

Why Entanglement?

Huw Price* and Ken Wharton†

12th January 2023

Abstract

In this piece, written for a general audience, we propose a mechanism for quantum entanglement. The key ingredient is the familiar statistical phenomenon of collider bias, or Berkson's bias. In the language of causal models, a collider is a variable causally influenced by two or more other variables. Conditioning on a collider typically produces non-causal associations between its contributing causes, even if they are actually independent. It is easy to show that this phenomenon can produce associations analogous to Bell correlations, in suitable post-selected ensembles. It is also straightforward that such collider artefacts may become real connections, resembling causality, if a collider is 'constrained' (e.g., by a future boundary condition). We consider the time-reversed analogues of these points in the context of retrocausal models of QM. Retrocausality yields a collider at the source of an EPR-Bell particle pair, and in this case constraint of the collider is possible by normal methods of experimental preparation. It follows that connections resembling causality may be expected to emerge across such colliders, from one branch of the experiment to the other. Our hypothesis is that this constrained retrocausal collider bias is the origin of entanglement. The piece is based on a suggestion we first made in arXiv:2101.05370v4 [quant-ph].

Keywords: Entanglement, collider bias, retrocausality

*University of Bonn and Trinity College, Cambridge, UK; email hp331@cam.ac.uk.

†Department of Physics and Astronomy, San José State University, San José, CA 95192-0106, USA; email kenneth.wharton@sjsu.edu.

1 Introduction

Quantum entanglement has been in the news recently. The physicists Alain Aspect, John Clauser and Anton Zeilinger won the 2022 Nobel Prize for, as the citation puts it, ‘experiments with entangled photons, establishing the violation of Bell inequalities and pioneering quantum information science.’

Within the foundations of physics community, the prize has also been welcomed as a tribute to John Stewart Bell (1928–1990), the Irish physicist who argued in the 1960s that quantum mechanics implies that the world is unavoidably ‘non-local’ – that Einstein’s ‘spooky action at a distance’ is unavoidable, if quantum theory is correct. Clauser and Aspect won their Nobel Prizes for pioneering work in the 1970s and 1980s to test the relevant predictions of quantum mechanics. Such tests are now called *Bell experiments*. Many increasingly sophisticated versions, including Zeilinger’s ‘loophole-free’ experiments, have confirmed that the world does indeed behave as quantum theory predicts.

But *why* does the world behave this way? Bell experiments rely on quantum entanglement, which was first identified by Erwin Schrödinger in 1935. He called it ‘not *one* but rather *the* characteristic trait of quantum mechanics.’ [Schrödinger 1935a] More recently, it has been described as ‘the essential fact of quantum mechanics’ [Susskind & Friedman 2014], and ‘perhaps its weirdest feature’ [Weinberg 2013]. The work of Clauser, Aspect, Zeilinger and others confirms the reality of entanglement – as Aspect himself put it, in his [speech](#) at the Nobel Prize banquet, ‘entanglement is confirmed in its strangest aspects’ – but it doesn’t tell us what it is, or where it comes from.

Our [research](#) [Price & Wharton 2021] suggests a surprisingly simple answer. Entanglement may rest on a familiar statistical phenomenon known as *collider bias*. In the light of collider bias, we think, entanglement is not really mysterious at all. It is what we might have expected, if we’d taken seriously the time-symmetry of the microworld.

Our proposal needs just three other ingredients, as well as collider bias – as we said, it is a simple recipe. All these ingredients are available off the shelf (though admittedly, in one case, from a niche corner of the shelf). But as far as we know, it has not been noticed that they can be combined in this way, to throw new light on the central puzzle of the quantum world.

2 What is collider bias?

Let's start with the main ingredient. Collider bias was first described by Joseph Berkson (1899–1982), a Mayo Clinic physicist, physician and statistician. In the 1940s Berkson noted an important source of error in statistical reasoning used in medicine. In some circumstances, the selection of a sample of patients produces misleading correlations between their medical conditions. Taken at face value, these correlations can suggest that one condition prevents another. Berkson pointed out that these apparent causal connections may not be real. They may be artefacts of the way the sample has been selected [Berkson 1946].

Simplifying Berkson's own example, imagine that all the patients admitted to Ward C have similar symptoms, caused by one of two rare infections, Virus A or Virus B. Ward C specialises in treating those symptoms, so all its patients have at least one of these diseases. Some may have both, but everyone on the ward who doesn't have Virus A is certain to have Virus B, and vice versa.

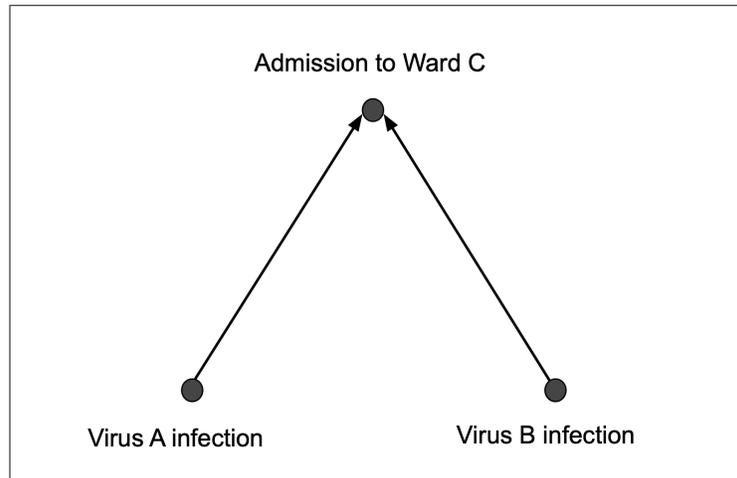


Figure 1. A simple collider

Berkson's point was that this isn't evidence that avoiding one virus leads to infection with the other one. The patients on Ward C are a very biased sample. In the general population, having a vaccine for Virus A won't make you more likely to catch Virus B.

This means that if a patient on Ward C with Virus A says to himself, “I’m on Ward C, so if I hadn’t caught Virus A I would have caught Virus B”, then he’s making a mistake. If he hadn’t caught Virus A then (most likely) he wouldn’t have either virus, and he wouldn’t have been admitted to the ward.

This statistical effect is now called Berkson’s bias, or *collider bias*. The term ‘collider’ comes from causal modelling, the science of inferring causes from statistical data. Causal modellers use diagrams called Directed Acyclic Graphs (DAGs), made up of nodes linked by arrows. The nodes represent events or states of affairs, and the arrows represent causal connections between those events. When an event has two independent contributing causes, like being a patient on Ward C, it is shown in a DAG as a node where two arrows ‘collide’ – see Figure 1.

If we just look at a sample of cases in which the event at a collider happens, we’ll often see a correlation between the two independent causes. It may look like these causes are influencing one another, but they are not. It is a selection artefact, as causal modellers say. That’s collider bias. The correlation stems from the way in which the event at the collider depends on the two causes – in our simple example it needed one cause or the other.

3 Rock-paper-scissors

We want to take collider bias in the direction of physics – ultimately, in the direction of the kind of Bell experiments for which Clauser, Aspect and Zeilinger won their Nobel Prizes. We want to propose an explanation for what may be going on in those experiments, and other cases of quantum entanglement.

We’ll get there via a series of toy examples. For the first of them, imagine that two physicists Alice and Bob play rock-paper-scissors, sending their calls to a third observer, Charlie. Charlie makes a list of the results: Alice wins, Bob wins, or it’s a draw.

Suppose that Charlie likes Alice and dislikes Bob. He throws away most of the results when Bob wins. In the remaining ‘official’ results Alice wins a lot more often than Bob. The correlation looks the way it would if Alice actually had some influence over Bob’s choice – as though Alice choosing scissors makes it a lot less likely that Bob will choose rock, and so on. If Alice and Bob are far apart, this could look like spooky action at a distance.

But there's no real Alice-to-Bob causation involved. It is just collider bias at work. The event at the collider – whether Charlie retains or throws away the result – is influenced both by Alice's choice and by Bob's choice, giving us the same kind of converging arrows as in Figure 1.

Suppose that in a particular round of the game Alice chooses paper and Bob chooses rock. As in the medical case, Alice would be making a mistake if she says, "If I had chosen scissors instead, Bob would probably not have chosen rock." The right thing for her to say is, "If I had chosen scissors, then Charlie would probably have discarded the result – so my choice didn't make any difference to Bob's choice."

4 Constrained colliders

Now to our second ingredient. It is the least familiar of all, though it, too, is already on the shelf, if you know where to look. In the game just described, Charlie could only favour Alice by discarding some results. Let's see what happens if we rig the game in Alice's favour, *without throwing any results away*.

In our world this isn't going to happen naturally, so for now, let's imagine it happening supernaturally. Suppose God also likes Alice more than Bob, so he tweaks reality to give her an advantage. Perhaps he arranges things so she never loses when she plays the game on Sundays. How does God do it? It doesn't matter for our story, which doesn't need to be realistic at this point, but here's one possibility. In a deterministic universe everything that happens is determined by the initial conditions at the very beginning of time. If God gets to choose the initial conditions, and – relying on a Laplacian calculation or simply Divine foreknowledge – knows exactly what follows from them, then he can simply choose the initial conditions so that Alice never loses on Sundays.

Readers who prefer a God-free version could imagine that Alice and Bob live in a simulation, and that the superintelligence (AGI) that runs the simulation favours Alice on Sundays. Some serious thinkers have suggested that we ourselves may live in a simulation, so it would be hasty to say that this version is inconceivable. But again, our example doesn't need to be realistic at this point. Later in our argument, when realism matters, we won't have to rely on God or simulations.

To invent some terminology, let's say that God (or the AGI) *constrains the*

collider – just on Sundays, in this version of the story. To see what difference this makes, think again about a round of the game where Alice chooses paper and Bob chooses rock. Is Alice still making a mistake if she says, “If I had chosen scissors instead, Bob would not have chosen rock”? It now depends what day of the week it is. This is still a mistake Monday through Saturday. On those days, the right thing for Alice to say is, “If I had chosen scissors, Bob would still have chosen rock (and I would have lost).” But Sunday is different. On Sunday Alice never loses, so if she had chosen scissors, Bob could not have chosen rock.

Let’s suppose that Alice knows that the game works this way. Perhaps she figured it out after years of experiments, and now makes a comfortable living as a gambler, working one day a week. From her point of view, it looks like she can *control* Bob’s choices, to some extent (and only on Sundays). By choosing scissors she can *prevent* Bob from choosing rock, and so on.

With a *constrained* collider, then, we would have something that looks a lot like real causality across the collider, from one of the pair of incoming causes to the other. True, it would be a very strange kind of causality. For one thing, it would work the other way, too, from Bob to Alice (though less happily, from his point of view). By choosing rock on a Sunday Bob could prevent Alice from choosing scissors, and so on.

For our purposes, it isn’t going to matter whether this would be real causality, or even whether the question makes sense, in this case. If we press too hard on a toy example like this, it is liable to fall apart at the seams. Could we still speak of both Alice and Bob as making free choices, for example, if the choices are linked in this way?

All we need from the example is the following lesson. If there were cases in nature in which something restricted the options at a collider, we should expect to find a new kind of dependence between the normally independent causes that feed into that collider. To keep our terminology non-committal on the question whether it would count as causality, we’ll call this new kind of relation *connection across a constrained collider* (CCC).

CCC is our second ingredient, and certainly the least familiar one, for most readers. But it, too, is already on the shelf, in the sense that there’s at least one place in physics where it has actually been proposed. It is the key to a suggestion by Maldacena and Horowitz [[Horowitz & Maldacena 2004](#)] for solving the so-called black hole information paradox (made famous by Stephen Hawking’s [bet](#) with Kip Thorne and John Preskill).

The Maldacena-Horowitz hypothesis relies on the proposal that special ‘future boundary conditions’ inside black holes constrain a collider (in our terminology) at that point. Maldacena and Horowitz suggest that this creates a zigzag causal path through time, along which information can escape from a black hole.

Discussing the Maldacena-Horowitz hypothesis recently, the Cambridge physicist Malcolm Perry says

[t]he interior of the black hole is therefore a strange place where one’s classical notions of causality . . . are violated. This does not matter as long as outside the black hole such pathologies do not bother us. [Perry 2021, 9]

As we’ll explain, our proposal is going to be that such pathologies are actually extremely common, if you know where to look. In the other direction of time, they are the basis of quantum entanglement – and they don’t need black holes.

So far, then, we have two ingredients on the table: collider bias itself, and CCC – connection across a constrained collider. Before we introduce the two remaining ingredients, let’s get a bit closer to the physics of the quantum world.

5 From rock-paper-scissors to Bell experiments

As we noted at the beginning, the recent Nobel Prizes were awarded for Bell experiments. These confirmed the strange correlations, predicted by quantum theory, which Bell took to show that the quantum world is unavoidably non-local. Given that these so-called Bell correlations were important enough to win Nobel Prizes, readers may be surprised to learn that they can easily be reproduced in a version of our rock-paper-scissors game. The only change we need is to have Alice and Bob each flip a coin before they make their choice.

In this variant – call it *quantum* rock-paper-scissors (QRPS) – Alice and Bob each send two pieces of information to Charlie: their choice of rock-paper-scissors, and the result of their coin flip. So Charlie gets four values, two choices and two coin outcomes. This is precisely the same amount of information generated in each run of a Bell experiment. In that

context, these values are called the two *measurement settings* and the two *measurement outcomes*. The Bell correlations are particular relationships between these four values, in long lists of experimental results. (In one kind of Bell experiment, for example, they specify among other things that whenever the two settings are the same, the two outcomes must be different.)

In QRPS it is very easy for Charlie to set up a filter, keeping some results and throwing away others, to make sure that the set of results he keeps satisfies the Bell correlations. By using the right filter, Charlie can ensure that the selected results look *exactly* like the data generated in the familiar kind of Bell experiment of Clauser, Aspect and Zeilinger.¹

Of course, this doesn't mean that there is any strange non-locality in QRPS. As in the earlier version, the correlations are a selection artefact, a result of collider bias. But since the results are now a perfect copy of real Bell experiments, it is worth paying careful attention to differences between the two.

The most obvious difference is that quantum rock-paper-scissors and real Bell experiments look like time-reversed versions of each other. In QRPS Alice and Bob send their choices *to* Charlie, later in time. In a spacetime diagram with time running up the vertical axis, the structure looks like an upside-down 'V' – see the left hand side of Figure 2. We'll say that cases like this are ' \wedge -shaped'. In real Bell experiments, Alice and Bob receive their particles *from* the source, which is earlier in time. So the structure looks like \vee , as in the right hand side of Figure 2 – we'll say that they are ' \vee -shaped'.

In a moment we're going to introduce a \vee -shaped version of QRPS, to eliminate this difference. But first let's summarise what we learned from the \wedge -shaped case. We saw that it is easy for Charlie to set up a filter, to make sure that the results he keeps satisfy the Bell correlations. But we don't need any non-locality between Alice and Bob, to explain what's going on. The correlations are simply collider bias. As in the original rock-paper-scissors case (without constraint at the collider) they are simply a selection artefact.

¹See [Bacciagaluppi & Hermens 2021] for the same point in the real QM context.

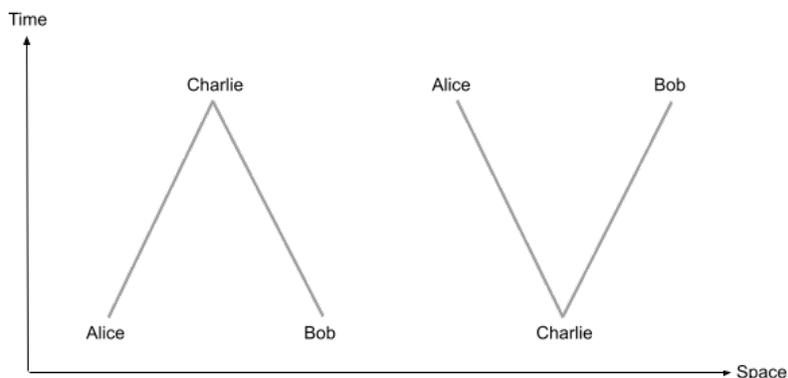


Figure 2. \wedge -shaped and ∇ -shaped experiments

We could reintroduce God or an AGI at this point, to add a constrained collider to QRPS. There would be one interesting difference from the original game. In that case, we saw that the effect of the constraint was to give Alice and Bob control over each other's choices, making it hard to maintain that they both had freedom to choose. In QRPS, as in the analogous real Bell experiments, that problem goes away: Alice and Bob each get some influence over the result of the other's coin toss, but we can still treat both of their own choices as completely free.

That difference aside, a constrained version of QRPS would be as unrealistic as for the original game, and wouldn't tell us anything new. So let's turn instead to a ∇ -shaped version of QRPS, where realistic constrained colliders will be much easier to find.

6 Flipping the game

As we said, QRPS looks like a time-reversed version of a Bell experiment, \wedge -shaped rather than ∇ -shaped. Instead of two particles *leaving* a common source and *going to* separate observers, it's the other way round. The information travels *from* Alice and Bob, *to* Charlie at the common point in the future.

Can we flip this quantum rock-paper-scissors, to make it ∇ -shaped not \wedge -shaped? It might look easy. We can have Charlie toss the two coins and send them to Alice and Bob, so that the results (heads or tails) become Alice and Bob's measurement outcomes. But if that's all we do, Charlie won't know what choices Alice and Bob are going to make when he sends out

the coins. That means there's no way for him to put bias into the results, in the way that he could in the \wedge -shaped case. There's no way that Charlie can produce the Bell correlations, in other words.

But suppose we let Charlie *know in advance* what choices Alice and Bob are going to make – we give him a crystal ball, say. Then it is very easy for him to manage the coins so that the net results, gathered over many plays of the game, satisfy the Bell correlations. The trick is for Charlie to toss one coin, and then choose the result for the other coin based on a rule that takes into account Alice and Bob's future choices.

So this game, too, generates the same kind of Bell correlations as the famous experiments of Clauser, Aspect, and Zeilinger. Let's ask the same question we did about \wedge -shaped QRPS. Does the new \vee -shaped version involve some kind of spooky action at a distance from Alice to Bob, and vice versa?

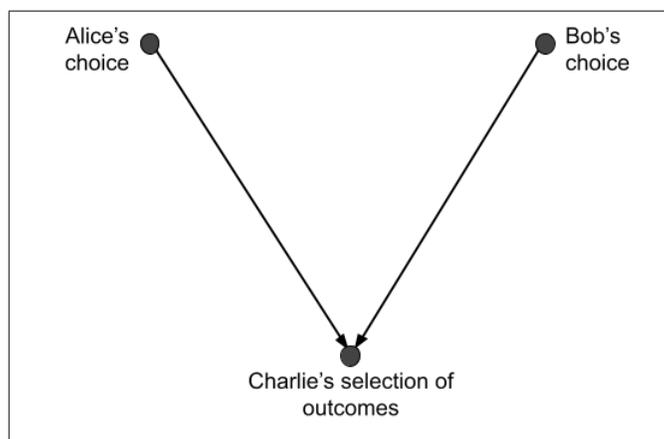


Figure 3. A past collider

We hope that readers will be inclined to say 'No' to this question. After all, the basic causal structure of the new \vee -shaped version is something like Figure 3. Thanks to Charlie's crystal ball, and the rule he uses for selecting outcomes, Alice's and Bob's choices both influence what outcomes Charlie produces, in every case. This means that Charlie's selection procedure is a collider, and we have to be on our guard for collider bias.

For this reason, attentive readers might suspect that collider bias plays the same role in explaining the results the new \vee -shaped QRPS as it did in the \wedge -shaped case. But there's one very big difference between these two cases – which brings us to our third ingredient.

7 Initial control

In the \wedge -shaped version of QRPS, Charlie had to apply a filter, and throw away results he didn't want. But in the \vee -shaped case, he gets to *choose* the results, in the light of what he learns from the crystal ball and his (single) coin toss. *He doesn't have to throw anything away.* In this case, then, Charlie himself can constrain the collider. All he needs is an ordinary ability we take for granted, to control the initial conditions of an experiment.

Perhaps we shouldn't take this for granted. It is actually a remarkable ability, one that depends on the fact that we live in a place in which abundant low-entropy energy can be harnessed by creatures like us. But by ordinary standards, there's nothing surprising about it. We have much more control over the *initial* conditions of experiments than over their *final* conditions. It's easy to arrange the balls on a pool table into precise positions *before* the initial break, but virtually impossible to play the game so that they all *end up* in those positions.

Let's call this familiar fact *initial control*. It is the third ingredient in our recipe. Looking ahead a bit (so to speak), the final ingredient is going to be something that does the job of the crystal balls, in giving Charlie access to information about future choices by Alice and Bob. We're going to find that on the shelf already in the quantum world, under the name *retrocausality*. But before we go there, let's summarise what we learn from the new \vee -shaped QRPS.

Thanks to the crystal balls, we have the causal structure shown in Figure 3, with a collider in the past, where Charlie chooses the outputs. Thanks to initial control, it is easy to make it a constrained collider, so that Charlie doesn't need to throw any results away. And this means that there can be connection across the constrained collider (CCC), from Alice to Bob, and vice versa.

In other words, the combination of the collider structure in Figure 3 and the constraint provided by initial control gives us CCC. If we are happy to use causal language, we can say that it gives us the kind of zigzag causal connection shown in Figure 4. There's also a zigzag path from Bob's choice to Alice's outcome, of course. (Why do we call it a Parisian Zigzag? More on that in a moment.)

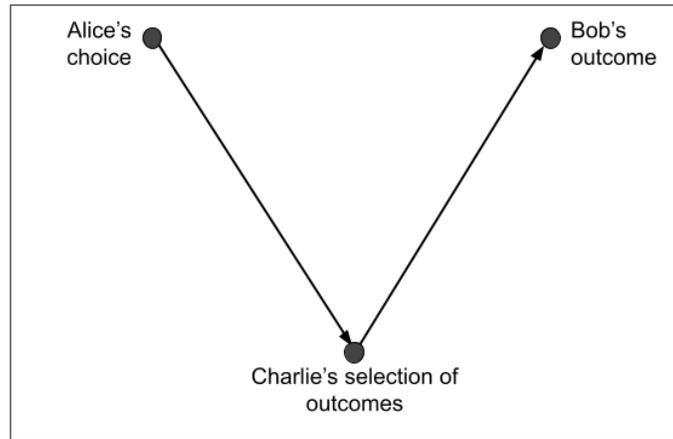


Figure 4. The Parisian Zigzag

In the light of this, let's go back to this question. Does \vee -shaped QRPS involve some kind of spooky action at a distance from Alice to Bob, and vice versa? At this point we need to be careful about what we mean by action at a distance. As we have just seen, there is indeed some influence, or connection, from Alice to Bob, and vice versa. Since they are at a distance from each other, and a direct connection might need to be faster than light, we might still want to call it action at a distance, or nonlocality. (One of the recent Nobel laureates once told us that he thought such a zigzag should still count as a nonlocal effect.)

However, the connection between Alice and Bob is indirect, and depends entirely on processes which don't themselves require anything faster than light. So whatever we call it, it doesn't have the relativity-challenging character normally associated with so-called spooky action at a distance in QM. And it is not very mysterious: we know exactly what it is, namely, connection across a constrained collider.

The crystal balls were pretty mysterious, of course, but once we gave ourselves those, the explanation of the connection between Alice and Bob is straightforward. Imagine if something like this could explain the results of real Bell experiments – that would be a nail in the coffin of the quantum spooks!

8 Retrocausality

Well, let's see. We need our final ingredient. In \vee -shaped QRPS, we gave Charlie a crystal ball, to allow causation to work backwards – in other words, to allow Alice and Bob's choices to feed into the algorithm Charlie uses to select the measurement outcomes. In the real world, of course, we don't find magical crystal balls on any actual shelf.

In the quantum world, however, retrocausality is a familiar hypothesis. In that sense, it is certainly available off the shelf. It was first proposed in the late 1940s by the Parisian physicist Olivier Costa de Beauregard [[Costa de Beauregard 1953](#)], who suggested that in the quantum world causal influence might follow a zigzag path, as in Figure 4 – that's why we called it the Parisian Zigzag.

Retrocausality remained a niche idea for many years, though it has long had some distinguished proponents. In the 1950s one of them, at least briefly, was the British physicist Dennis Sciama [[Sciama 1958](#)], who taught an astonishing generation of physicists, including Stephen Hawking. Sir Roger Penrose, himself a recent Nobel laureate, has long been sympathetic to the idea. There's a story from the 1990s of Penrose drawing a zigzag at a quantum workshop at the Royal Society in London, and joking 'I can get away with proposing this kind of thing, because I'm already a Fellow here.' (Now that he has a Nobel Prize it is even easier, presumably!)² Another famous long-term proponent is the Israeli physicist Yakir Aharonov, whose interest dates from a well-known 1964 paper [[Aharonov, Bergmann & Lebowitz 1964](#)].

More recently, we ourselves have written in *Aeon* and [elsewhere](#) about the advantages of retrocausal approaches to QM, both in avoiding action at a distance, and in respecting time-symmetry. We argued that quantum theory provides new reasons to think that microscopic time-symmetry requires retrocausality [[Price & Wharton 2016](#), [Wharton & Price 2016](#)].

Recent popular discussions of the retrocausal approach to QM may be found [here](#), [here](#) and [here](#), for example.³ It is now a sufficiently familiar proposal that no adequate survey of the puzzles of quantum theory can afford to ignore it. Some commentators on the recent Nobel Prize announcement

²For a recent version of Penrose's views on this topic, see [[Penrose 2023](#)].

³See [[Merali 2022](#), [Becker 2018](#), [Zyga 2017](#)]. For specialist audiences, recent discussions include the following: [[Price & Wharton 2015](#), [Leifer & Pusey \(2017\)](#), [Leifer \(2017\)](#), [Friedrich & Evans 2019](#), [Wharton & Argaman 2020](#), [Norsen & Price 2021](#), [Adlam 2022](#)].

got themselves into trouble for doing so [[Hance & Hossenfelder 2022](#)].

So retrocausality in QM is a well-known idea, and has well-known points in its favour. However, an additional striking advantage seems to have been overlooked. Retrocausality suggests a simple mechanism for ‘*the characteristic trait of quantum mechanics*’ (Schrödinger), ‘*its weirdest feature*’ (Weinberg) – in other words, for the strange connections between separated systems called quantum entanglement.

Starting with retrocausality, our proposal goes like this, in four easy steps. We’ve highlighted the use of our four ingredients.

1. **Retrocausality** automatically introduces colliders into Bell experiments, at the point where the two particles are produced.
2. That’s interesting because colliders produce **collider bias** and causal artefacts – correlations that look like they involve causation, but really don’t.
3. But **constraining a collider** can turn a causal artefact into a real connection across the collider.
4. In the case of colliders in the past, as in Figure 3, constraint is easy. It just follows from normal **initial control** of experiments.

Taken together, these steps suggest a simple explanation for the Parisian Zigzag, and the strange kind of non-local connections in the quantum world, revealed by Bell’s arguments. It is *connection across constrained colliders*, where the colliders result from retrocausality and the constraints from ordinary initial control of experimental setups.

Of course, more work is needed to show that this simple mechanism can actually explain quantum entanglement. We don’t mean to claim that it is a trivial step from \vee -shaped QRPS to real Bell experiments. What we do claim, and what we take to be demonstrated by \vee -shaped QRPS, is that the combination of retrocausality and initial control can give rise to a connection between separated systems that looks very similar to entanglement. In our view, this is such a striking fact – and entanglement is otherwise such a strange and mysterious beast – that we propose the following hypothesis:

Hypothesis (E=CCC): Quantum entanglement is retrocausal collider bias, constrained by initial control.

If this hypothesis turns out to be true, then in place of spooky action at a distance, we'll get Costa de Beauregard's zigzag connections – which, as he always emphasised, are much easier to reconcile with relativity – but now explained as CCC. It will still be true that quantum theory gives us a new kind of connection between the properties of distant systems. The experiments of Clauser, Aspect and Zeilinger provide very convincing evidence that quantum entanglement is a real phenomenon. But it would no longer look mysterious. On the contrary, any world that combines retrocausality and initial control would be expected to look like this.

In earlier work, as we noted above, [we](#) and [others](#) have argued that time-symmetry requires retrocausality, once the world becomes quantised [[Price & Wharton 2016](#), [Zyga 2017](#)].⁴ The same line of argument now seems to lead to entanglement, once initial control is added to the picture. That's what we meant at the beginning, when we said that entanglement is what we might have expected, if we'd taken seriously the time-symmetry of the microworld.

9 Avoiding causal loops and signalling

Finally, a note for readers who are worried that the cure is worse than the disease – that retrocausality opens the door to a menagerie of paradoxes and problems. The note says: Well spotted! As we described \vee -shaped QRPS, with the crystal balls, you are absolutely right. For one thing, the crystal balls give Charlie options much like those of the famous time-traveller, encountering his own grandfather long before his parents met. What's to stop Charlie interfering with the course of history, say by bribing Bob to make a different choice than the one shown in the crystal ball? (In the causal loop literature, this is called 'bilking'.)

Also – less dramatic, maybe, but especially interesting in comparison to QM – the crystal balls allow Alice and Bob to send messages to Charlie, and hence potentially, with his help, to signal to each other. This isn't possible in real Bell experiments, where Alice and Bob can't signal to each other, despite having some influence on each other's measurement outcomes. So isn't this bad news for retrocausality?

These are good objections, but it is easy to modify the \vee -shaped QRPS

⁴See also [[Price 2012](#), [Price & Wharton 2015](#), [Leifer & Pusey \(2017\)](#), [Leifer \(2017\)](#)].

game to avoid them. What we need to do is to split Charlie's functions into two parts. Most of what he does gets replaced by a simple algorithm, inside a black box, that takes in information about the two future measurement settings, and spits out the two measurement outcomes. Charlie himself can't see inside the black box, and doesn't have access to the future settings. But he still has a vital job to do. The box has a knob on the front, with a small number of options. Charlie controls that knob, and if he wants the device to produce the Bell correlations, he needs to choose the right option. In the terminology of QM, that's called 'preparing the initial state'. If that's all that Charlie does, and the quantum black box takes care of the rest, the door to the menagerie is closed. Alice and Bob can no longer signal to Charlie, or to each other. Everything works as in orthodox QM, except that we now have the prospect of an explanation of entanglement.

This means that if nature wants retrocausality without retrosignalling (and without the paradoxes that retrosignalling would lead to), it is going to need black boxes – places in nature where observers like Charlie can't see the whole story. In normal circumstances, such black boxes would seem like another kind of magic. Charlie is a clever guy, after all. What's to stop him taking a peek inside any kind of box?

But in the quantum case, many readers will already know the answer to this question. What's to stop Charlie taking a peek is Werner Heisenberg, or more precisely his famous Uncertainty Principle. Ever since Heisenberg, quantum theory has been built on the idea that there are new limits to what it is possible to know about physical reality. One of the central questions is whether this is just a restriction on our knowledge of reality, or whether reality itself is somehow fuzzy. As Schrödinger put it in 1935, after describing his famous Cat Experiment: 'There is a difference between a shaky or out-of-focus photograph and a snapshot of clouds and fog banks.' [Schrödinger 1935b]

The Cat Experiment was supposed to support the out-of-focus photograph option, the view that the Uncertainty Principle is just a restriction on our *knowledge* of reality. Schrödinger thought it was obvious that the cat couldn't actually be somehow neither alive nor dead. Like Einstein, Schrödinger favoured the view that the quantum description is incomplete, and that reality contains further details, hidden behind Heisenberg's veil.

In the decades since 1935, most physicists who care about these issues have concluded that Einstein and Schrödinger were wrong. Bell's Theorem, together with the quantum predictions being confirmed by Clauser, Aspect,

Zeilinger and many others, has often been interpreted as showing that the spooky action at a distance which Einstein hoped to avoid with additional ‘hidden variables’, is an inevitable part of the quantum world.

Retrocausality is already the most interesting challenge to that view. By taking the first option on Schrödinger’s list – by treating QM as an unavoidably fuzzy picture of a sharper reality – it can allow the kind of quantum black boxes needed to avoid retrosignalling and paradoxes. How satisfying, then, if it also explains the other thing that Schrödinger put his finger on in 1935, when he invented the term ‘entanglement’, and called it ‘the characteristic trait of quantum mechanics.’⁵

References

- [Adlam 2022] Adlam, E. Two Roads to Retrocausality. arXiv:2201.12934 [physics.hist-ph]. doi.org/10.48550/arXiv.2201.12934
- [Aharonov, Bergmann & Lebowitz 1964] Aharonov Y., Bergmann, P. & Lebowitz, J. Time symmetry in the quantum process of measurement. *Physical Review*, 134, B1410. doi:10. 1103/PhysRev.134.B1410.
- [Bacciagaluppi & Hermens 2021] Bacciagaluppi, G. & Hermens, R. Bell Inequality Violation and Relativity of Pre- and Postselection. arXiv:2002.03935
- [Berkson 1946] Berkson, J. Limitations of the application of fourfold table analysis to hospital data. *Biometrics Bulletin*, 2, 47–53.
- [Becker 2018] Becker, Adam. Quantum time machine: How the future can change what happens now. *New Scientist*, 14 February 2018. <https://www.newscientist.com/article/mg23731652-800-quantum-time-machine-how-the-future-can-change-what-happens-now/>
- [Costa de Beauregard 1953] Costa de Beauregard, O. Mécanique quantique. *Comptes Rendus Académie des Sciences* 236, 1632–34 (1953).

⁵We are grateful to Emily Adlam, Guido Bacciagaluppi and Nathan Argaman for comments on previous versions.

- [Friedrich & Evans 2019] Friederich, S. & Evans, P. Retrocausality in quantum mechanics. In *The Stanford Encyclopedia of Philosophy* (Summer 2019 Edition); Zalta, E., Ed. <http://plato.stanford.edu/archives/sum2019/entries/qm-retrocausality/>
- [Hance & Hossenfelder 2022] Hance, J.R. & Hossenfelder, S. Bell’s theorem allows local theories of quantum mechanics. *Nature Physics* 18, 1382. doi.org/10.1038/s41567-022-01831-5
- [Horowitz & Maldacena 2004] Horowitz, G. and Maldacena, J. The black hole final state, *JHEP* 0402:008 (2004). arXiv:hep-th/0310281
- [Leifer & Pusey (2017)] Leifer, M. & Pusey, M. Is a time symmetric interpretation of quantum theory possible without retrocausality? *Proc. R. Soc. A*, **2017**, 473, 20160607. arXiv:1607.07871
- [Leifer (2017)] Leifer, M. Time Symmetric Quantum Theory Without Retrocausality? A Reply to Tim Maudlin. arXiv:1708.04364
- [Merali 2022] Merali, Z.. The Sudoku universe. John Templeton Foundation, February 2022. <https://www.templeton.org/news/the-sudoku-universe>
- [Norsen & Price 2021] Norsen, T. & Price, H. Lapsing quickly into fatalism: Bell on backward causation. *Entropy*, 23(2021), 251.
- [Penrose 2023] Penrose, R., New physics for the Orch-OR consciousness proposal, in Gao, S. (ed), *Consciousness and Quantum Mechanics*, Oxford University Press, 317–362.
- [Perry 2021] Perry, M. No Future in Black Holes. arXiv:2106.03715.
- [Price 2012] Price, H. Does time-symmetry imply retrocausality? How the quantum world says “maybe”. *Studies in History and Philosophy of Modern Physics*, 43, 75–83. arXiv:1002.0906
- [Price & Wharton 2015] Price, H. & Wharton, K. Disentangling the quantum world. *Entropy* 17:11, 7752–7767. arXiv:1508.01140

- [Price & Wharton 2016] Price, H. & Wharton, K. Taming the quantum spooks. *Aeon*, 14 September 2016. <https://aeon.co/essays/can-retrocausality-solve-the-puzzle-of-action-at-a-distance>
- [Price & Wharton 2020] Price, H. & Wharton, K. A live alternative to quantum spooks. *International Journal of Quantum Foundations*, 6, 1–8. arXiv:1510.06712
- [Price & Wharton 2021] Price, H. & Wharton, K. Entanglement Swapping and Action at a Distance. arXiv:2101.05370v4 [quant-ph]. doi.org/10.48550/arXiv.2101.05370
- [Price & Weslake 2010] Price, H. & Weslake, B. The time-asymmetry of causation. In Helen Beebe, Christopher Hitchcock and Peter Menzies (eds), *The Oxford Handbook of Causation* (OUP), 414–443.
- [Schrödinger 1935a] Schrödinger, E. Discussion of probability relations between separated systems. *Mathematical Proceedings of the Cambridge Philosophical Society*, 31, 555–563.
- [Schrödinger 1935b] Schrödinger, E. Die gegenwärtige Situation in der Quantenmechanik. *Naturwissenschaften*, 23, 807–812. doi:10.1007/BF01491891.
- [Sciama 1958] Sciama, D. W., ‘Determinism and the Cosmos’, in S. Hook (ed.), *Determinism and Freedom in the Age of Modern Science*, N.Y. Univ. Press, New York, 76–78.
- [Susskind & Friedman 2014] Susskind, L. & Friedman, A. *Quantum Mechanics: The Theoretical Minimum*, Basic Books.
- [Weinberg 2013] Weinberg, S. *Lectures on Quantum Mechanics*, Cambridge University Press.
- [Wharton & Argaman 2020] Wharton, K. & Argaman, N. Bell’s Theorem and locally-mediated reformulations of quantum mechanics. *Reviews of Modern Physics*, 92, 21002. arXiv:1906.04313.
- [Wharton & Price 2016] Wharton, K. & Price, H. To understand your past, look to your future. *Nautilus*, 12 May 2016. <https://nautil.us/to-understand-your-past-look-to-your-future-235937/>

[Zyga 2017] Zyga, L. Physicists provide support for retrocausal quantum theory, in which the future influences the past. *Phys.org*, 5 July 2017. <https://phys.org/news/2017-07-physicists-retrocausal-quantum-theory-future.html>