Foundations of Physics (2021) 51:12
https://doi.org/10.1007/s10701-021-00412-5

How Quantum is Quantum Counterfactual Communication?

Jonte R. Hance[1] · James Ladyman[2] · John Rarity[1]

Received: 3 June 2020 / Accepted: 4 December 2020 / Published online: 4 February 2021

Abstract

Quantum Counterfactual Communication is the recently-proposed idea of using quantum physics to send messages between two parties, without any matter/energy transfer associated with the bits sent. While this has excited massive interest, both for potential 'unhackable' communication, and insight into the foundations of quantum mechanics, it has been asked whether this process is essentially quantum, or could be performed classically. We examine counterfactual communication, both classical and quantum, and show that the protocols proposed so far for sending signals that don't involve matter/energy transfer associated with the bits sent must be quantum, insofar as they require wave-particle duality.

1 Introduction

Quantum Counterfactual Communication is the combination of counterfactual circumstances (where "things… might have happened, although they did not in fact happen" [1]) with quantum physics, to send information between two parties without any matter/energy transfer associated with the bits sent (although in any practical implementation of it there is matter/energy transfer in the protocol as a whole). Given its interesting foundational implications, and potential for 'unhackable' communication, it has excited massive interest in recent years [2–65].

What, if anything, makes these protocols essentially quantum [14]? To answer this, we need to determine the underlying structure of classical counterfactual communication for comparison (Sect. 2), and give a sufficient condition for a protocol to be quantum (Sect. 3).

Section 4 then examines the quantum counterfactual communication protocols proposed so far, to assess their non-classicality. We show what separates them from

✉ Jonte R. Hance
jonte.hance@bristol.ac.uk

[1] Quantum Engineering Technology Laboratories, Department of Electrical and Electronic Engineering, University of Bristol, Woodland Road, Bristol BS8 1US, UK

[2] Department of Philosophy, University of Bristol, Cotham House, Bristol BS6 6JL, UK

classical counterfactual communication protocols and how they meet the condition for being quantum.

We identify two essential differences between classical and quantum counterfactual communication. The first is that only one bit-value (e.g. ‘0’) can be sent in a classical protocol without matter or energy transfer associated with the bit being sent. The second is that the two-value quantum protocols require wave-particle duality to be able to send either bit value of each bit sent.

2 Classical Counterfactuality

Counterfactual communication long predates quantum mechanics. For instance, in the Sherlock Holmes story, *Silver Blaze*, Holmes infers a racehorse was abducted by its own trainer, as the stable dog didn’t bark. As Holmes puts it, “the curious incident of the dog in the night-time” was that the dog did nothing [66]. A more recent fictional example is the Bat-Signal. If there were a major crime being committed the Bat-Signal would appear in the sky, and so the Bat-Signal’s absence counterfactually communicates to Bruce Wayne that all is well. Whenever we receive information from a sign’s absence we are being signalled to counterfactually (e.g. the signal that an engine’s components are functioning as they should is that the warning light is off).

Obviously in each of these cases a single bit is transmitted, and the bit value is signalled without the transfer of matter or energy. However, only one bit-value can be communicated by an absence in this way. Had a stranger kidnapped the racehorse, the dog would have barked, and energy would have have been transferred through the communication channel; correspondingly of course the Bat-Signal and other warning lights involve the transmission of energy when they are on. A sign’s absence can transmit one value of a bit, only if the sign always occurs for the other bit value [67]. This is counterfactual communication based on counterfactual inference.[1] Counterfactual inference is not rare but ubiquitous in everyday life and in science. For example, if there was an ether then the Michelson-Morley experiment would not have a null result.

The structure of classical counterfactual communication as above is as follows. Were A to happen, B would happen. B did not happen. Therefore A did not happen. B not happening is a signal that A did not happen, only because B happening signals that A happened.

Formally,

$$A \supset B; \neg B; \therefore \neg A \tag{1}$$

Any instance of this structure in which B doesn’t happen can be thought of as counterfactual communication of A’s not happening. However, typically, we want a

[1] It is not required that the signalled event’s non-occurrence directly causes the the sign’s absence, as there could be other common causal factors. Also a channel may be more reliable for signalling one bit value than for another.

one-to-one correspondence between the signalling event A and the inferred event B, so we always recognise the inferred event's absence. For this, we need the further condition that, were A not to happen, B would not happen ($\neg A \supset \neg B$).

3 Quantum as Non-classical

Next, to evaluate the proposed protocols we need a sufficient condition for something being quantum. There are many differences between classical and quantum physics. For optics (which all protocols so far have used), classical physics is everything up to and including Maxwell's equations. These formulate light as the evolution of waves whose intensity can be split continuously [68]. In contrast to this, in the quantum optics needed for many situations, we must consider light as photons [69], which are quanta of the electromagnetic field that are detected as discrete packets of absorbed energy. Despite this discrete particle-like behaviour, in propagation photons retain wave-like properties such as interference. Therefore, in the context of optics it is appropriate to take a protocol to be quantum if it requires using both wave- and particle-like features by combining interference with single photon detection. The latter nullifies the splitting of light intensity across different detectors, and forces it to end in a single location.

4 Protocol Evaluation

4.1 Salih et al.'s Protocol

Of the protocols proposed so far, only one has been shown counterfactual by both Weak Trace and Consistent Histories - Salih et al's [34, 41]. We show this protocol in Fig. 1 and give a detailed description in the associated caption. Above we argued that a sufficient condition for a protocol to be quantum is that it requires both discreteness and path-interference - which, for light, only single photons can do. We now consider whether this protocol has to meet this condition in order to be counterfactual.

While the limit of many single photons may generate the same results as coherent states, the way in which they produce them differs. This is due to the discreteness discussed, which is not considered when using coherent states, but is when using Fock states (i.e. single photons).

In the quantum case, beamsplitters split a photon's probability amplitude between the two eigenstates that correspond to the photon going in each direction; in the classical case, they split the beam intensity (and field). As interference still occurs, when Bob does not block, waves on both sides still destructively interfere, so the light never returns to Alice. However, Bob's D_3 and Alice's D_0 both detect light simultaneously. Similarly, when he blocks, light goes to his blockers and Alice's D_1 simultaneously. Therefore, in both cases, as light goes between Alice and Bob, it is not counterfactual. While the amount of light going to Bob's D_3 may be infinitesimal for an infinite number of outer cycles, and that

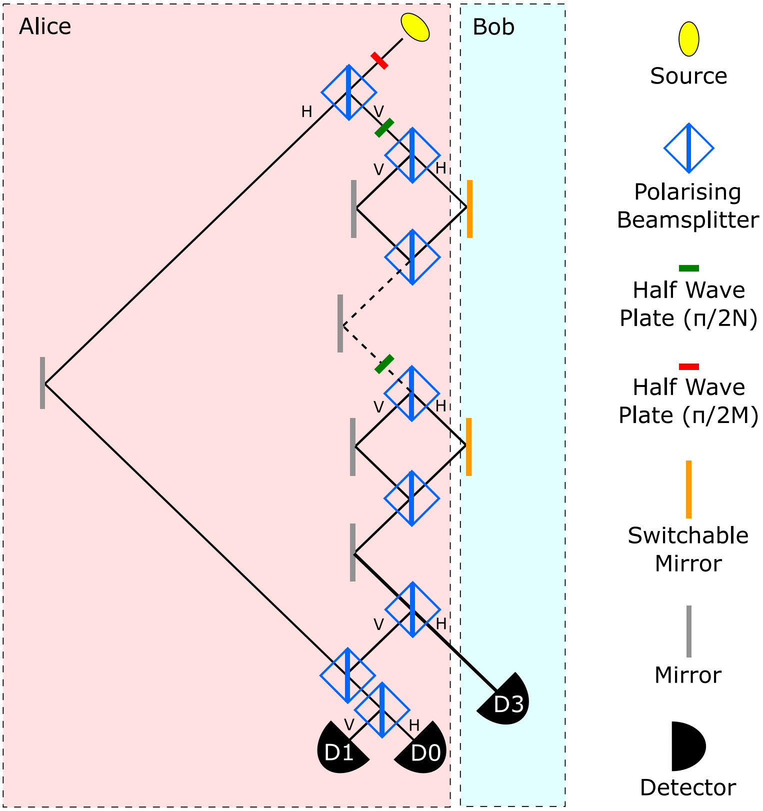

Fig. 1 A schematic diagram of Salih et al.'s protocol for counterfactual communication, where, for every bit communicated, provably no photons have been to Bob. This version shows for one outer interferometer cycle ($M = 1$), and multiple (N) inner interferometer cycles. The region of Alice is shown in pink, and the region of Bob in blue. Each Polarising Beamsplitter (PBS) reflects any vertically-polarised (V) light, and transmits any horizontally-polarised (H) light. The Half-Wave Plates (HWPs) rotate polarisation between horizontal and vertical polarisation unitarily, by an angle of either $\pi/2M$ (the HWPs before each outer interferometer) or $\pi/2N$ (the HWPs before each inner interferometer). This means, when Alice injects a horizontally-polarised photon from the source into her apparatus, the outer (left) path contains only horizontally-polarised light (labelled H), while a small amount of V-polarised light is created by the first HWP, and injected into the inner interferometer chain. In each inner interferometer, the left (Alice) path contains only V-polarised light, and the right (at Bob) only H-polarised light. If Bob does not block his paths, the chain of N 'π/N' rotations turn the V-polarised component in the inner interferometer chain to H, and so anything in the inner chain is sent to a loss channel D_3. This means Alice can only receive the photon if it went via her outer path, and so arrived at her D_0. If Bob blocks his paths, he absorbs this inner-chain H-polarised light, so the light is continually reset to V-polarised at the end of each inner interferometer, which stays on Alice's side, and reaches her D_1 as a V-polarised photon, having never travelled to Bob. A small number of photons are absorbed at Bob, reducing the efficiency. The probability of this happening decreases as M and N increase. Here, M is 1, but in general $M \geq 2$, $N \geq 2$. The only way Alice's D_1 can click is if Bob blocks; and in the infinite limit of chained outer cycles, the only way her D_0 can click is if he doesn't. Unlike the classical case, both the '0' and '1' bit-values are received without energy transfer across the channel, and so both are sent counterfactually [34, 41]

going to his blockers infinitesimal for infinite inner cycles, this is not the same as no light going there in either case - so, regardless the number of cycles, with classical light, the protocol isn't counterfactual. This may seem obvious, but many have not realised this and claimed this protocol could be performed classically (e.g. [14]).

The only way to avoid this is to force the light to end at only one point - to postselect, with information only travelling when nothing goes between Alice and Bob. Only single photons can do this. Therefore, the only way to make the protocol counterfactual is to use these, and so make the protocol quantum.

4.2 Vaidman's Protocol

Alongside the proven protocol of Salih et al, Vaidman recently proposed one [54] which, while not yet proven valid by Consistent Histories, has been shown to be valid by the Weak Trace criterion. We show this protocol in Fig. 1 and give a detailed description in the associated caption. Similarly to Salih et al's, it relies on both interference and single-photon detection - without the use of single photons, when Bob blocks the paths on his side of the inner interferometers, light could reach both Bob's blocker and Alice's detector D_1. Further, when Bob doesn't block, light could reach both the loss channels (marked 'D_L' in Fig. 2) and Alice's D_0 - in both these cases, the protocol would definitely not be counterfactual.

Extrapolating from these protocols, any counterfactual protocol (valid or not) whereby the interference from Bob blocking or not blocking his side of an interferometer affects the destination of light on Alice's side, without them both simultaneously detecting that light, relies on both wave-like and particle-like properties. This means, all these protocols, from the Elitzur-Vaidman Bomb Detector, to Noh's counterfactual cryptography scheme, to even Arvidsson-Shukur et al's proposal (despite only sending one bit-value counterfactually), are essentially quantum by our definition above.

5 Conclusion

We have shown Quantum Counterfactual Communication is essentially quantum. This confirms that both particle-like behaviour and path interference are necessary for schemes where *both* bit-values are sent counterfactually. In all schemes demonstrated so far, this is the only way it is quantum (though protocols which send quantum information have been proposed theoretically [37, 38, 40, 70, 71]). Quantum Counterfactual Communication allows us to look at principles at the heart of the foundations of quantum physics - self-interference and counterfactual non-definiteness [72] - in a new and exciting way, and will hopefully motivate new thought experiments and experimental work based on this seemingly impossible phenomenon.

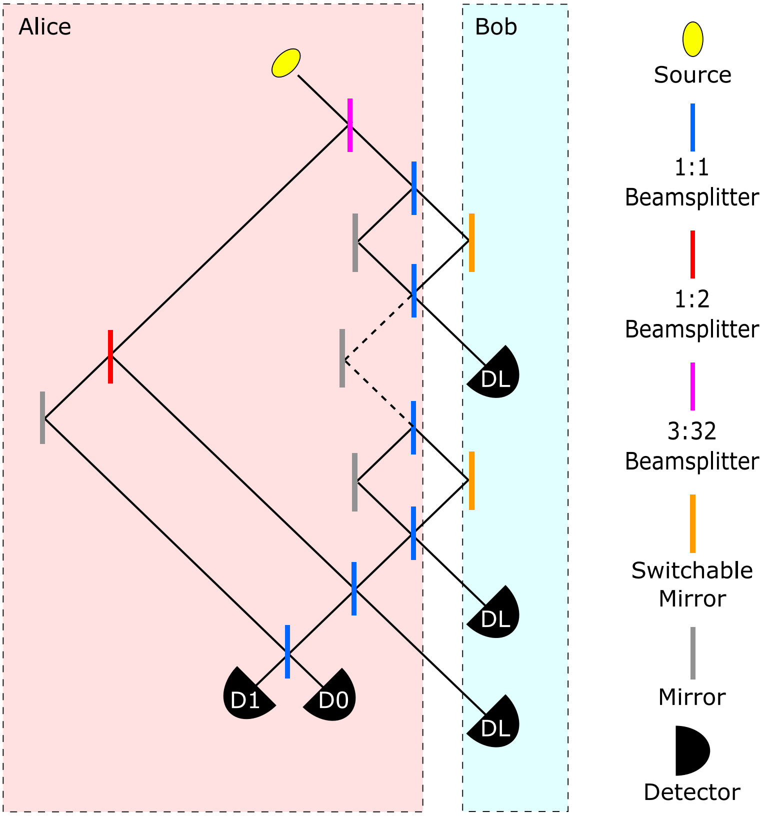

Fig. 2 A schematic diagram of Vaidman's recent protocol for counterfactual communication [54]. The region of Alice is shown in pink, and the region of Bob in blue. Unlike Salih et al's, it doesn't use polarised light, instead making use of ordinary (non-polarising) beamsplitters. By their design, when Bob doesn't block his side of the channel, each inner interferometer outputs into loss channels (here, the D_L s). This means (unless the photon is lost) the interferometer on the outer arm (starting with the red beamsplitter) always outputs at D_0. However, when Bob blocks, waves coming out of the inner interferometer chain negatively interfere at the final beamsplitter, causing the photon (if it stays at Alice) to go to D_1. The beam-splitting values given in this example make the losses equal, for two inner interferometers, regardless of whether or not Bob blocks his path (sends 1 or 0). We discuss this further in Appendix 1

A Path of a Quantum Particle

To investigate Quantum Counterfactual Communication we need a criterion for a quantum particle being/not being somewhere. This will allow us to establish if a protocol is actually counterfactual.

Classical Approach

When considering where a quantum particle has been, our first instinct is to treat it as spatially local. However, as quantum phenomena are not solely particle-like, but as also have wave-like properties (as the Two-Slit Experiment shows [73]), this is not necessarily the case. Therefore, we need a criterion for counterfactuality that takes into account quantum phenomena.

No-Signal Approach

Based on this need for a stronger criterion, a common view posed (typically on initially hearing of quantum counterfactual communication), is denying counterfactual communication is possible, by saying something can only be counterfactual if no information can pass between A and B. However, as we showed above, there are many classical protocols which are considered counterfactual communication, despite not fulfilling this criterion. Based on counterfactuality referring to something which “might have happened, although they did not in fact happen”, there is no reason to only call something counterfactual if no information can be transferred. Counterfactual communication is demonstrably possible - we just look at an extension of it.

Density Matrix Approach

Our first non-classical approach to a particle’s location is using the entire quantum mechanical description of the system. We can do this by looking at the density matrix, which shows the whole of a state—it provides all the information that exists about a particle at a given moment. By associating them with physical positions, we can see the spread of the particle’s possible locations.

However, some parts of the density matrix correspond to paths lost when the wavefunction is collapsed at the protocol’s end, and so information not being sent. We need some way to sort these possible paths into those where information is sent (that the particle could have been on when counterfactual communication occurred) and those where it is not. This requires post-selection (selection based on the final state they lead to), rather than just pre-selection from an initial state.

Consistent Histories

To resolve this, we could use the Consistent Histories approach [74]. This involves creating histories (tensored chains of projectors, each a way the system could evolve). A family is a set of these histories, that form a projective decomposition of the identity operator over the whole evolution time. By refining certain projectors (replacing them with projectors summing to them), we can model a situation. These histories are consistent if they all mutually commute—if they do not, it is meaningless to ask which history was the ‘correct’ model for the system’s evolution, as we cannot assign positive relative probabilities [75].

Therefore, we can say a particle has not gone between Alice and Bob when all histories where it travels between the two (where information is sent), have probability zero. This requires us to analyse all histories in this family, as Griffiths does for several protocols [15]. In essence, if a particle can go between Alice and Bob when information is transmitted, by Consistent Histories it is not counterfactual; if not all the histories in the family are consistent with one another, it is meaningless to talk about the counterfactuality of the situation.

Weak Trace

Next, we consider weak measurement [76], which examines the state between measurements, without collapsing it. Weak measurement involves lightly coupling a system to a measuring device, so while little information is gathered over one run of the system, over many runs a probability distribution is obtained. This contrasts with Von Neumann measurements, which cause a system to collapse into an eigenstate of the measured operator. Weak measurement allows us to collect information that would be lost were the system strongly measured [77]. To do this we apply the expectation value of the evolution operator to the initial state, which we can interpret as evaluating all forward-evolving paths from that state.

However, rather than working forwards, can also work back from a given result (post-select), to investigate the paths the system may have evolved through. If we pre- and post-select like this, we say a particle leaves a weak trace (indicating possible presence) wherever this weak measurement value is non-zero. To approximate this trace, we can trace the initial vector forward, and the final vector backward, in time, and see where they overlap. If we represent the evolution of the state by operator $\hat{O}$, we get this approximate value as

$$O_w = \frac{\langle \psi_f | \hat{O} | \psi_i \rangle}{\langle \psi_f | \psi_i \rangle} \tag{2}$$

where $|\psi_i\rangle$ is the initial state of the particle, and $|\psi_f\rangle$ the final. This approximation of the weak value (to the first order in trace size, $\mathcal{O}(\epsilon)$) is the Two-State Vector Formalism (TSVF), as it goes from/to two states—that at the beginning of the protocol, and that at the end. This gives both pre- and post-selection needed.

If an operator returns a non-zero TSVF value, there is to $\mathcal{O}(\epsilon)$ a weak trace along the path it describes. If we trace the paths a quantum particle could evolve along from its initial, and those it could have come from to get to its final, state, there is a weak trace where they overlap—so we cannot say the particle was not there [51] (see Fig. 3).

However, this has unintuitive results. While, with Consistent Histories, a path needs to link the initial and final states, here, it only requires paths from the initial and final states overlap at some point. This means Bob can have a weak trace on his side of the transmission channel, without any in the channel itself. This was demonstrated using weak measurements in nested Mach-Zehnder Interferometers (MZIs) by Danan et al [78]. If one accepts the weak trace as a valid indicator of a particle's path, this leads to peculiarities—such as particles jumping discontinuously between locations [79]. This caused Sokolovski to doubt the formalism, in favour of continuous paths. [80–82]. However, these peculiar results don't contradict standard quantum theory [83].

A more compelling counterargument is that the TSVF ignores the non-$\mathcal{O}(\epsilon)$ weak trace, and so does not give the particle's entire path. Vaidman admits this, saying the TSVF only gives the weak trace to $\mathcal{O}(\epsilon)$. Further, analysis of Danan et al's data shows smaller, $\mathcal{O}(\epsilon^2)$ peaks not visible in their original presentation [84]. Vaidman explains this by saying the non-local trace on any particle is also of $\mathcal{O}(\epsilon^2)$, and so this applies in any set-up, even if objects are physically separated. Further, as there are no non-local interactions in nature, this non-local weak trace cannot be strong enough to mediate any effects, so neither can a local second-order weak trace [51]. Despite this, Vaidman still claims this constitutes a weak trace for Salih et al's one-outer-cycle protocol [54].

Quantum Counterfactual Communication Protocols

In this Appendix, we look at Quantum Counterfactual Communication protocols proposed so far. Since Elitzur and Vaidman first discovered quantum counterfactuality [85], and Kwiat et al allowed loss to be made effectively nil [86], researchers have tried to exploit it for communication. Despite this, all protocols until recently fell into three broad categories: where communication is counterfactual only for one bit-value; where photons travel between Alice and Bob, but in the opposite direction to the information passed between them; and where no photons pass between Alice and Bob when information flows, but the error/loss rates vary with the bit-value Bob sends.

Elitzur–Vaidman Bomb Tester

All quantum counterfactual communication protocols stem from the Elitzur-Vaidman Bomb-Detector [85]. Here (Fig. 4), an MZI has a potentially faulty bomb along one of its paths, which can only be detonated by a non-demolition single-photon detection.

If the photon goes along the bomb's side of the MZI (and the bomb works), it detonates, and the photon (and everything else) is destroyed; if the bomb is faulty, the photon travels to the merging beamsplitter normally. However, if the photon traverses the other side, the bomb working changes the interference pattern, making it able to go to a detector it previously couldn't access. This allows us to test if the bomb would have worked, without detonating it, by putting it on the path the photon could have, but did not, travel down. Unlike classical counterfactual communication, both options are transmitted counterfactually—the photon's path is from the source to the detectors without going via the object (bomb) under

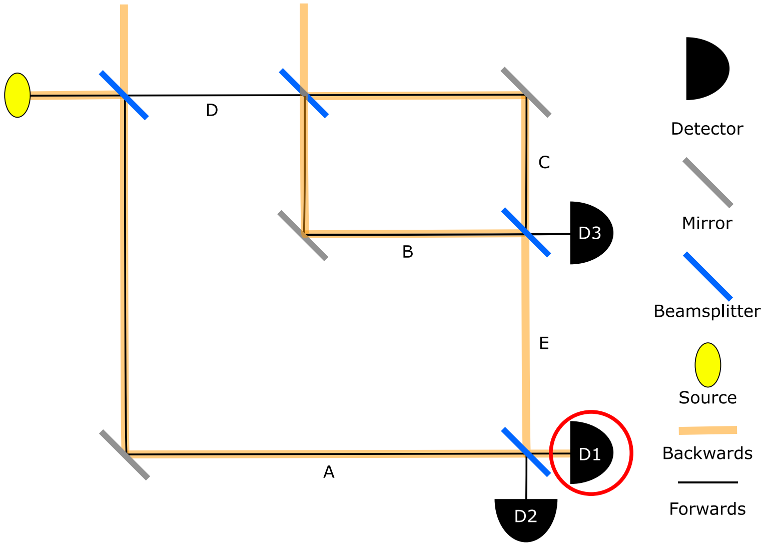

Fig. 3 The Two-State Vector formalism applied to a nested interferometer. Forwards-travelling paths are marked by thin black lines, and backwards-travelling paths by thick orange lines. Though no forward or backwards travelling path goes from the source, to Bob (along path C) and into D2, they do overlap over C, meaning there is a weak trace at Bob. This illustrates the peculiar property of the TSVF where particles can jump between regions (e.g. between the inner interferometer and the outer arm) [51]

evaluation. However, it is not necessarily always counterfactual. This is as, while the photon *can* carry the information without going via the bomb, it does not necessarily have to. This means the Bomb-Tester is not fully counterfactual.

Counterfactual only for One Bit

Here, the protocol is counterfactual for one bit-value, but the photon goes between Alice and Bob for the other.

The first of this type of protocol, and indeed the first Quantum Counterfactual Communication protocol proposed was Noh's (Fig. 5) [31], (barring Guo's E-V Bomb Detector adaption, where photons travel between Alice and Bob for both bit-values [18]). For matched polarisations, if Alice gets a click, the photon has remained on her side. However, for orthogonal polarisations, it has both been to, and returned from, Bob—so is not counterfactual. Despite this, the work generated a lot of interest [7, 25–29, 32, 43–47, 49, 50, 55–58, 60–62]. While plenty of these focus on reducing loss by reducing the proportion of the photon sent to Bob [50], this must be non-zero for the protocol to function. Therefore, the system will never be fully counterfactual.

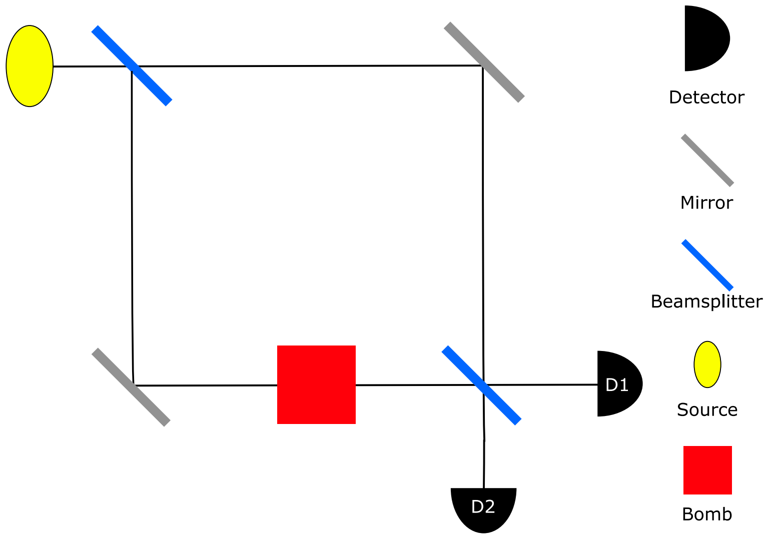

Fig. 4 The Elitzur–Vaidman Bomb Tester. A photon is emitted from the source (top-left), enters the balanced Mach-Zehnder interferometer, and is spread across both paths equally. If the bomb is faulty, the photon recombines at the second beam-splitter, and always enters D_1. If the bomb works, and is activated, it destroys the set-up. If the bomb would work, but the photon went down the bomb-free path, the photon has a 50:50 chance of going to either detector

Information and Photon Travel in Opposite Directions

In the next category, the photon can cross the channel, but in the opposite direction to the information being sent. Here, for one bit-position the photon destructively interferes across the quantum channel, keeping it at Alice, and for the other, it constructively interferes, allowing it through to Bob. Based on if she detects a photon, Alice can determine which bit Bob sent.

The only protocol of this sort is Arvidsson-Shukur et al's. They propose a device formed of chained MZIs, which use the Quantum Zeno effect to (for many MZIs) keep the photon at Alice if Bob blocks, and force it to go to Bob if he does not [4]—identically to Kwiat et al's Interaction Free Measurement protocol [86]. Ignoring the high chance of Alice wrongly believing Bob did not block (due to the necessarily finite number of MZIs causing some chance of Bob's blocker absorbing the photon), the photon still travels at the same time as the information. Waves carrying information in the opposite direction to travel is a well-known classical phenomenon [54]. Therefore, this protocol only seems quantum when you consider light as local—where, for one possibility, the photon travels from Alice to Bob. This obviously creates a weak trace at Bob, and so is not counterfactual. Arvidsson-Shukur et al attempt to advocate their protocol by saying it tolerates error better than others [6], and by calling other protocols classical using a classical model with Alice and Bob having extra, non-trivial resources (e.g. a shared clock) [5]. However, they give no reason to view their protocol as true

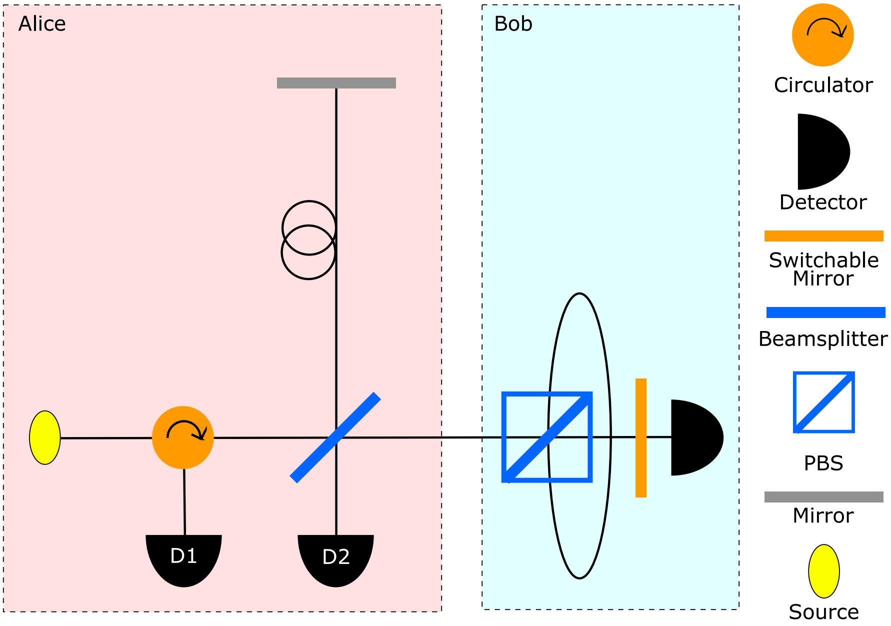

Fig. 5 Noh's counterfactual cryptography protocol—Alice randomly polarises a photon, which passes through a beam-splitter, with one of the outputs going to Bob. Bob uses a PBS and delay to time-separate possible polarisations arriving it him, and picks one to reflect and one to absorb. If the photon is Bob's reflect-polarisation, it reflects back, and interferes into Alice's D_2. However, if the photon is his absorb-polarisation, it is sent into his detector. If this clicks, the protocol is aborted; if not, the photon goes into D_1 [31]

Quantum Counterfactual Communication, and so their work with Calafell et al [8] just demonstrates classical counterfactuality.

Photon only Travels Erroneously

The next set of protocols have Alice receiving a photon for both bit values, which has never been to Bob. This means the photon cannot go to Bob when Alice gains information, as then Alice would be unable to see what was sent. Therefore, when photons go to Bob, the protocol is aborted and retried, creating a source of loss.

Unequal Losses

For these protocols, this loss varies with the bit-value sent. This leads us to ask if Alice can, by knowing loss probabilities, guess the bit Bob sends solely based on if any of her detectors light up. This would make the same as the last category. We can also ask if this post-selection is to blame for any peculiar effects observed, even with this loss enforced according to the protocol.

Salih et al's 2013 protocol was the first claiming to be fully counterfactual for both bit values [33, 34]. It is formed of a chain of outer interferometers, each containing a chain of inner interferometers (see Fig. 1). However, Vaidman claimed this was not counterfactual, when assessed by the Weak Trace criterion (see Fig. 3) [51]. The TSVF gives a weak trace on Bob's side of the channel when Bob does not block this, meaning we cannot say the photon was not there [52, 53]. However, this is only for the simplest (polarisation-free) form of the protocol. Using polarisation (as shown in Fig. 1 and [41]) avoids a weak trace on Bob's side by ensuring the only waves that go to Bob are H-polarised, which are lost via D_3 on Bob's side, restarting the protocol [36]. Then, when Bob does not block, the differences in polarisation between the forward- and backward-travelling states keep them separate on Bob's side—giving no Weak Trace there. This was tested practically through weak measurement, using Danan et al's method [78] and shown to have no weak trace from Bob's side visible at Alice's detectors. Griffiths also claimed it is not counterfactual by Consistent Histories, as a history with a non-zero probability could be traced to Bob's side and back when Bob does not block [15, 16, 74]. However, again, Griffiths only considered physical paths, rather than polarisations, which provide an extra degree of freedom [39]. Griffith later showed, when using more than one outer cycle, the family became inconsistent, and so calling the protocol counterfactual was meaningless [17]. However, Salih notes the final outer cycle is counterfactual, while identical earlier ones are meaningless, which seems paradoxical [40]. Once Salih et al published their protocol, various implementations began to appear [3, 9–13, 19–24, 30, 35, 37, 38, 40, 42, 47, 48, 59, 63, 65, 70, 71, 87]. While many of these don't make use of polarisation, some do.

Zhang et al proposed a protocol, based on Salih et al's, for probabilistic counterfactual communication. Their protocol isn't always counterfactual, but they claim the chance of the photon being at Bob can be reduced to nil, and losses (from noise/blocking) are lower [64]. However, they assume photons only trace one path, which isn't always true—so it is not counterfactual.

Despite originally claiming counterfactual communication of both bit-values was impossible, at roughly the same time as Salih et al defended their protocol using polarisation, Vaidman, alongside Aharonov, released a protocol allowing just this [2]. This method is effectively the same as in Salih et al's original protocol—however, to avoid a weak trace in this set-up, where there is no polarisation degree of separation, at least two inner interferometers are needed. Alongside this, Aharonov and Vaidman make repeated reference to a double-sided mirror in the protocol; but all this does is connect the two inner interferometers, and fold the outer path to reduce physical space used, and so it is irrelevant to the protocol's counterfactuality. Also, unlike Salih et al's, this protocol is not counterfactual by Consistent Histories [40].

Equal Losses

Shortly after publishing with Aharonov, Vaidman created another weak trace-free counterfactual communication protocol. For one outer cycle, this protocol avoids the risk of an erroneous reading that Salih et al's, and their earlier, protocol has [54]. It is again based on a chained MZI set-up, but uses interference from light passing through the inner interferometer chain to alter which detector the photon ends up at when Bob blocks (see Sect. 4.2). This allows Alice, when she receives a bit, to be certain it is the same value Bob sent. Another benefit of the protocol is that, for certain beam-splitter values, losses were the same whether or not Bob blocked. This means Alice cannot infer if Bob blocked, just based on if she receives a photon. Therefore, the protocol cannot be reduced to the information and photon travelling simultaneously, in opposite directions, and so it seems counterfactual. However, it remains to be seen if it is fully counterfactual by the Consistent Histories criterion.

Acknowledgements We thank Hatim Salih, Will McCutcheon, Paul Skrzypczyk and Robert Griffiths for useful discussions. This work was supported by the Engineering and Physical Sciences Research Council (Grants EP/P510269/1, EP/T001011/1, EP/R513386/1, EP/M013472/1 and EP/L024020/1).

References

1. Penrose, R.: Shadows of the Mind, vol. 4. Oxford University Press, Oxford (1994)
2. Aharonov, Y., Vaidman, L.: Modification of counterfactual communication protocols that eliminates weak particle traces. Phys. Rev. A. **99**, 010103(R) (2019). https://doi.org/10.1103/PhysRevA.99.010103
3. Al Amri, M., Li, Z.H., Zubairy, M.S.: Second International Seminar on Photonics, Optics, and Its Applications (ISPhOA 2016), vol. 10150. International Society for Optics and Photonics, 2016. https://www.spiedigitallibrary.org/conference-proceedings-of-spie/10150/101501N/Quantum-teleportation-without-classical-channel/10.1117/12.2267245.short?SSO=1
4. Arvidsson-Shukur, D.R.M., Barnes, C.H.W.: Quantum counterfactual communication without a weak trace. Phys. Rev. A. **94**, 062303 (2016). https://doi.org/10.1103/PhysRevA.94.062303
5. Arvidsson-Shukur, D.R.M., Barnes, C.H.W.: Postselection and counterfactual communication. Phys. Rev. A. **99**, 060102(R) (2019). https://doi.org/10.1103/PhysRevA.99.060102
6. Arvidsson-Shukur, D.R.M., Gottfries, A.N.O., Barnes, C.H.W.: Evaluation of counter factuality in counterfactual communication protocols. Phys. Rev. A. **96**, 062316 (2017). https://doi.org/10.1103/PhysRevA.96.062316
7. Brida, G., Cavanna, A., Degiovanni, I.P., Genovese, M., Traina, P.: Experimental realization of counterfactual quantum cryptography. Laser Phys. Lett. **9**(3), 247 (2012). https://doi.org/10.1002/lapl.201110120/meta
8. Calafell, I.A., Strömberg, T., Arvidsson-Shukur, D., Rozema, L., Saggio, V., Greganti, C., Harris, N., Prabhu, M., Carolan, J., Hochberg, M., Baehr-Jones, T., Englund, D., Barnes, C., Walther, P.: Trace-free counterfactual communication with a nanophotonic processor. NPJ Quant. Inf. **5**(1), 61 (2019)
9. Cao, Y., Li, Y.H., Cao, Z., Yin, J., Chen, Y., Ma, X., Peng, C.Z., Pan, J.W.: CLEO: QELS_Fundamental Science. Optical Society of America (2014)
10. Cao, Y., Li, Y.H., Cao, Z., Yin, J., Chen, Y.A., Yin, H.L., Chen, T.Y., Ma, X., Peng, C.Z., Pan, J.W.: Direct counterfactual communication via quantum Zeno effect. Proc. Natl. Acad. Sci. (2017). https://doi.org/10.1073/pnas.1614560114

11. Chen, Y., Gu, X., Jiang, D., Xie, L., Chen, L.: Counterfactual quantum cryptography network with untrusted relay. Int. J. Mod. Phys. B **29**(20), 1550134 (2015). https://doi.org/10.1142/S0217979215501349
12. Chen, Y., Gu, X., Jiang, D., Xie, L., Chen, L.: Tripartite counterfactual entanglement distribution. Optics Express **23**(16), 21193 (2015)
13. Chen, Y., Jiang, D., Gu, X., Xie, L., Chen, L.: Counterfactual entanglement distribution using quantum dot spins. J. Opt. Soc. Am. B. **33**(4), 663 (2016). https://doi.org/10.1364/JOSAB.33.000663
14. Gisin, N.: Optical communication without photons. Phys. Rev. A **88**, 030301(R) (2013). https://doi.org/10.1103/PhysRevA.88.030301
15. Griffiths, R.B.: Particle path through a nested Mach-Zehnder interferometer. Phys. Rev. A **94**, 032115 (2016). https://doi.org/10.1103/PhysRevA.94.032115
16. Griffiths, R.B.: Reply to "Comment on 'Particle path through a nested Mach-Zehnder interferometer' ". Phys. Rev. A **95**, 066102 (2017). https://doi.org/10.1103/PhysRevA.95.066102
17. Griffiths, R.B.: Reply to "Comment on 'Particle path through a nested Mach-Zehnder interferometer' ". Phys. Rev. A **97**, 026102 (2018). https://doi.org/10.1103/PhysRevA.97.026102
18. Guo, G.C., Shi, B.S.: Quantum cryptography based on interaction-free measurement. Phys. Lett. A **256**(2–3), 109 (1999)
19. Guo, Q., Cheng, L.Y., Chen, L., Wang, H.F., Zhang, S.: Counterfactual distributed controlled-phase gate for quantum-dot spin qubits in double-sided optical microcavities. Phys. Rev. A **90**(4), 042327 (2014). https://doi.org/10.1103/PhysRevA.90.042327
20. Guo, Q., Cheng, L.Y., Chen, L., Wang, H.F., Zhang, S.: Counterfactual entanglement distribution without transmitting any particles. Optics Express **22**(8), 8970 (2014)
21. Guo, Q., Cheng, L.Y., Chen, L., Wang, H.F., Zhang, S.: Counterfactual quantum-information transfer without transmitting any physical particles. Sci. Rep. **5**, 8416 (2015)
22. Guo, Q., Zhai, S., Cheng, L.Y., Wang, H.F., Zhang, S.: Counterfactual quantum cloning without transmitting any physical particles. Phys. Rev. A **96**, 052335 (2017). https://doi.org/10.1103/PhysRevA.96.052335
23. Guo, Q., Cheng, L.Y., Wang, H.F., Zhang, S.: Counterfactual entanglement swapping enables high-efficiency entanglement distribution. Optics Express **26**(21), 27314 (2018)
24. Hance, J., McCutcheon, W., Yard, P., Rarity, J.: Modal, truly counterfactual communication with on-chip demonstration proposal. Quantum information and measurement (QIM) V: quantum technologies, Optical Society of America (2019). https://doi.org/10.1364/QIM.2019.T5A.50
25. Jiang, M., Sun, S., Liang, L.: Practical stabilization of counterfactual quantum cryptography. J. Quant. Inf. Sci. **1**(03), 116 (2011)
26. Li, Y.B., Wen, Q.Y., Li, Z.C.: Security flaw of counterfactual quantum cryptography in practical setting. arXiv:1312.1436 (2013)
27. Li, Y.B.: Analysis of counterfactual quantum key distribution using error-correcting theory. Quant. Inf. Process. **13**(10), 2325 (2014). https://doi.org/10.1007/s11128-014-0786-y
28. Liu, Y., Ju, L., Liang, X.L., Tang, S.B., Tu, G.L.S., Zhou, L., Peng, C.Z., Chen, K., Chen, T.Y., Chen, Z.B., Pan, J.W.: Experimental demonstration of counterfactual quantum communication. Phys. Rev. Lett. **109**, 030501 (2012). https://doi.org/10.1103/PhysRevLett.109.030501
29. Liu, X., Zhang, B., Wang, J., Tang, C., Zhao, J., Zhang, S.: Eavesdropping on counterfactual quantum key distribution with finite resources. Phys. Rev. A **90**(2), 022318 (2014). https://doi.org/10.1103/PhysRevA.90.022318
30. Liu, C., Liu, J., Zhang, J., Zhu, S.: Improvement of reliability in multi-interferometer-based counterfactual deterministic communication with dissipation compensation. Optics Express **26**(3), 2261 (2018)
31. Noh, T.G.: Counterfactual quantum cryptography. Phys. Rev. Lett. **103**, 230501 (2009). https://doi.org/10.1103/PhysRevLett.103.230501
32. Ren, M., Wu, G., Wu, E., Zeng, H.: Experimental demonstration of counterfactual quantum key distribution. Laser Phys. **21**(4), 755 (2011). https://doi.org/10.1134/S1054660X11070267
33. Salih, H., Li, Z.H., Al-Amri, M., Zubairy, M.S.: The Rochester Conferences on Coherence and Quantum Optics and the Quantum Information and Measurement Meeting. Optical Society of America (2013). https://doi.org/10.1364/CQO.2013.T4B.1
34. Salih, H., Li, Z.H., Al-Amri, M., Zubairy, M.S.: Protocol for direct counterfactual quantum communication. Phys. Rev. Lett. **110**, 170502 (2013). https://doi.org/10.1103/PhysRevLett.110.170502
35. Salih, H.: Tripartite counterfactual quantum cryptography. Phys. Rev. A **90**(1), 012333 (2014)

36. H. Salih, Z.H. Li, M. Al-Amri, M.S. Zubairy. Salih et al.: Reply: Phys. Rev. Lett. **112**, 208902 (2014). https://doi.org/10.1103/PhysRevLett.112.208902
37. Salih, H. Protocol for counterfactually transporting an unknown qubit. arXiv:1404.2200v1 (2014)
38. Salih, H.: Protocol for counterfactually transporting an unknown qubit. Frontiers Phys. **3**, 94 (2016). https://doi.org/10.3389/fphy.2015.00094/full
39. Salih, H.: Comment on "Particle path through a nested Mach-Zehnder interferometer". Phys. Rev. A **97**, 026101 (2018). https://doi.org/10.1103/PhysRevA.97.026101
40. Salih, H.: From a quantum paradox to counterportation. arXiv:1807.06586 (2018)
41. Salih, H., McCutcheon, W., Hance, J., Rarity, J.: Do the laws of physics prohibit counterfactual communication? arXiv:1806.01257 (2018)
42. Salih, H.: Counterfactual quantum erasure: spooky action without entanglement. R. Soc. Open Sci. **5**(2), 171250 (2018). https://doi.org/10.1098/rsos.171250
43. Shenoy, H.A., Srikanth, R.: The wave-function is real but nonphysical: a view from counterfactual quantum cryptography. arXiv:1311.7127 (2013)
44. Shenoy, A., Srikanth, R., Srinivas, T.: Counterfactual quantum key distribution without polarization encoding. In: QIS 2013 Proceedings (2013)
45. Shenoy-Hejamadi, A., Srikanth, R., Srinivas, T.: Semi-counterfactual cryptography. EPL (Europhys. Lett.) **103**(6), 60008 (2013). https://doi.org/10.1209/0295-5075/103/60008
46. Shenoy-Hejamadi, A., Srikanth, R., Srinivas, T.: Counterfactual quantum certificate authorization. Phys. Rev. A **89**, 052307 (2014). https://doi.org/10.1103/PhysRevA.89.052307
47. Shenoy-Hejamadi, A., Srikanth, R.: Counterfactual distribution of Schrödinger cat states. Phys. Rev. A **92**, 062308 (2015). https://doi.org/10.1103/PhysRevA.92.062308
48. Shukla, C., Pathak, A.: Orthogonal-state-based deterministic secure quantum communication without actual transmission of the message qubits. Quant. Inf. Process. **13**(9), 2099 (2014). https://doi.org/10.1007/s11128-014-0792-0
49. Song, Y.Q., Yang, L.: Quantum bit commitment protocol based on counterfactual quantum cryptography. arXiv:1709.08490 (2017)
50. Sun, Y., Wen, Q.Y.: Counterfactual quantum key distribution with high efficiency. Phys. Rev. A **82**(5), 052318 (2010). https://doi.org/10.1103/PhysRevA.82.052318
51. Vaidman, L.: Past of a quantum particle. Phys. Rev. A **87**, 052104 (2013). https://doi.org/10.1103/PhysRevA.87.052104
52. Vaidman, L.: Comment on "Protocol for direct counterfactual quantum communication". Phys. Rev. Lett. **112**, 208901 (2014). https://doi.org/10.1103/PhysRevLett.112.208901
53. Vaidman, L.: Counterfactual quantum protocols. Int. J. Quant. Inf. (2016). https://doi.org/10.1142/S0219749916400128
54. Vaidman, L.: Analysis of counter factuality of counterfactual communication protocols. Phys. Rev. A **99**(5), 052127 (2019). https://doi.org/10.1103/PhysRevA.99.052127
55. Wang, T.Y., Li, Y.P., Zhang, R.L.: Analysis of counterfactual quantum certificate authorization. Int. J. Theoret. Phys. **55**(12), 5331 (2016). https://doi.org/10.1007/s10773-016-3152-2
56. Yang, X., Wei, K., Ma, H., Sun, S., Du, Y., Wu, L.: Trojan horse attacks on counterfactual quantum key distribution. Phys. Lett. A **380**(18–19), 1589 (2016)
57. Yin, Z.Q., Li, H.W., Chen, W., Han, Z.F., Guo, G.C.: Security of counterfactual quantum cryptography. Phys. Rev. A **82**(4), 042335 (2010). https://doi.org/10.1103/PhysRevA.82.042335
58. Yin, Z.Q., Li, H.W., Yao, Y., Zhang, C.M., Wang, S., Chen, W., Guo, G.C., Han, Z.F.: Counterfactual quantum cryptography based on weak coherent states. Phys. Rev. A **86**(2), 022313 (2012). https://doi.org/10.1103/PhysRevA.86.022313
59. Zaman, F., Jeong, Y., Shin, H.: Counterfactual Bell-state analysis. Sci. Rep. **8**(1), 14641 (2018)
60. Zhang, S., Wang, J., Tang, C.J.: Counterfactual attack on counterfactual quantum key distribution. EPL (Europhys. Lett.) **98**(3), 30012 (2012). https://doi.org/10.1209/0295-5075/98/30012
61. Zhang, S., Wang, J., Tang, C.J.: Security proof of counterfactual quantum cryptography against general intercept-resend attacks and its vulnerability. Chin. Phys. B **21**(6), 060303 (2012). https://doi.org/10.1088/1674-1056/21/6/060303
62. Zhang, J.L., Guo, F.Z., Gao, F., Liu, B., Wen, Q.Y.: Private database queries based on counterfactual quantum key distribution. Phys. Rev. A **88**, 022334 (2013). https://doi.org/10.1103/PhysRevA.88.022334
63. Zhang, S., Zhang, B., Liu, X.T.: Improved direct counterfactual quantum communication. arXiv:1410.2769 (2014)

64. Zhang, S.: Probabilistic direct counterfactual quantum communication. Chin. Phys. B **26**(2), 020304 (2017). https://doi.org/10.1088/1674-1056/26/2/020304
65. Zubairy, M.S., Li, Z., Al-Amri, M.D., Salih, H.A.: Method and apparatus for direct counterfactual quantum communication. US Patent **8**(891), 767 (2014)
66. Doyle, A.C.: The Memoirs of Sherlock Holmes. George Newnes, London (1894)
67. Maudlin, T.: Quantum Non-Locality and Relativity: Metaphysical Intimations of Modern Physics. Blackwell, London (2002)
68. Griffiths, D.J.: Introduction to Electrodynamics. AAPT, Boston (2005)
69. Einstein, A.: Über einen die Erzeugung und Verwandlung des Lichtes betreffenden heuristischen Gesichtspunkt. Ann. Phys. **322**, 132 (1905). https://doi.org/10.1002/andp.19053220607
70. Salih, H., Hance, J.R., McCutcheon, W., Rudolph, T., Rarity, J.: Exchange-free computation on an unknown qubit at a distance. New J. Phys. **23**(1), 013004 (2021). https://doi.org/10.1088/1367-2630/abd3c4
71. Salih, H., Hance, J.R., McCutcheon, W., Rudolph, T., Rarity, J.: Deterministic teleportation and universal computation without particle exchange. arXiv:2009.05564 (2020)
72. Hance, J.R.: Counterfactuality, Definiteness and Bell's Theorem. arXiv:1909.06608 (2019)
73. Young, T.: The Bakerian lecture: experiments and calculations relative to physical optics. Philosoph. Trans. R. Soc. Lond. Ser. https://doi.org/10.1098/rstl.1804.0001
74. Griffiths, R.B.: Consistent histories and the interpretation of quantum mechanics. J. Stat. Phys. **36**(1–2), 219 (1984). https://doi.org/10.1007/BF01015734
75. Griffiths, R.B.: The Stanford Encyclopedia of Philosophy. In: Zalta, E.N. (ed.) Summer 2019. Stanford University, Metaphysics Research Lab (2019)
76. Aharonov, Y., Albert, D.Z., Vaidman, L.: How the result of a measurement of a component of the spin of a spin-1/2 particle can turn out to be 100. Phys. Rev. Lett. **60**, 1351 (1988). https://doi.org/10.1103/PhysRevLett.60.1351
77. Tamir, B., Cohen, E.: Introduction to weak measurements and weak values. Quanta **2**(1), 7 (2013)
78. Danan, A., Farfurnik, D., Bar-Ad, S., Vaidman, L.: Asking photons where they have been. Phys. Rev. Lett. **111**(24), 240402 (2013). https://doi.org/10.1103/PhysRevLett.111.240402
79. Aharonov, Y., Cohen, E., Landau, A., Elitzur, A.C.: The case of the disappearing (and re-appearing) particle. Sci. Rep. **7**(1), 531 (2017)
80. Sokolovski, D.: Are the weak measurements really measurements? Quanta **2**(1), 50 (2013).
81. Englert, B.G., Horia, K., Dai, J., Len, Y.L., Ng, H.K.: Past of a quantum particle revisited. Phys. Rev. A **96**, 022126 (2017). https://doi.org/10.1103/PhysRevA.96.022126
82. Sokolovski, D., Akhmatskaya, E.: An even simpler understanding of quantum weak values. Ann. Phys. **388**, 382 (2018). https://doi.org/10.1016/j.aop.2017.11.030
83. Peleg, U., Vaidman, L.: Comment on "Past of a quantum particle revisited". arXiv:1805.12171 (2018)
84. Sokolovski, D.: Asking photons where they have been in plain language. Phys. Lett. A **381**(4), 227 (2017)
85. Elitzur, A.C., Vaidman, L.: Quantum mechanical interaction-free measurements. Found. Phys. **23**(7), 987 (1993). https://doi.org/10.1007/BF00736012
86. Kwiat, P., Weinfurter, H., Herzog, T., Zeilinger, A., Kasevich, M.A.: Interaction-free measurement. Phys. Rev. Lett. **74**, 4763 (1995). https://doi.org/10.1103/PhysRevLett.74.4763
87. Hance, J., Rarity, J.: Counterfactual ghost imaging. arXiv:2010.14292 (2020)