History and Philosophy of the Sciences and the Humanities and Freudenthal Institute, Utrecht University; IHPST and SPHERE, Paris; email: g.bacciagaluppi@uu.nl .

Specker's seer meets Popescu and Rohrlich

At the Assyrian school for prophets in Arba'ilu, there taught in the age of king Asarhaddon a sage from Nineveh. He was an outstanding representative of his discipline (solar and lunar eclipses), who except for the heavenly bodies had thoughts almost only for his *two* daughters. His teaching success was modest; the discipline was considered dry, and also required previous mathematical knowledge that was scarcely available. If thus in his teaching he failed to capture the attention he wanted from the students, he received it overabundantly in a different area: no sooner had his daughters reached the marriageable age than he was flooded with requests for their hands from students and young graduates. And even though he did not imagine wishing to keep them with him forever, yet they were still far too young, and the suitors in no way worthy of them. And so that they should all convince themselves of their unworthiness, he promised their hands to those two who could perform a set prophecy task. The suitors were led in front of two tables on each of which stood three boxes in a row, and urged to say which boxes contained a gem and which were empty. Yet, as many as would attempt it, it appeared impossible to perform the task. Indeed, after they both had made their prophecies, each of the suitors was urged by the father to open two boxes that he had named as both empty or as both not empty: it always proved to be that one contained a gem and the other did not, and in fact the gem lay now in the first, now in the second of the boxes that were opened; and yet, every time the suitors both opened the first box, or both the second, or both the third, whatever one of them found (or failed to find) in one box, the other would also find (or fail to find) in the corresponding box – which showed that the gems needs must have been contained in the boxes in the first place. But how should it be possible, out of three boxes, to name no two as empty or as not empty? Thus indeed the daughters would have remained unmarried until their father's death, had they not encouraged two bright young students they were secretly in love with to attempt the task, whose names, let it be known, were Sandu and Daniel. Now, Sandu and Daniel were not renowned at the time as having any particular gift for prophecy, but they were very ingenious and hard-working, and Daniel's uncle was a prophet of considerable standing, so Daniel hoped to have inherited some of the same gift. Besides, the two friends were both desperately in love: and so they picked up their courage and sought an audience with their beloved ones' father. When he heard Sandu and Daniel asking him for his daughters' hands, the father smiled to himself, and (judging they had not the slightest chance of succeeding in the task) declared that not only were they welcome to come upon the morrow to test their skills, but that he would give them not one, nor two, but one hundred chances if they should care to try. And so the father was awake all night setting up two hundred little tables and six hundred little boxes and placing gems from his collection in at least two hundred of them (for he had won many prizes in prophecy competitions over the years). And the next morning Sandu and Daniel duly presented themselves to the sage's home and embarked on making prophecies. As soon as they had both performed a set of prophecies, at the father's urging each of them opened two boxes that he had named as both empty or as both not empty, and, lo and behold: it always proved to be that one contained a gem and the other did not, and the gem lay now in the first, now in the second of the boxes that were opened; and yet, every time they both opened the first box, or both the second, or both the third, whatever one of them found (or failed to find) in a box, the other would also find (or fail to find) in the corresponding box – just as had happened to every pair of suitors before them (and as the two daughters had long grown accustomed to). And so they laboured all morning (for prophesying is a strenuous task, especially if you are still a student), becoming ever more disconsolate at each unsuccessful attempt. At the end of fifty trials, the father called a break, and kindly offered Sandu and Daniel some tea and cucumber sandwiches of which they partook, grateful of a rest and of the opportunity to exchange some precious words with their beloved ones. When the hour was over and the final fifty trials were to begin, the younger of the two daughters walked up to her father and spoke to him thus: 'Father, Sandu and Daniel realise that you are a far better prophet than they are, and that if you choose the boxes that are to be opened, then in no single trial will their prophecies stand. My sister and I therefore humbly beseech you (for we are very fond of them, even though you must be laughing at their youth and inexperience) to give them one small chance of success. If you will, demand they be successful in all fifty trials that still lie before them, but please let them choose themselves which boxes shall be opened and prophesy which shall be found empty or full. If they both fail, they will lay down any claim to our hands forever. But if at least one of them succeeds (for neither of us could ever be happy if the other is not), oh! please let your loving daughters be married to them, and we shall forever and a day be grateful to you, dear Father'. Now, well knowing how tiny the chance was of either Sandu or Daniel successfully predicting fifty trials at equal odds (unless they had a very great gift for prophecy, indeed), but equally moved by this eloquence (because the younger daughter who had thus spoken had her own subtle gift of winding her father around her little finger), the father consented to this alteration of the remaining trials, of course as long as Sandu and Daniel still made their prophecies before seeing the other's boxes opened. And so Sandu and Daniel embarked again on making prophecies for the rest of the day. Upon each trial, first

Daniel announced which two boxes he would open (which by then all expected would contain one gem between them) and prophesied which one would be full and which one would be empty. Then Sandu chose his first box, prophesied whether this would be full or empty, and opened the box to verify whether or not his prophecy was correct. Then he chose a second box, which of course he correctly prophesied to be full if the first had been empty and empty if the first had been full. Finally Daniel opened his two chosen boxes and verified his previous prophecy. Upon the first trial, both Sandu's and Daniel's prophecies proved to be successful, and so did they upon the second trial, and upon the third. Upon the fourth trial, however, Sandu's first prophecy was falsified, and in the subsequent trials it became apparent that the first of Sandu's prophecies met now with success, now with failure, as if he had no gift for prophecy at all (of course his second prophecy was always successful). But the father had already turned pale by the end of the second trial, realising he had been outsmarted. And, indeed, Daniel – whether his family gift had found his way to him at last or through some other artifice, and to the increasing surprise and delight of the numerous bystanders (for by this time the father's servants and neighbours had long started gathering around) – kept meeting with success, trial upon trial upon trial. Just before the hundredth and last trial, the father interjected, protesting weakly that he had not meant Sandu to announce his second prophecy only after having already verified the first one. To this now the elder of his daughters replied: what difference did that make to Sandu's chances of success, for it made no difference to his first prophecy, and the second stood or fell with the first? And so, grumbling, the father let the final trial proceed, and from the last of his boxes (which he had prophesied to be full) Daniel triumphantly extracted a sparkling emerald that matched one that Sandu had just extracted from his. The four young people were married the very next day, and henceforth and for the rest of their lives both Sandu and Daniel enjoyed a reputation as formidable prophets. Meanwhile, the father consoled himself in the knowledge of having brought up two very clever daughters, indeed.

1 Introduction

Specker's principle is one of the names given to a condition on sets of 'propositions which are not simultaneously decidable' (Specker 1960). The original version of the condition is that sets of pairwise compatible propositions must be jointly compatible, and as such it enters as an axiom in the definition of partial Boolean algebras by Kochen and Specker (1965). In the literature up to the 1980s this condition is also known as 'coherence' or 'regularity', and a central result of quantum logic is that every coherent orthomodular poset is canonically isomorphic to a transitive partial Boolean algebra (Finch 1969, Gudder 1972). A weaker version of the same condition is that pairwise orthogonal propositions must be jointly orthogonal. This is known as 'orthocoherence' in the earlier literature, and is a key feature of the discussion of tensor products in quantum logic (Foulis and Randall 1981).[1] Roughly since the appearance of the paper by Liang, Spekkens and Wiseman (2011), this weaker version of the principle has been the object of renewed discussion in the context of the search for physical principles characterising quantum mechanics. Slightly different variants have been used to this end by Fritz *et al.* (2013), Henson (2012) and Cabello (2012, 2013), under the names of 'local orthogonality', 'consistent exclusivity', and 'Specker's principle'. (I am adopting the last name in this paper.)

While Specker's principle is easy to formulate and rather powerful in applications,[2] its physical justification has remained rather obscure. For instance, Kochen himself has stated: 'Ernst and I spent many hours discussing the principle [...]. The difficulty lies in trying to justify it on general physical grounds, without already assuming the Hilbert space formalism of quantum mechanics. We decided to incorporate the principle as an axiom in our definition of partial Boolean algebras [...]. I have never found a general physical justification for [it]' (as reported in Cabello 2012).[3]

It is easy to construct toy examples that violate Specker's principle, and the prime one is Specker's (1960) own fable of the seer of Nineveh, adapted above and discussed in detail by Liang, Spekkens and Wiseman (2011) (a related example is the 'firefly box', to which I return in Section 3.2). In the original variant, Specker tells the tale of a seer who sets a prophecy task to his (single) daughter's suitors. The suitors are asked to prophesy which of three boxes will reveal a gem when opened and which will be empty. After their prophecy, the father requires them to open two boxes that they had indicated as being both full or both empty, but of these one always turns

[1]For a helpful overview of classic issues in quantum logic, see Hughes (1989, Chapter 7).

[2]See for instance Cabello's derivation of the quantum bound for the Klyachko inequality (2013).

[3]The one exception I know is provided in the paper by Chiribella and Yuan (2014), which gives a proof of Specker's principle in the context of the general probabilistic theories of Chiribella, D'Ariano and Perinotti (2010). The proof relies on a notion of sharp measurements and on the assumption that each measurement can be implemented through such a sharp measurement. Chiribella, D'Ariano and Perinotti's framework also includes a requirement called causality, i.e. no operational retrocausation. Note that the assumptions in our proof also lead to the impossibility of any retrocausal mechanisms manifesting themselves at the operational level (see Section 3.3).

out to be full and one empty, so that the suitors' prophecies are always falsified. One day, however, the daughter without warning opens two boxes predicted one to be full and one to be empty, which turns out to be correct – while the third box cannot be opened.[4] This is an example of a violation of Specker's principle (both the weaker and the stronger versions), because one can always open boxes A and B together, or B and C, or C and A, and the results are always mutually exclusive, i.e. the corresponding propositions are pairwise orthogonal (and thus in particular pairwise compatible); but the three propositions are clearly not jointly compatible (and thus in particular not jointly orthogonal), because there is no joint probability distribution over the three corresponding events that would return the flat distribution $p = 1/2$ for the three marginals.

This paper provides one possible justification for Specker's principle (in the weaker version) – in fact a surprisingly simple one – deriving it from a combination of the existence of an analogue of maximally entangled states, the existence of 'combinable' or 'non-maximal' measurements, and no-signalling. (The first two conditions will be made suitably precise.) Insofar as maximally entangled states are a nonlocal feature of quantum theory, the derivation presented here can be seen as a variant of Popescu and Rohrlich's (1994) strategy of trying to derive quantum theory from a suitable combination of nonlocality and no-signalling. There are further connections with Popescu and Rohrlich's work, specifically with Popescu–Rohrlich nonlocal boxes (PR boxes), on which I shall also comment.

I present the main result in Section 2. Then in Section 3 I discuss the three assumptions in turn, and give examples of failures of each of the three. These examples display analogies with various approaches to the foundations of quantum mechanics, in particular retrocausal ones. Finally, in Section 4 I briefly address the connection with PR boxes.

2020 marked the centenary of Ernst Specker's birth. I had the honour and the pleasure of having him as one of my most inspiring teachers at ETH Zürich, as well as of hearing the tale of the seer of Nineveh from his own mouth sometime around the spring of 1985. I fondly and respectfully dedicate this paper to Professor Specker's memory.

[4] There are two full translations of Specker's (1960) paper that contains the original fable, by Stairs (1975) and by Seevinck (2011). Liang, Spekkens and Wiseman (2011) include an amalgam of the two. I have based myself on my own translation of the fable in Bacciagaluppi (2016).

2 The proof

The formulation of Specker's principle that we shall use is: for any n, if a set of propositions $A_1, \ldots, A_n$ are pairwise orthogonal, they are jointly orthogonal.

As the framework for the proof we shall assume any propositional structure allowing for a suitable notion of orthogonality, in the sense that orthogonality of a pair or a set of propositions implies that the propositions are jointly compatible (i.e. there is an experiment in which they can be jointly measured) and mutually exclusive (i.e. in all experiments in which they can be jointly measured they are mutually exclusive in all states). The proof does not depend on the exact details of the structure chosen, except that we assume additionally that there is some appropriate notion of composition defined for it, which we symbolically denote by a tensor product (bearing in mind of course that one of the main lessons of the quantum logic programme is that well-behaved tensor products of propositional structures are difficult to define in general).[5]

The explicit assumptions to be used in our proof are:

(a) *Existence of maximal entanglement*: We require that there exist the composition of two copies of the same system and a state on the resulting composite in which any non-trivial proposition A of a single system will have non-trivial probability,[6] and any measurement results for copy 1 of the system and for copy 2 of the system will be perfectly correlated (even though the propositions of a single system need not be all jointly compatible).

(b) *Existence of combinable measurements*: For any proposition A, there is a set $\mathcal{C}_A$ of measurements of A that can be carried out locally (i.e. only involve a procedure on the system in question), and such that if the propositions A_i are jointly

[5]For definiteness we could take orthoalgebras – one of the best current candidates for providing an abstract setting for a generalised probability theory (for a good discussion, see Hardegree and Frazer (1981)). Every orthocoherent orthoalgebra is automatically an orthomodular poset (so every coherent orthoalgebra is automatically a transitive partial Boolean algebra) and if enough states exist one can define tensor products of orthoalgebras, even though such constructions need not preserve orthocoherence. Our proof, however, will go through also in a more general setting: for instance, as long as orthogonal effects are jointly compatible and mutually exclusive in the above sense, it will apply also to effect algebras (cf. Foulis and Bennett 1994).

[6]That is, excepting any trivially true or trivially false propositions, which have probability 1 or 0 in *all* states.

compatible, it is possible to perform any measurements in the respective $\mathcal{C}_{A_i}$ in any order (or simultaneously), yielding a joint measurement of the propositions A_i.

(c) *No-signalling*: No experimental procedure carried out locally on one system may affect the probabilities for results of experimental procedures carried out locally on another system if there is no interaction between the systems.

We shall discuss these assumptions in detail in Section 3. Their conjunction will now enforce Specker's principle on the subsystems of the composite.

Assume we have a system violating Specker's principle, i.e. a system with a set of n pairwise orthogonal propositions $A_1 \ldots, A_n$ that are not jointly orthogonal. Assume further that this set is minimal, in the sense that any proper subset is jointly orthogonal. Otherwise select a subset that is indeed minimal and redefine n. We may also assume that $n = 3$. Otherwise redefine A_3 as the coarse-graining (disjunction) of $A_3, \ldots, A_n$. If the set $\{A_1, \ldots, A_n\}$ is a minimal non-Specker set, so is the new set $\{A_1, A_2, A_3\}$.

Now take an ensemble of pairs of such a system in a maximally entangled state as specified in assumption (a): assuming that all measurements performed are combinable measurements as specified in assumption (b), we shall now construct a protocol in which Alice can signal to Bob, thus violating (c).

In our version of the fable Sandu and Daniel share pairs of maximally entangled systems with a three-element non-Specker set. Their propositions, however, are not only pairwise exclusive but also pairwise exhaustive. In this special case Sandu adopts the following protocol. If Daniel chooses A and B for his prophecy, Sandu chooses the *third* box C, checks whether it is full or empty, and *depending on the result* (regardless of whether he correctly prophesies it) he chooses which box to measure next. If his C is full, his next box will be empty, which implies that also the corresponding box on Daniel's side will. And if Sandu's box C is empty, his next box will be full, as will the corresponding one on Daniel's side. In this way, Sandu can ensure that Daniel's prophecy is always fulfilled.

This protocol of course violates no-signalling (c), but Alice cannot use it in general because in our case pairs of propositions on the same side need not be exhaustive. Nevertheless, the protocol we now construct generalises the one in the fable: Alice measures one proposition, then depending on the result chooses to combine this with a measurement of one of the other two propositions she can measure, and this will affect the probabilities for a measurement of those other two propositions on Bob's side.

We have the following. Any two propositions A_i and A_j ($i = 1, 2, 3$) can be measured together. In the maximally entangled state, the probability p_i of A_i is strictly positive (because of (a)), and $p_i + p_j \leq 1$ (because the propositions are pairwise orthogonal). Unlike in the fable, there is no requirement that $p_i + p_j = 1$: there may be some third proposition A_k^* outside of the given set such that A_i, A_j and A_k^* are jointly orthogonal.[7]

If a measurement of A_i and A_j is performed by either Alice or Bob, the conditional probabilities for the results are clearly:

$$p(A_i|A_j) = 0 \quad \text{and} \quad p(\neg A_i|A_j) = 1 \tag{1}$$

(because the propositions are pairwise orthogonal) and

$$p(A_i|\neg A_j) = \frac{p(A_i \wedge \neg A_j)}{p(\neg A_j)} = \frac{p(A_i)}{p(\neg A_j)} = \frac{p_i}{1 - p_j} \tag{2}$$

(again because the propositions are pairwise orthogonal), from which we have also

$$p(\neg A_i|\neg A_j) = 1 - \frac{p_i}{1 - p_j} = \frac{1 - p_i - p_j}{1 - p_j} \ . \tag{3}$$

We may assume that these conditional probabilities remain the same on Bob's side if a measurement of A_k is performed by Alice. Otherwise, we already have a protocol by which Alice can signal to Bob. Even so, if Alice measures A_k, the conditional probabilities on Bob's side need only remain the same on average, and may well depend on the outcome of Alice's measurement. In particular:

(*) The conditional probability $p(A_i|\neg A_j)$ on Bob's side for the case in which Alice obtains outcome A_k or $\neg A_k$, respectively, may take on arbitrary values $\alpha_{ij}^k, \beta_{ij}^k \in [0, 1]$, respectively, as long as they satisfy $p_k \alpha_{ij}^k + (1 - p_k)\beta_{ij}^k = \frac{p_i}{1-p_j}$.

Now assume that Alice attempts to affect the probabilities $p(A_1)$ and $p(A_2)$ on Bob's side by first measuring A_3, then choosing to measure either A_1 or A_2. Recall that $0 < p_3 < 1$, so there are four possible cases:

[7] Indeed, perhaps surprisingly the proof goes through even in the case $p_1 + p_2 + p_3 \leq 1$. But some reflection will reveal that the proof does not rely on the incompatibility of A_1, A_2, A_3 in the sense that no joint probability measure for these three quantities exists, but in the sense that the joint probability for A_i and A_j is generally disturbed by an intermediate measurement of A_k.

(I) Alice obtains outcome A_3, then measures A_1;

(II) Alice obtains outcome A_3, then measures A_2;

(III) Alice obtains outcome $\neg A_3$, then measures A_1;

(IV) Alice obtains outcome $\neg A_3$, then measures A_2.

In these four cases, the probabilities for A_1 and A_2 on Bob's side are, respectively:

$$p^{\mathrm{I}}(A_1) = 0 \ , \ \text{because of (1) on Alice's side and (a),} \tag{4a}$$
$$p^{\mathrm{I}}(A_2) = \alpha^3_{21} \ , \ \text{because of (*);} \tag{4b}$$
$$p^{\mathrm{II}}(A_1) = \alpha^3_{12} \ , \ \text{because of (*),} \tag{4c}$$
$$p^{\mathrm{II}}(A_2) = 0 \ , \ \text{because of (1) on Alice's side and (a);} \tag{4d}$$
$$p^{\mathrm{III}}(A_1) = \tfrac{p_1}{1-p_3} \ , \ \text{because of (2) on Alice's side and (a),} \tag{4e}$$
$$p^{\mathrm{III}}(A_2) = \beta^3_{21} \cdot \tfrac{1-p_1-p_3}{1-p_3} \ , \ \text{because of (3) on Alice's side, (a) and (*);} \tag{4f}$$
$$p^{\mathrm{IV}}(A_1) = \beta^3_{12} \cdot \tfrac{1-p_2-p_3}{1-p_3} \ , \ \text{because of (3) on Alice's side, (a) and (*),} \tag{4g}$$
$$p^{\mathrm{IV}}(A_2) = \tfrac{p_2}{1-p_3} \ , \ \text{because of (2) on Alice's side and (a).} \tag{4h}$$

Assume specifically that Alice wishes to minimise the probability of $p(A_1)$. Using (4) we now see that if she obtains A_3, choosing (I) over (II) gives the smaller probability because $0 \leq \alpha^3_{12}$, indeed strictly so unless $\alpha^3_{12} = 0$. Further, if instead she obtains $\neg A_3$ it would be clearly useless for her to choose (III) over (IV), because then the total probability on Bob's side would be $0 \cdot p_3 + \frac{p_1}{1-p_3} \cdot (1-p_3) = p_1$, as if she had not performed any measurement at all. Thus, in order to have a chance of success, she has to choose (IV) over (III). And since she wishes to minimise $p(A_1)$, the worst-case scenario is when β^3_{12} is maximal, i.e. again when $\alpha^3_{12} = 0$ (because of (*)).

In order to see whether Alice succeeds, we thus only need to check whether $\beta^3_{12} \cdot \frac{1-p_2-p_3}{1-p_3} < \frac{p_1}{1-p_3}$, or

$$\beta^3_{12}(1-p_2-p_3) < p_1 \tag{5}$$

for the worst-case scenario $\alpha^3_{12} = 0$, i.e. (by (*)) for the case that

$$\beta^3_{12} = \frac{p_1}{(1-p_2)(1-p_3)} \ . \tag{6}$$

But inserting (6) into (5) it is immediately clear that

$$\frac{p_1(1-p_2-p_3)}{(1-p_2)(1-p_3)} = \frac{p_1(1-p_2-p_3)}{1-p_2-p_3+p_2p_3} < p_1 \tag{7}$$

since $p_2, p_3 > 0$ (again by assumption (a)).

Therefore, given the existence of a non-Specker set and assumptions (a) and (b), we have constructed a protocol by which Alice can signal to Bob, and the assumption of a non-Specker set of propositions is inconsistent with the conjunction of assumptions (a), (b) and (c). *QED.*

3 The assumptions

3.1 Maximal entanglement

Our assumption (a) is meant to generalise the existence of maximally entangled states in quantum mechanics: we have perfect correlations between any pair of matching quantities, even though the quantities on each side are not all jointly compatible. In the case in which we have two sets of 'Specker boxes' and the same two boxes are opened on each side, say A and B, Sandu and Daniel will either both find A full and B empty or they will both find A empty and B full. In the case in which only one of the boxes is the same on the two sides, say Sandu opens C and A and Daniel opens C and B, then either both find C full or both find C empty.

In this second case, assumption (a) on its own does not yet determine what Sandu will find in A and Daniel will find in B, but unless measurements on one side disturb the conditional probabilities on the other (thus violating (c)), it must be that if they both find C full, they must both find their other box empty, and if they find C empty they must find their other box full. Since C and A are compatible and C and B are compatible, if we assume also that opening individual boxes can be combined to yield a measurement of a pair of boxes (assumption (b)), then this will hold irrespective of the order in which boxes are opened. If Sandu opens first A and then C (e.g. finding A empty and C full), and Daniel opens first B then C, Daniel will still find B empty, because he still has to find C full. And if first Sandu opens A (e.g. finding it empty), then Daniel opens B, and then both open C, Daniel will also find B empty, because they both have to find C full.[8]

Given that the boxes form a non-Specker set, our proof shows that assuming (a) and (b), (c) cannot in fact be consistently upheld. But our scenario is not the only one in the literature featuring some kind of 'entangled Specker boxes'. Liang, Spekkens and Wiseman (2011, Section IV) consider a two-party one-query scenario, with two sets of three boxes but only *one* box opened on each side. If the same box is opened on both sides, the results are perfectly correlated; if different boxes are opened, opposite results are obtained. They also consider a one-party two-query scenario in which one opens two boxes in succession on a single side, keeping fixed that one box turns out to be full and one empty (their Section III). If we combine these two scenarios, we get a two-party two-query scenario like in our fable, but the behaviour of the boxes is different. For instance, if C is measured on either or both sides and found empty, then subsequent measurements of A or B on either side will reveal them to be full (like in our fable). But if, for instance, A and B are measured on the two sides, in which case by assumption they produce opposite results, then subsequent measurements of C on the two sides *also* need to produce opposite results (*unlike* in our fable).

In terms of our three assumptions, it is clear that this new scenario *satisfies no-signalling* (c): the average behaviour on each side is simply that of a single set of Specker boxes, irrespective of what measurements are made on the other side. Assumption (b) is also satisfied, because the order in which one opens two individual boxes does not affect the distribution of outcomes for the pair. What is violated (and has to in order for Specker's principle to be violated) is our assumption (a): when A and B are measured on the two sides (yielding opposite results), the results of subsequent measurements of C on the two sides are *not* perfectly correlated, even though measuring BC on one side and CA on the other are both measurements of C.

As suggested to me by Allen Stairs,[9] we can see this extended Liang, Spekkens and Wiseman scenario as exemplifying a different generalisation of maximal entanglement. Instead of requiring that measurements of the same quantity always give the same results on both systems, we can alternatively require that performing (compatible) measurements on one system affects the other system *as if* the measurements

[8] If this seems impossible because the choice of opening C may still lie in the future, recall that the pattern of outcomes already seems impossible if the boxes are opened simultaneously. We postpone a full analysis of the scenario of the fable until Section 3.3.

[9] Email communication, 10 June 2017.

had been performed on that system itself.[10]

This behaviour is somewhat analogous to that in the standard collapse formulation of quantum mechanics (where Specker's principle is of course satisfied). One can imagine that opening the first box (say, A on Alice's side with outcome 'empty') instantaneously 'collapses' the joint state of all the boxes to a 'disentangled' state (which we can write as $|\text{empty}, \text{full}, \text{full}\rangle \otimes |\text{empty}, \text{full}, \text{full}\rangle$). One can then take the resulting state on each side as locally determining the result of the next measurement. This will generally induce some further collapse, but again only locally. For instance, a measurement of B on Bob's side will have outcome 'full' with certainty and collapse the state to $|\text{empty}, \text{full}, \text{full}\rangle \otimes |\text{empty}, \text{full}, \text{empty}\rangle$, so that further measurements of C on the two sides will have indeed outcomes 'full' and 'empty', respectively.[11]

3.2 Combinable measurements

Assumption (b) might also be called 'existence of (non-disturbing) non-maximal measurements', by analogy with the situation in quantum mechanics.[12] We shall now see what kind of violation of Specker's principle we can construct by dropping this assumption.

Note first of all that for Sandu to be able to signal to Daniel, it is crucial that he be able to choose to open A or B *after* having opened C and seen the result. The reason is that the probabilities for Daniel's outcomes are in fact independent of whether Sandu performs a measurement of BC or of CA. It is only the *con-*

[10] It would be interesting to investigate theories in which the two alternative versions of assumption (a) coincide. (See also the remark at the end of Section 3.2.)

[11] One should not take this analogy too far. The toy model is well-behaved only if the maximally entangled state is its only state. If we extend it to include also collapsed states such as $|\text{empty}, \text{full}, \text{full}\rangle \otimes |\text{empty}, \text{full}, \text{full}\rangle$, then B and C will no longer be compatible, because measuring them on the same side produces different outcomes depending on the order of the measurements. For the closest quantum analogue, think of measurements of spin along three directions pairwise at angle α on a pair of spin-1/2 systems in the singlet state. These are pairwise incompatible observables, but their commutator is zero in the singlet state. Note there is a quantum analogue also for the scenario in our fable, namely measurements on a pair of spin-1 systems in a maximally entangled state, where perfect correlations (or anticorrelations) between corresponding one-dimensional projectors are obtained independently of whatever other compatible measurements are performed alongside.

[12] Where it is in fact a non-trivial feature – so much so that it went unnoticed until the now well-known paper by Lüders (1950).

ditional probabilities for Daniel's outcomes given Sandu's that depend on the local measurement context on Sandu's side, if one regards his procedures as measurements of C in the context of also measuring B or also measuring A. (We shall discuss this dependence further in Section 3.3.)

The way the story is told, the ability to make this choice appears trivial. That it is not so can be seen by adapting what is usually taken to be a different version of the same example, namely the 'firefly box'.[13]

Take a triangular box, with translucent sides, and consider the following three experiments. It is dark, and you approach the box from any of the three sides, holding a lantern in your hand: under these conditions you always observe that something starts glowing in either of the two corners you can observe (in each case with probability 1/2). This is the same example as Specker's three boxes in the sense that observations of any two corners are compatible with each other, the results of these observations are always opposite, and of course there is no joint probability distribution with the given marginals. But the way the story is told in the case of the firefly box, C can only be measured *simultaneously* with A or with B. One cannot choose to measure A or B after having observed whether or not there is a glow in corner C of the box.

We can extend also this example to a bipartite system of two Specker-violating firefly boxes.We shall require that they be perfectly correlated in the sense of our assumption (a): say, if corner C is observed both on Alice's firefly box (say together with B) and on Bob's firefly box (say together with A), then C will glow on Alice's side if and only if it glows on Bob's side. This is analogous to the case of our fable, but the only procedures that Alice can implement are *simultaneous* observations of two corners. That is, there are no procedures for measuring A, B and C individually that one can combine to obtain procedures for measuring AB, BC and CA, thus violating assumption (b). By the same token, Alice cannot exploit the context-dependence of the conditional probabilities in order to signal to Bob, and condition (c) is satisfied.

Rather than by analogy to collapse, we can think of firefly boxes by analogy to another well-known approach to the foundations of quantum mechanics, namely de Broglie–Bohm theory. The name 'firefly box' comes about because there is a simple mechanism for explaining the results of your observations, at least in the

[13]My special thanks to Alex Wilce for pressing me on this point. I thank Alex also for the information that the firefly box was apparently invented by Dave Foulis one day that Eugene Wigner was visiting his research group, as an illustration of the kind of work they were doing.

case of a single box. You imagine that somewhere in the box there sits a firefly, which mistakes the glow of your lantern for a potential mating partner, then *moves towards the side you are approaching from*, and starts glowing. Thus, while each experiment corresponds in fact to a random variable on the space of initial positions of the firefly, the three experiments correspond to *three different random variables.* In particular, the event 'C glowing when we approach from CA' is different from the event 'C glowing when we approach from BC'. This means that – even though there is no non-contextual model for the three experiments – the firefly mechanism provides a *contextual hidden variables model* for them. The model is rather like de Broglie–Bohm theory in which it is the position of the particle that plays the role of hidden variable.

The model can also be extended to two firefly boxes that are entangled in the sense of (a), but – again like in de Broglie–Bohm theory – in order to reproduce the desired perfect correlations we need to postulate some action at a distance: Alice's approaching the box from a particular side in general needs to force Bob's firefly to behave in a way it would not otherwise. Even so, this action at a distance *cannot be used for signalling*, as long as Alice remains ignorant of the initial position of her firefly.[14]

This analysis of the firefly box also suggests a possible interpretation for assumption (b). Indeed, what happens here is that we originally misidentify, say, 'C glowing when we approach from CA' and 'C glowing when we approach from BC' as one and the same event 'C glowing' because they are equiprobable. However, once we consider the hidden mechanism behind the experiments, we recognise that they are two different events (which might be distinguished given a non-standard distribution for the hidden variables). That is, even though there might initially appear to be a well-defined quantity C because probabilities for outcomes of measurements of C are independent of the chosen measurement procedure (BC or CA), there is in fact *no* such physical quantity independent of the context of observation. The suggested interpretation of assumption (b) is thus that it is a plausible condition for an observable to be a genuine physical quantity.

[14]For an insightful description of the aspects of de Broglie–Bohm theory that form the basis for the analogy, see Barrett (1999, Chapter 5). See in particular his description of measurements of spin, where the hidden variable is the initial position of the particle, the results of measurements are context-dependent (since they depend on the choice of polarity of the magnetic field, or of direction of the field gradient), the contextuality yields action at a distance in the case of entangled pairs, but we are unable to signal if we have no knowledge of the hidden variables.

Note that unlike in the case of Section 3.1, in the case of the entangled firefly boxes our assumption (a) is now perfectly compatible with the idea that performing a measurement on one firefly box affects both firefly boxes in exactly the same way. That is, we are free to postulate that if Alice measures BC and observes a glow in C, she will again find a glow in C if she then performs CA while she will fail to find a glow in B if she then performs AB – exactly the results Bob will obtain for the same measurements on his own firefly box.

3.3 No-signalling

We finally return to assumption (c) and our fable of the seer of Nineveh. The condition of no-signalling is well known and perfectly transparent, but in the context of our fable its violation is nevertheless unusual. This is because, as in Specker's original version, the mechanism that suggests itself to explain the results is neither instantaneous collapse nor action at a distance, but *retrocausation.* As devised by the father, the trials are meant to test whether the suitors are better prophets than himself (only then would he gladly surrender his daughters to them). But the father is an extremely good prophet indeed, and always prophesies correctly *which two boxes will be opened on each table* (say, AB on both sides, or AB and BC). He then flips a coin to decide where to place the first gem (say, in either A or B on the one side), and he places another gem accordingly on the other table (so as to satisfy (a)). This means that opening the boxes the next day retrocausally influences which boxes are filled the previous night. Assumption (b) is also clearly satisfied: opening two boxes is just the combination of opening the boxes singly, and whether a box is empty or full is a perfectly genuine physical question, albeit one that can be influenced retrocausally. And now, if Daniel makes a prophecy about A and B, once Sandu knows that his C is, say, empty, he knows that by choosing his second box to be, say, B, he will be causing the father to have put a gem inside it the previous night – as well as a gem in the corresponding box on Daniel's side (and similarly in the other cases). *This* is how he manages to freely choose to make Daniel's prophecy turn out to be correct, thereby violating (c).[15]

[15]The model can be generalised further in various ways. Knowing which four boxes are to be opened on each trial allows the father to place gems according to any arbitrary distributions he might choose for any two pairs of boxes. Also, one can restrict or extend the model to the case in which fewer or more than two boxes get opened on at least one side. Clearly, based on the proportion of gems he has placed in, say, the A-boxes in the cases in which two boxes are opened on both sides, the father can place the same proportion of gems in the A-boxes for the case only these

As remarked already in Section 3.2, Sandu's strategy relies on the fact that the conditional probabilities for Daniel's outcomes given Sandu's outcomes will depend on Sandu's local measurement context, if we regard Sandu's procedure as a measurement of C in the context of measuring also A or of measuring also B. The unconditional probabilities for Daniel's outcomes are instead unaffected by Sandu performing these measurements (which are indeed measurements of CA or of BC). But we can equivalently think of Sandu's procedure as a single *generalised* measurement (analogous to a quantum mechanical POVM) that does affect the unconditional probabilities on Daniel's side. Specifically, if A, B and C were some self-adjoint operators with spectral resolutions

$$A = A_+ - A_- \ , \qquad B = B_+ - B_- \qquad \text{and} \qquad C = C_+ - C_- \ , \tag{8}$$

Sandu's procedure (say, if he wants to guarantee that Daniel's box A is empty) would simply be a measurement of the POVM with resolution of the identity

$$C_+A_+C_+ + C_+A_-C_+ + C_-B_+C_- + C_-B_-C_- = \mathbf{1} \tag{9}$$

(and similarly with C_+ and C_- interchanged if he wants to guarantee that Daniel's box A is full).

Although the protocol in the fable involves some classical communication because Sandu wants to know whether he should perform (9) or interchange the roles of C_+ and C_-, this is not essential to ensure signalling. Already the simple fact of performing one of these two generalised measurements affects the probabilities for the outcomes of a possible measurement of A and B on Daniel's side.

It is important to stress that Sandu's ability to signal is not due to a violation of the well-known condition of parameter independence: *given* the complete state of the boxes (i.e. whether they are full or empty), the probability for Daniel finding a gem in a particular box does *not* depend on anything that Sandu might do. In terms of the probabilistic conditions standardly used in discussing distant correlations in quantum mechanics, our version of the fable involves neither violations of outcome independence (as in collapse theories or in our version of the Liang, Spekkens and Wiseman scenario) nor violations of parameter independence (as in de Broglie–Bohm

are opened on one side. And if we understand incompatibility in the sense that observed statistics depend on whether other measurements have been performed, it does not matter once two boxes have been opened on one side what the probability for finding the third one full is or what the correlations between the second and the third are, so the father can place further gems in the third box as he sees fit.

theory or in our entangled firefly boxes). It rather involves violations of *measurement independence*: the distribution of gems in the boxes depends on the choice of the future measurement contexts, and it is this mechanism that mediates the dependence of Daniel's outcomes on Sandu's choice of procedure.[16]

A further striking feature of the entangled Specker boxes in our fable is that – unlike the original case of only one daughter and only one suitor at a time, in which the retrocausal mechanism remains hidden – in the entangled version retrocausation becomes *manifest*. Indeed, we have chosen a case in which it becomes exploitable for signalling, but one can imagine even weirder things happening. Consider Daniel opening box A and getting, say, $+1$ and Sandu opening box B and getting, say, -1. We have seen that, given any possible choices of two boxes on either side, the distribution of gems in the boxes is such as to ensure *both* perfect anticorrelations on each side *and* perfect correlations across the two sides. It follows that Sandu and Daniel *cannot both open box C*, because their boxes would have to be both full and empty. If Sandu and Daniel wanted to thwart the father's prediction, they would not be able to. Something (presumably perfectly natural) would always prevent them from opening box C – just like something always prevents you from killing your grandfather if you travel back to a time before your parents' birth. But there is nothing mysterious about it: Daniel's box A being full and Sandu's box B being empty simply reflects that the father knew already the previous night that Sandu and Daniel in fact would *not* both open box C.[17]

As an alternative assumption for deriving Specker's principle we could thus have chosen a 'free will assumption' or a 'no manifest retrocausation assumption'.[18] However, no-signalling is a very convenient choice, because 'free will' assumptions can

[16]For a detailed treatment of the relation between measurement dependence and signalling in the context of hidden variables theories, see Bacciagaluppi, Hermens and Leegwater (in preparation).

[17]This is also the reason why, somewhat mystifyingly, at the end of the original fable the third box cannot be opened: *not* that the three boxes can never be opened simultaneously or in sequence (incompatibility merely means that opening two of them must disturb the statistics of the third), but that on this particular occasion the father had correctly prophesied that precisely *those two* would be opened. The cause for the failure to prise open the third is quite separate and presumably mundane, like the lid being jammed or the lock being defective.

[18]Note again that the latter is a background assumption in Chiribella and Yuan's (2014) proof of the principle.

be misleading,[19] and retrocausation is a notion that is often misunderstood,[20] while no-signalling is formulated at the level of measurements and correlations.

In conclusion, the three examples we have described in this section – the extended Liang *et al.* scenario, the entangled firefly boxes, and the tale of the two suitors – can be seen as the three canonical examples of violations of Specker's principle obtained by relaxing in turn the three assumptions (a), (b), (c) of our theorem, and provide further insights into the meaning of these assumptions.

4 Popescu and Rohrlich meet Specker's seer

The two suitors in our tale are named after Sandu Popescu and Daniel Rohrlich, who in their 'Quantum nonlocality as an axiom' (Popescu and Rohrlich 1994) investigated the combination of nonlocality with no-signalling. As is well known, Popescu and Rohrlich discovered that nonlocality and no-signalling allow for correlations that are stronger than quantum mechanical ones, and their work inaugurated the field of investigating such super-quantum correlations and the ways of distinguishing them from quantum correlations. Their main example has come to be known as a 'PR

[19] Free will assumptions are just suitable independence assumptions between measurement settings and initial states – in our case that any measurement scenario be compatible with any positioning of the gems. For an excellent analysis along these lines, see Cator and Landsman (2014).

[20] Huw Price is a notable champion of retrocausation, and his Price (1996) provides a lucid discussion of how retrocausation makes sense in the first place and of how it might provide an explanation for the violations of the Bell inequalities. It should be noted that, while the fable of the seer provides an excellent illustration of some of the implications of retrocausation, it remains silent on the mechanism behind the father's prophetic gift. For all we know, the causal link from opening the boxes one day and filling them the previous night might be operating at temporal distance. If that were a generic feature of retrocausal mechanisms, why prefer them to action at *spatial* distance in order to explain nonlocal quantum correlations? But that is just an association evoked by the idea of prophecy. The reason one might expect retrocausal effects in nature lies in the time symmetry (or CPT symmetry) of the laws of physics: for any process that may be given a description in terms of forwards causation, there is a corresponding process that can be given a description in terms of backwards causation. Thus one can conceive of an agent who may be able to exploit such backwards causal processes to affect events in their own past. If that is so, however, then retrocausal mechanisms to be expected on the basis of such time-symmetry arguments will be *just as local as all familiar forwards causal mechanisms*. If they are at all nonlocal in time it will be on the basis of physical processes that are equally nonlocal in the forwards time direction, e.g. non-Markovian stochastic evolutions. (Many thanks to Wayne Myrvold for raising this issue and to Huw Price for discussion.)

box'. One imagines a black box with two pairs of alternative settings a, a' and b, b' such that for each combination $ab, ab', a'b, a'b'$, the box produces a pair of random results ± 1, each individual result having probability $1/2$, and one pair being perfectly correlated and the other three perfectly anti-correlated (or *vice versa*). This in fact yields eight different PR boxes, each of which saturates the $S = 4$ bound in the Bell inequalities.

The scenario by Liang Spekkens and Wiseman (2011) is an example of a realisation of a PR box when one considers opening alternatively boxes A and B on one side and A and C on the other – which yields perfect correlations for the pair AA and perfect anticorrelations for the other three pairs.

The scenario in our fable also provides such a realisation (and the seer must have found it amusing to prophesy that another Sandu and another Daniel were going to give their names to these contraptions). Let Daniel open either A and B, or C and A on his side (label these two choices a and a'), and let Sandu open either A and B, or B and C on his side (label these two choices b and b'). If we interpret Daniel's measurements, respectively, as a measurement of A (in the context of measuring also B) or of C (in the context of measuring also A), and Sandu's measurements, respectively, as a measurement of A (in the context of measuring also B) or of B (in the context of measuring also C), then we get perfect correlations for the pair ab and perfect anticorrelations for the other three pairs of measurements. Thus we have again a realisation of a PR box. (The argument applies equally to entangled firefly boxes, since we have not used assumption (b).)

In fact, we can realise all eight PR boxes, as suggested to me by Jeff Bub.[21] Indeed, by changing the interpretation of Daniel's a to that of a measurement of B (in the context of measuring also A), we get perfect correlations for the pair ab' and perfect anticorrelations for the other three. If we further change the interpretation of Sandu's b' to that of a measurement of C (in the context of measuring also B), we get perfect correlations for the pair $a'b'$ and perfect anticorrelations for the other three. If we change back the interpretation of Daniel's a, we get perfect anticorrelations for the pair $a'b$ and perfect correlations for the other three. Continuing in this fashion by considering Daniel's and Sandu's measurements alternatively as measurements of the first or of the second box (in the context of opening also the other one), we get realisations of all eight Popescu–Rohrlich boxes (each of them twice – because

[21] Private communication, Tarquinia (Italy), May 2017. The argument again applies equally to entangled firefly boxes.

reinterpreting both of Daniel's and Sandu's measurements in fact yields again the same box).

Popescu and Rohrlich's original aim was to explore whether the combination of nonlocality (here implemented by our (a)) with no-signalling (our (c)) could be used to characterise quantum mechanics. Our proof of Specker's principle shows that combining these axioms with a condition on what should count as a genuine physical quantity (our (b)) comes one modest step closer to realising this aim.

Acknowledgements

This paper has had a long gestation, originating on the one side in the pedagogical use of Specker boxes, firefly boxes and related examples in my teaching, various talks, and my handbook article (Bacciagaluppi 2016); on the other in an unfinished plan to discuss the comparison between the uses of Specker's principle in the older and newer literature in joint work with Alex Wilce, to whom my special thanks go for extensive discussion and feedback and for being my habitual consultant on matters quantum logical. I would further like to thank Ehtibar Dzhafarov and Sonja Smets for organising two conferences where this material was presented in preliminary form, as well as the very perceptive audiences there, and Ehti also for the invitation to contribute this paper to the present special issue. Finally, I am very much indebted to Jeff Bub, Huw Price, Allen Stairs and Ken Wharton for discussion of ideas related to this paper and comments on previous versions.

References

Bacciagaluppi, G. (2016), 'Quantum probability: An introduction', in A. Hájek and C. Hitchcock (eds), *The Oxford Handbook of Probability and Philosophy* (Oxford: Oxford University Press), pp. 545–572. Extended version at `http://philsci-archive.pitt.edu/10614/`.

Bacciagaluppi, G., Hermens, R., and Leegwater, G. (in preparation), 'Nonlocality and measurement independence'.

Barrett, J. A. (1999), *The Quantum Mechanics of Minds and Worlds* (Oxford: Oxford University Press).

Cabello, A. (2012), 'Specker's fundamental principle of quantum mechanics', `https://arxiv.org/abs/1212.1756`.

Cabello, A. (2013), 'Simple explanation of the quantum violation of a fundamental inequality', *Physical Review Letters* **110**(6), 060402/1–5.

Cator, E., and Landsman, K. (2014), 'Constraints on determinism: Bell versus Conway–Kochen', *Foundations of Physics* **44**, 781–791.

Chiribella, G., D'Ariano, G. M., and Perinotti, P. (2010), 'Probabilistic theories with purification', *Physical Review A* **81**(6), 062348/1–40.

Chiribella, G. and Yuan, X. (2014), 'Measurement sharpness cuts nonlocality and contextuality in every physical theory', `https://arxiv.org/abs/1404.3348`.

Finch, P. D. (1969), 'On the structure of quantum logic', *Journal of Symbolic Logic* **34**, 275–282. Reprinted in Hooker (1975), pp. 415–425.

Foulis, D. J., and Bennett, M. K. (1994), 'Effect Algebras and Unsharp Quantum Logics', *Foundations of Physics* **24**, 1331–1352.

Foulis, D. J., and Randall, C. H. (1981), 'Empirical logic and tensor products', in H. Neumann (ed.), *Interpretations and foundations of quantum mechanics* (Mannheim: Bibliographisches Institut), pp. 9–20.

Fritz, T., Sainz, A. B., Augusiak, R., Bohr Brask, J., Chaves, R., Leverrier, A., and Acín, A. (2013), 'Local orthogonality as a multipartite principle for quantum correlations', *Nature Communications* **4**, 2263/1–7.

Gudder, S. P. (1972), 'Partial algebraic structures associated with orthomodular posets', *Pacific Journal of Mathematics* **41**, 717–730.

Hardegree, G. M., and Frazer, P. J. (1981), 'Charting the labyrinth of quantum logics: A progress report', in E. Beltrametti and B. van Fraassen (eds), *Current Issues in Quantum Logic* (New York: Plenum Press), pp. 53–76.

Henson, J. (2012), 'Quantum contextuality from a simple principle?', `https://arxiv.org/abs/1210.5978`.

Hooker, C. A. (1975), *The Logico-Algebraic Approach to Quantum Mechanics*, Vol. 1 (Dordrecht: Reidel).

Hughes, R. I. G. (1989), *The Structure and Interpretation of Quantum Mechanics*, (Cambridge, Mass.: Harvard University Press).

Kochen, S., and Specker, E. P. (1965), 'Logical structures arising in quantum theory', in L. Addison, L. Henkin and A. Tarski (eds), *The Theory of Models* (Amsterdam: North-Holland), pp. 177–189. Reprinted in Hooker (1975), pp. 263–276.

Liang, Y. C., Spekkens, R. W., and Wiseman, H. M. (2011), 'Specker's parable of the overprotective seer: A road to contextuality, nonlocality and complementarity', *Physics Reports* **506**(1–2), 1–39.

Lüders, G. (1950), 'Über die Zustandsänderung durch den Meßprozeß', *Annalen der Physik* **443**(5–8), 322–328.

Popescu, S., and Rohrlich, D. (1994), 'Quantum nonlocality as an axiom', *Foundations of Physics* **24**(3), 379–385.

Price, H. (1996), *Time's Arrow and Archimedes' Point: New directions for the physics of time* (Oxford: Oxford University Press).

Seevinck, M. P. (2011), 'The logic of non-simultaneously decidable propositions', `http://arxiv.org/abs/1103.4537`. Translation of Specker (1960).

Specker, E. P. (1960), 'Die Logik nicht gleichzeitig entscheidbarer Aussagen', *Dialectica* **14**, 239–246. Translated as Stairs (1975) and as Seevinck (2011).

Stairs, A. (1975), 'The logic of propositions which are not simultaneously decidable', in Hooker (1975), pp 135–140. Translation of Specker (1960).