Mathematics, Physics, and Computer Science
Faculty Articles and Research

Science and Technology Faculty Articles and
Research

8-6-2021

A Dynamical Quantum Cheshire Cat Effect and Implications for Counterfactual Communication

Yakir Aharonov

Eliahu Cohen

Sandu Popescu

A Dynamical Quantum Cheshire Cat Effect and Implications for Counterfactual Communication

Comments

Creative Commons License

Copyright

ARTICLE

A dynamical quantum Cheshire Cat effect and implications for counterfactual communication

Yakir Aharonov[1,2], Eliahu Cohen[3] & Sandu Popescu[4✉]

Here we report a type of dynamic effect that is at the core of the so called "counterfactual computation" and especially "counterfactual communication" quantum effects that have generated a lot of interest recently. The basic feature of these counterfactual setups is the fact that particles seem to be affected by actions that take place in locations where they never (more precisely, only with infinitesimally small probability) enter. Specifically, the communication/computation takes place without the quantum particles that are supposed to be the information carriers travelling through the communication channel or entering the logic gates of the computer. Here we show that something far more subtle is taking place: It is not necessary for the particle to enter the region where the controlling action takes place; it is enough for the controlled property of the particle, (i.e., the property that is being controlled by actions in the control region), to enter that region. The presence of the controlled property, without the particle itself, is possible via a quantum Cheshire Cat type effect in which a property can be disembodied from the particle that possesses it. At the same time, we generalize the quantum Cheshire Cat effect to dynamical settings, in which the property that is "disembodied" from the particle possessing it propagates in space, and leads to a flux of "disembodied" conserved quantities.

[1] School of Physics and Astronomy, Tel Aviv University, Tel Aviv, Israel. [2] Schmid College of Science and Technology and Institute for Quantum Studies, Chapman University, Orange, CA, USA. [3] Faculty of Engineering and the Institute of Nanotechnology and Advanced Materials, Bar Ilan University, Ramat Gan, Israel. [4] H. H. Wills Physics Laboratory, University of Bristol, Bristol, UK. ✉email: s.popescu@bristol.ac.uk

Quantum mechanics is a notoriously counterintuitive theory. More than ninety years after its basic laws were discovered, there is a widespread agreement that we do not yet really understand what is going on. Nevertheless, in recent years considerable progress has been made, especially in understanding entanglement and Bell-type non-locality. However, the focus of that research is essentially kinematic—these effects depend only on the Hilbert space structure of the set of quantum states. On the other hand, the fundamental dynamic aspects of the theory have been far less investigated, though, undoubtedly, they are equally rich, if not richer. Here we describe a "Dynamic Quantum Cheshire Cat effect".

In the original Quantum Cheshire Cat effect[1], it was shown that physical properties can be disembodied from the objects to which they belong. For example, we may find an electron in one location and its spin in a different location. The original effect is however essentially kinematic. Here we show that the disembodied property has a dynamics of its own. Once disembodied from the particle to which it belongs, it can be subsequently affected by external actions even though the particle is not present. In particular, we describe the spatial propagation of such a property, specifically a flux of spin without its corresponding particle.

The motivation that led us to this effect is the desire to better understand "counterfactual computation"[2–8] and especially "counterfactual communication" quantum effects[9–24]. These are information processing tasks with a very strange property: the particles that are supposed to be the information carriers seem to not actually enter the information processing devices. The starting point of this class of phenomena was the discovery by Elitzur and Vaidman of the so-called "interaction-free measurement"[25], arguably one of the most striking effects in quantum physics. The basic set-up consists of an object—a "bomb" in the original example—that is ultra-sensitive to photons. Whenever a photon impinges on it, the bomb explodes. What Elitzur and Vaidman showed is that the bomb could nevertheless be investigated with photons without triggering it, if photons are prepared in a coherent superposition of impinging and not impinging on it.

Subsequent work[26] refined the original interaction-free measurement protocol, which had a failure probability (the bomb exploding) of 50% and raised the probability of success infinitesimally close to 1.

The ideas behind the interaction-free measurements have then been used to design information-processing protocols such as computation and communication, having the same basic characteristic, namely particles being affected by what happens in regions where they do not enter, similar to the probe in the interaction-free measurement. Ever since the discovery of interaction-free measurement and of the subsequent counterfactual information processing protocols there has been an intensive effort to understand what is behind these very puzzling effects.

In the present paper, inspired by the above, we construct a different set-up, that leads us to formulate the dynamic quantum Cheshire Cat effect. In this effect, a physical property can be disembodied from the particle to which it belongs, and can be subsequently affected by external actions even though the particle is not present. Clearly, this could be the key for understanding what happens in counterfactual information processing effects, where information can be accumulated, processed, stored and communicated without particles being present. We believe that in fact, this is the core of all such effects.

The main body of the paper is focused on the dynamic quantum Cheshire Cat effect itself, since it is important and interesting in itself, and its implications are likely to extend well beyond the counterfactual effects that motivated this research.

The connection with counterfactual communication will be discussed in the "Discussion" section.

Results

The set-up. The experiment at the core of our paper is deceptively simple. We will be interested in some particular cases of the following set-up, illustrated in Fig. 1. Consider a spin 1/2 particle in a box of length D. The left wall of the box is completely reflective; the right wall is spin-dependent: it is completely transparent for $|\uparrow_z\rangle$ and it completely reflects $|\downarrow_z\rangle$. For example, we can take the interaction Hamiltonian between the particle and the spin-dependent wall to be

$$\hat{H}_{\text{int}} = \frac{1 - \sigma_z}{2} V_0 \delta(x - x_R) \qquad (1)$$

where x denotes the location of the particle, x_R is the location of the right wall, and V_0 is very large, essentially infinite. This can be arranged physically by, say, having inside the wall a strong magnetic field along the z-axis, plus some other, constant, non-spin dependent, potential. In addition, the box is divided in the middle by a spin-independent partition wall through which the particle can tunnel, but whose transmission coefficient will be taken to be very small.

Suppose the particle is prepared initially in a small wavepacket moving with velocity v. The particle will move back and forth in the box, colliding with the box walls and the partition, and leaking out of the box via the right wall. For simplicity, we take the particle to have a large mass, and the velocity v high enough, so that during the entire duration of the experiment the spread of the wavepacket is negligible.

For the partition, we take the reflection and transmission amplitudes to be $\cos\epsilon$ and $\sin\epsilon$ respectively, with $\epsilon = \pi/2N$ and N integer. Eventually, we will take ϵ to be very small (N to be large).

The experiment. We are interested in what happens in the particular case when we start with the particle next to the left wall, moving towards the partition, a state that we denote $|L\rangle$, and being polarized $|\uparrow_z\rangle$. We are interested in what happens by time $t = 2NT$, where T is the time required to move the distance D, the length of the box, (which is also the time required to bounce once back and forth inside half of the box).

To start with, we note that $|\uparrow_z\rangle$ is a constant of motion: the only place where there is any spin-dependent interaction is at the right wall, and $|\uparrow_z\rangle$ is an eigenstate of that interaction. Hence the particle's spin remains $|\uparrow_z\rangle$ for all times. In this situation, the right wall is completely transparent for the particle.

The time evolution can easily be solved. It is convenient to consider the state of the particle at times $t_n = nT$ with n a positive integer. Apart from the state $|L\rangle$ defined above, other states that will be of interest for us are $|k\rangle$, with k a positive integer, which denotes the wavepacket situated outside the box, at a distance kD from its right end, moving away from the box. In particular, the

Fig. 1 The set-up. A spin 1/2 particle in a box with a completely reflective wall on the left, a spin-independent, partially transparent partition in the middle with, very small transmissivity, and a spin-dependent wall on the right, which is totally reflective for spin $|\uparrow_z\rangle$ and totally transparent for $|\downarrow_z\rangle$. The particle starts in a wavepacket $|L\rangle$, next to the left wall, moving right and with spin $|\uparrow_z\rangle$.
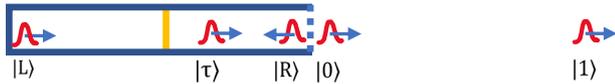
Fig. 2 States of interest. Various wavepackets mentioned in the experiment. The arrows denote the direction of movement of the wavepackets. These are various wavepackets that may result at various stages, when a particle starts in an initial wavepacket, $|L\rangle$ next to the left wall, moving towards right, and then is reflected by, or tunnels through, a semi-transparent partition in the middle or a spin-dependent wall at the right end of the box. $|R\rangle$ is a wavepacket inside the box, next to the right wall, moving towards left and $|0\rangle$ is a wavepacket next to the right wall but outside the box, moving to the right. The wavepacket $|1\rangle$ is a distance D, equal to the length of the box, away from the right wall; it is here where the wavepacket $|0\rangle$ will get after a time $T = D/v$. Other states $|2\rangle$, $|3\rangle$,... are situated further away, to the right, at intervals D from one another and are not illustrated in the figure. The wavepacket $|\tau\rangle$ represents a particle localized in the right half of the box, moving towards right, at a distance $v\tau$ from the left wall, with $T/2 \leq \tau < T$. Note that not all these wavepackets are simultaneously present during the experiment. Also, the spin degree of freedom, and the normalization with which various wavepackets may appear at various times in the experiment are not marked in the figure.

state $|0\rangle$ means the particle is just outside the right wall (see Fig. 2).

Let $\hat{U}$ denote the time evolution operator for a time T. We have

$$\hat{U}|L\rangle|\uparrow_z\rangle = \cos\epsilon|L\rangle|\uparrow_z\rangle + \sin\epsilon|0\rangle|\uparrow_z\rangle$$
$$\hat{U}|k\rangle|\uparrow_z\rangle = |(k+1)\rangle|\uparrow_z\rangle \qquad (2)$$

What happens is that when the particle starts in the state $|L\rangle$, next to the left wall, it moves towards the partition and collides with it at time $T/2$. As a result of this collision, with amplitude $\cos\epsilon$, the wavepacket is reflected, moves back towards the left wall, and at time T collides with it and is reflected, ending in the state $|L\rangle$, as it was initially, but with corresponding reduced amplitude, while with amplitude $\sin\epsilon$, the wavepacket is transmitted through the partition and at time T it emerges from the box, ending in state $|0\rangle$ with the corresponding reduced amplitude. On the other hand, if the particle is already outside the box, in state $|k\rangle$, then after time T it evolves to $|(k+1)\rangle$, a state further away from the box by the distance D. The operator $\hat{U}$ describes this entire process.

It is now trivial to calculate what happens at time nT, with $1 \leq n$, when $|\Psi(0)\rangle = |L\rangle|\uparrow_z\rangle$. The state becomes

$$|\Psi(nT)\rangle = \hat{U}^n|L\rangle|\uparrow_z\rangle$$
$$= \cos^n\epsilon|L\rangle|\uparrow_z\rangle + \sum_{k=0}^{n-1}\sin\epsilon\cos^k\epsilon|n-1-k\rangle|\uparrow_z\rangle \qquad (3)$$

After every collision with the partition, the wavepacket in the left side of the box loses a fraction of its amplitude, generating a wavepacket that tunnels to the right through the partition, and which then goes out of the box. Eventually, the particle will leak out of the box almost entirely, but this takes a long time, of order of N^2T.

Importantly for us, however, in the limit of small transmissivity (large N), for times of order NT the particle is essentially still on the left side, leaking out of the box only with an infinitesimal probability of order $O(1/N)$, i.e.,

$$\big|\Psi(2NT)\big\rangle = \hat{U}^{2N}|L\rangle|\uparrow_z\rangle = |L\rangle|\uparrow_z\rangle + |O(1/N)\rangle \qquad (4)$$

where $|O(1/N)\rangle$ denotes corrections of order $1/N$.

We can see this by directly calculating the probability of remaining on the left side, $\cos^{2N}\epsilon = \cos^{2N}(\pi/2N) = 1 - O(1/N)$ but it is actually illuminating to calculate instead the probability of leaving the box. The important feature is that the tunnelled wavepackets do not overlap with each other so when determining the total probability of leaving they add up in probability, not in amplitude

$$\sum_{k=0}^{N-1}(\sin\epsilon\cos^k\epsilon)^2 \leq \sum_{k=0}^{N-1}\sin^2\epsilon \approx N\epsilon^2 = \frac{\pi}{2}\epsilon = \frac{\pi^2}{4N} \qquad (5)$$

which goes to zero in the limit of small transmissivity (large N).

The paradox. The situation seems completely trivial: the particle bounces back and forth in the left half of the box, leaking infinitesimally small wavepackets via tunnelling, and its spin remains undisturbed, $|\uparrow_z\rangle$. Yet, at a closer look, we see that something quite interesting happens.

Consider the spin along x. Since the particle is prepared polarized in the z-direction, the x component is undefined: it is an equal superposition of $|\uparrow_x\rangle$ and $|\downarrow_x\rangle$, namely $|\uparrow_z\rangle = \frac{1}{\sqrt{2}}(|\uparrow_x\rangle + |\downarrow_x\rangle)$. Furthermore, since $|\uparrow_z\rangle$ is a constant of motion, this superposition of $|\uparrow_x\rangle$ and $|\downarrow_x\rangle$, remains the same at all times. However, each component changes with time. Specifically, as we will show, if at time $2NT$ we find the particle in the left half of the box—which we can arrange to happen with probability as close to 1 as we desire—the spin along the x-axis flips: $|\uparrow_x\rangle \to |\downarrow_x\rangle$ and $|\downarrow_x\rangle \to |\uparrow_x\rangle$.

Another way to look at the situation is to note that when we look at the x-spin observable, i.e., at the σ_x operator, its Heisenberg equation of motion tells us that if we find the particle in the left half of the box then $\sigma_x(t=2NT) = -\sigma_x(t=0)$ (see the proof in the next section).

The fact that there are situations in which the quantum state of a system does not change but nevertheless some observables change is actually extremely common, though, as far as we know, the very existence of this effect and its implications are not widely appreciated (for various discussions of this see refs. [27–29] where the notion of two-time observables was introduced and analyzed, and [30–34] where the disturbance due to measurements and its implications for the Heisenberg uncertainties are studied). This situation occurs every time when a particle is in an eigenstate of a non-zero Hamiltonian. This is not what is interesting in our example. What is interesting is how this happens in our case.

To put things in perspective, a trivial example is that of a spin 1/2 particle prepared in the state $|\uparrow_z\rangle$ and placed in a magnetic field oriented along the z-axis. In this case, the state is an eigenstate of the Hamiltonian and remains unchanged (up to an unobservable overall phase factor), but all the spin components in the plane orthogonal to z undergo Larmor precession, which we see if we solve their Heisenberg equations of motion. If we were to measure their values at any time, the outcomes of the measurement would be completely random, but if we look at the correlations between their values at different moments in time, we find that they change.

The above case is trivial, and we understand very well why these spin components change: the magnetic field acts upon them. And here lies our paradox: if at the end of the experiment we find the particle in the left-side of the box, the particle must have stayed there for the entire duration of the experiment. But there is no magnetic field there—the only place where there is a magnetic field is in the spin-dependent wall, at the right-end of the box. Indeed, up to terms whose total magnitude we can make as small as we want by decreasing the transmissivity of the

partition, the wavefunction remained localized in the left-side of the box for the entire duration of the experiment where there is no magnetic field. Moreover, even the infinitesimally small wave-packets that leaked out of the left-side of the box and reached the spin-dependent wall on the right could not contribute to what is happening on the left-side, as for them the spin-dependent wall is fully transparent, so they exit the box and never return. Nevertheless, the Heisenberg equations of motion tell that the x-component of the spin changes.

Even more surprising, looking solely at the solution of the Schrödinger equation, the situation in our experiment seems to be identical to the case in which there is no wall at all closing the right end of the box. Indeed, when initially the particle is prepared $|\uparrow_z\rangle$ the solutions of the Schrödinger equation with spin-dependent wall or no wall at all are identical, as the particle simply does not see the wall even if the wall is there. Yet, when the wall is present the x-spin component flips, while when there is no wall, it does not.

The time evolution of σ_x. In the previous section we simply claimed that by time $2NT$ the x-spin component flips; we will now prove it. Rather than attempting to calculate the evolution of σ_x by solving the Heisenberg equations of motion directly, it is simpler to go via an indirect route. Consider first what happens if instead of starting with the initial state $|L\rangle|\uparrow_z\rangle$ we would start with $|\Phi(0)\rangle = |L\rangle|\downarrow_z\rangle$. In this case, the particle sees the box closed on the right by a completely reflecting wall, hence the box is now closed at the two ends.

With $|R\rangle$ denoting the wavepacket inside the box, next to the right wall, and moving to the left, towards the partition, we have

$$\hat{U}|L\rangle|\downarrow_z\rangle = \cos\epsilon|L\rangle|\downarrow_z\rangle + \sin\epsilon|R\rangle|\downarrow_z\rangle \quad (6)$$

$$\hat{U}|R\rangle|\downarrow_z\rangle = \cos\epsilon|R\rangle|\downarrow_z\rangle - \sin\epsilon|L\rangle|\downarrow_z\rangle. \quad (7)$$

What happens is that the particle, starting in state $|L\rangle$, next to the left wall and moving towards the centre, traverses the left side of the box, and hits the partition at time $T/2$. Then it gets partially reflected by the partition, with amplitude $\cos\epsilon$, traverses back the left side of the box collides with the left wall and at time T it gets again in state $|L\rangle$, with the reduced amplitude $\cos\epsilon$, while with amplitude $\sin\epsilon$ the particle tunnels through the partition, traverses the right side of the box and gets reflected by the right wall, ending at time T in state $|R\rangle$ with the corresponding amplitude, $\sin\epsilon$. A similar thing happens when starting in the state $|R\rangle$. The operator $\hat{U}$ describes this whole evolution.

The evolution for a time nT leads to

$$\hat{U}^n|L\rangle|\downarrow_z\rangle = \cos(n\epsilon)|L\rangle|\downarrow_z\rangle + \sin(n\epsilon)|R\rangle|\downarrow_z\rangle. \quad (8)$$

$$\hat{U}^n|R\rangle|\downarrow_z\rangle = \cos(n\epsilon)|R\rangle|\downarrow_z\rangle - \sin(n\epsilon)|L\rangle|\downarrow_z\rangle. \quad (9)$$

In particular, we see that starting in the state $|L\rangle|\downarrow_z\rangle$ the particle oscillates between the left and right side of the box, with period $T_{osc} = 4NT = 2\pi T/\epsilon$. We are interested in what happens at time $2NT$, at half the period. By this time the particle would have first transitioned to the right side of the box, in state $|R\rangle$, where it got at time NT and then, at $t = 2NT$, the particle is again in the left side, where it started, but in state $-|L\rangle$, the state has acquired a negative sign. In other words,

$$\hat{U}^{2N}|L\rangle|\downarrow_z\rangle = -|L\rangle|\downarrow_z\rangle \quad (10)$$

Suppose now that we start with the state $|L\rangle|\uparrow_x\rangle$. We then have

$$\hat{U}^{2N}|L\rangle|\uparrow_x\rangle = \hat{U}^{2N}|L\rangle\frac{1}{\sqrt{2}}\left(|\uparrow_z\rangle + |\downarrow_z\rangle\right)$$

$$= |L\rangle\frac{1}{\sqrt{2}}\left(|\uparrow_z\rangle - |\downarrow_z\rangle\right) + |O(1/N)\rangle \quad (11)$$

$$= |L\rangle|\downarrow_x\rangle + |O(1/N)\rangle.$$

That is, if we start with spin $|\uparrow_x\rangle$, after time $2N$ it flips to $|\downarrow_x\rangle$, up to corrections of $O(1/N)$. Similarly,

$$\hat{U}^{2N}|L\rangle|\downarrow_x\rangle = |L\rangle|\uparrow_x\rangle + |O(1/N)\rangle. \quad (12)$$

Hence we can now obtain the Heisenberg evolution:

$$\hat{U}^{\dagger 2N}|L\rangle\langle L|\sigma_x\hat{U}^{2N} = |L\rangle\langle L|(-\sigma_x) + O(1/N). \quad (13)$$

On the other hand, when there is no wall closing the right side of the box, the evolution of $|\uparrow_z\rangle$ and $|\downarrow_z\rangle$ is identical, so they accumulate no phase difference; hence their superposition $|\uparrow_x\rangle = \frac{1}{\sqrt{2}}\left(|\uparrow_z\rangle + |\downarrow_z\rangle\right)$ remains $|\uparrow_x\rangle$ all the time, i.e., σ_x does not flip.

Observational consequences. As discussed in the above sections, the Heisenberg equations of motion show that when at time $t = 2NT$ we find the particle in the left half of the box—which is almost always when the transmission coefficient ϵ is infinitesi-mally small—the spin component in the x-direction flips from what it was at the start of the experiment. In other words, the correlation between the initial and final values of the x-spin components changes due to the presence of the spin-dependent wall. But that was a purely mathematical analysis—no measure-ments were considered that could verify the effect. We will do that now. The challenge is to observe this without disrupting the original set-up.

Of course, if we attempt to start by measuring σ_x at $t = 0$, in order to see how it changes by $t = 2NT$, we disturb the initial state $|\uparrow_z\rangle$ and generate a $|\downarrow_z\rangle$ component. This component then propagates up to the spin-dependent wall and bounces back from there, ending up back on the left side at $t = 2NT$, so there would be no surprise that the particle's spin feels the magnetic field. To avoid this, we start by measuring σ_x "weakly", i.e., in a manner that will produce as small disturbance as we desire. Of course, such a measurement cannot be precise, but repeating the experiment many times, we can extract the desired information. (A similar procedure, in the context of Heisenberg's measurement−disturbance relation, was considered in ref. [33] and was experimentally implemented in ref. [34]).

To this end, we use a measuring device with a pointer $\hat{q}$, prepared in the initial state (up to normalization)

$$\phi(q) = e^{-\frac{q^2}{4\Delta^2}}, \quad (14)$$

where Δ determines the uncertainty in the initial position of the pointer. We couple this test particle to our spin by the interaction Hamiltonian

$$\hat{H}_{test} = \lambda\delta(t)\hat{p}\sigma_x. \quad (15)$$

where $\hat{p}$ is the momentum conjugate to $\hat{q}$ and λ is a numerical constant, which has the role of a coupling constant. We can then make the disturbance as small as we want by taking λ appropriately small.

The time evolution operator that describes the measuring interaction is

$$\hat{V} = e^{-\frac{i}{\hbar}\lambda\hat{p}\sigma_x} \quad (16)$$

which is the shift operator that shifts q by the value $\lambda\sigma_x$. In other

words, the pointer $\hat{q}$ moves proportionally to the value of σ_x, with proportionality constant λ.

After measuring σ_x in this weakly disturbing way at $t = 0$, just after we prepared the initial state, we let the particle and measuring device evolve until $t = 2NT$. At that moment their state is

$$\hat{U}^{2N}\hat{V}|\uparrow_z\rangle|L\rangle e^{-\frac{q^2}{4\Delta^2}}. \qquad (17)$$

We then check to see if it is on the left side of the box and perform there an ideal (strong) measurement of σ_x. We are interested to see how the σ_x values measured at $t = 0$ and $t = 2NT$ correlate.

When the final measurement finds $\sigma_x = +1$, the state of the pointer q, used for measuring σ_x at $t = 0$ is

$$\langle L|\langle\uparrow_x|\hat{U}^{2N}\hat{V}|\uparrow_z\rangle|L\rangle e^{-\frac{q^2}{4\Delta^2}} = \frac{1}{\sqrt{2}}e^{-\frac{(q+\lambda\sigma_x^w)^2}{4\Delta^2}} + \mathcal{O}(\lambda^2) \qquad (18)$$

where we have used the first-order approximation of $\hat{V}$ in λ and where

$$\sigma_x^w = \frac{\langle L|\langle\uparrow_x|\hat{U}^{2N}\sigma_x|\uparrow_z\rangle|L\rangle}{\langle L|\langle\uparrow_x|\hat{U}^{2N}|\uparrow_z\rangle|L\rangle} \qquad (19)$$

is the so-called "weak" value of σ_x between the initial state $|\uparrow_z\rangle|L\rangle$ and the final state $\langle L|\langle\uparrow_x|\hat{U}^{2N}$ (which is the state corresponding to the result of the final measurement, i.e., that the particle was found at $t = 2NT$ in the left side of the box, with $\sigma_x = +1$, propagated backwards in time to $t = 0$). This is an instance of the general result that the state of the pointer used in a weak measurement of an observable $\hat{A}$ indicates the weak value

$$A^w = \frac{\langle\Psi_f|\hat{W}(t_f, t_0)\hat{A}\hat{W}(t_0, t_i)|\Psi_i\rangle}{\langle\Psi_f|\hat{W}(t_f, t_i)|\Psi_i\rangle} \qquad (20)$$

where $|\Psi_i\rangle$ and $|\Psi_f\rangle$ are the pre and post-selected states, t_i and t_f are the times of the initial preparation and post-selection respectively, t_0 is the time when the measurement occurred and $\hat{W}$ is the time evolution operator[35] (see also "Methods", subsection Weak measurements).

Using Eqs. (11) and (12) we find that

$$\sigma_x^w = -1 + \mathcal{O}\left(\frac{1}{N}\right). \qquad (21)$$

Putting all together, we see that when the measurement performed at $t = 2NT$ found the spin in the left side of the box, (which happens with probability close to 1) and $\sigma_x = +1$, the pointer q of the measuring device used to measure the σ_x at $t = 0$ shifted to the value $-\lambda$. Of course, with λ small, the shift in the position of the pointer q is smaller than its uncertainty, but repeating the experiment many times we are able to determine the shift with as much precision as desired.

As the pointer $\hat{q}$ moves proportionally to the value of σ_x, with proportionality constant λ, this means that the pointer indicates that at $t = 0$, the x-spin component was $\sigma_x = -1$, confirming our claim that the value of σ_x at $t = 0$ is opposite to that at the final time $t = 2NT$.

A similar conclusion also holds when at time $t = 2NT$ we found the spin in the left side of the box but with $\sigma_x = -1$.

The flux of spin without particles. We now come to a very interesting aspect of our problem. Spin is a conserved quantity as long as the particle moves in a region where there is no magnetic field. As we have seen, the x-component of the spin in the left half of the box changes, but the only place where there is a magnetic field is in the right end of the box, in the spin-dependent wall. But

how does it get there? Recall that the overall probability of the particle to reach the spin-dependent wall can be made as small as we want. Furthermore, even this infinitesimally small amount of cases cannot have any effect on what happens on the left side of the box, since in all the cases when the particle did reach the spin-dependent wall they leave the box and never return to the left side, as the time evolution (3) shows. Hence, if at the end of the experiment we find the particle in the left half of the box, it must have been there all the time.

The answer, we will show, is that there is a flux of spin without particles that carry it, in a version of the "Quantum Cheshire Cat" effect[1], in which properties can be disembodied from the particle that possesses them, like the grin from the famous cat.

Even more interesting, the spin flux originates from the left half of the box and propagates to the right, and not from the region of the wall towards the particle. In other words, the particle has to originate this flux. But while the particle is in the left half of the box, it has no knowledge whether or not at the right there is a spin-dependent wall or not, which raises the question of how does the particle knows when to originate the flux and when not?

Equation (3), of course, describes the evolution of the particle when no measurement to test it is performed. Again, to observe it, we will perform weak measurements that only infinitesimally disturb the particle, similarly to that described before.

To observe the details of the spin flow, we need to analyze the experiment in more detail. Up to this point, we only looked at times $t = nT$, multiples of the period of the bouncing back and forth inside one half of the box. We need now to look at more intermediate times.

Let $U(\tau)$, with $0 \leq \tau < T$ represent the time evolution for a time τ shorter than the period T.

Suppose that at time $t = nT + \tau$ we test whether the particle is on the right-hand half of the box, given that we find it on the left side at time $t = 2NT$, (which is, as we have shown, almost always). The times of interest are in the second part of each period (i.e., $T/2 \leq \tau < T$), which is when a tunnelled wavepacket is formed and traverses the right half of the box on its way out.

We already know that the amplitude to find the particle on the right-hand side after the n-th collision with the partition is $\sin\epsilon\cos^n\epsilon$, so very low to start with. But if we further condition it by the fact that at the final time $t = 2NT$ we find the particle on the left side, we now expect the probability to find it at the intermediate time in the right side should be strictly zero: indeed, given that the spin is $|\uparrow_z\rangle$, each tunnelled wavepacket subsequently emerges from the box and never comes back. The only way to find it on the left side at the final time is therefore not to have tunnelled at all. This intuition is confirmed by the weak measurement: if we weakly measure P_R, the projector on the right side at any arbitrary time $t = nT + \tau$ we find that the pointer indicates zero. Indeed, the pointer will be displaced to the "weak value" of P_R (see Eq. (20))

$$\hat{P}_R^w = \frac{\langle L|\langle\uparrow_x|\hat{U}^{2N-n-1}\hat{U}(T-\tau)\hat{P}_R\hat{U}(\tau)\hat{U}^n|\uparrow_z\rangle|L\rangle}{\langle L|\langle\uparrow_x|\hat{U}^{2N}|\uparrow_z\rangle|L\rangle} = 0. \qquad (22)$$

This weak value is equal to zero because $\hat{P}_R\hat{U}(\tau)\hat{U}^n|\uparrow_z\rangle|L\rangle$ is a wavepacket localized in the right half of the box, originated from tunnelling during the n-th collision and, by the final time $t = 2NT$ is outside of the box, i.e., up to a normalization factor,

$$\hat{U}^{2N-n-1}\hat{U}(T-\tau)\hat{P}_R\hat{U}(\tau)\hat{U}^n|\uparrow_z\rangle|L\rangle = |\uparrow_z\rangle|N-n\rangle \qquad (23)$$

and it is orthogonal to $|\uparrow_z\rangle|L\rangle$, hence the numerator in (22) is zero.

(In the above we assumed that the measurement of the spin at the end of the measurement yielded $\sigma_x = 1$; the same holds if this measurement yields $\sigma_x = -1$.)

Therefore, indeed, in the case that at the end we find the particle in the left half, there is no flux of particles traversing the right half of the box. Yet, its x-spin component changes. The spin difference must leak out somehow.

Let us ask now whether there is a spin in the right half of the box. This seems silly, as we have just proven that there are no particles there at any time. Yet, let us ask. If we perform a weak, very non-disruptive measurement of $\hat{P}_R \sigma_x$ at time $t = nT + \tau$ with $T/2 \le \tau < T$, and suppose that at the end of the experiment, a measurement of position and spin of the particle finds the particle in the left side of the box, with spin $\sigma_x = 1$, the measurement at $t = T + \tau$ yields (see "Methods", subsection Spin measurements).

$$
\begin{aligned}
(\hat{P}_R \sigma_x)^w \\
= \frac{\langle L| \langle \uparrow_x | \hat{U}^{2N-n-1} \hat{U}(T-\tau) \hat{P}_R \sigma_x \hat{U}(\tau) \hat{U}^n |\uparrow_z\rangle |L\rangle}{\langle L| \langle \uparrow_x | \hat{U}^{2N} |\uparrow_z\rangle |L\rangle} \\
= -\frac{\cos^n \epsilon \sin \epsilon \sin(2N - n - 1)\epsilon}{\cos^{2N} \epsilon}
\end{aligned}
\tag{24}
$$

which is different from zero, even though there is no particle there.

We can in fact study the evolution in more detail. Instead of just checking for the existence of the spin in the whole right half of the box, we check its precise position and how it changes in time. We can do this by measuring at time $t = nT + \tau$ the operator $\hat{P}_{R,\tau} \sigma_x$ where $\hat{P}_{R,\tau}$ is the projector on the location in the right half of the box where the tunnelled wavepacket is at time τ after the tunnelling occurred. We find that the entire spin in the right half of the box is actually concentrated in this region, i.e.,

$$
(\hat{P}_{R,\tau} \sigma_x)^w = (\hat{P}_R \sigma_x)^w
\tag{25}
$$

while if at the same time we measure the spin at other locations in the right half of the box we find zero.

Finally, we can also check for the existence of a particle and spin outside the box. A similar calculation to the one above, taking into account that the propagation of a wavepacket is spin-independent, $\hat{U}|k\rangle |\uparrow_z\rangle = |k+1\rangle |\uparrow_z\rangle$ and $\hat{U}|k\rangle |\downarrow_z\rangle = |k+1\rangle |\downarrow_z\rangle$, shows that, given that at the end of the experiment we find the particle in the left half of the box, there is neither a particle nor a spin outside the box at any time.

We, therefore, see that every period T, a pulse of spin propagates from the partition towards the spin-dependent wall and it is absorbed by the wall.

The total spin that left the left half of the box, propagated to the wall and got absorbed there can be found by summing the spin carried in each pulse during the experiment:

$$
\sum_{n=0}^{2N-1} \frac{\cos^n \epsilon \sin \epsilon \sin(2N - n - 1)\epsilon}{\cos^{2N} \epsilon} = -2 + \mathcal{O}\left(\frac{1}{N}\right)
\tag{26}
$$

(see "Methods", subsection Spin flux), meaning that the spin in the left half of the box increased by $2 + \mathcal{O}(\frac{1}{N})$, which is precisely the difference between the initial spin, which we argued is $\sigma_x = -1$ and the final spin, which we measured to be $\sigma_x = +1$.

Discussion

In this paper we analyzed a set-up that is at the core of a host of "counterfactual" phenomena. What all these phenomena have in common is that events in a given space region depend on what happens in a different space region (the "control" region) despite the particle having infinitesimally small probability of ever entering that region. What we have shown here is that this apparent paradox can be explained by the fact that it is not important for the particle to enter the control region; it is enough for the controlled property, (i.e., the property that is being controlled by actions in the control region), of the particle to enter that region. The presence of the controlled property, without the particle itself, is possible via a quantum Cheshire Cat effect.

As we mentioned in the "Introduction", our work was inspired by the desire to better understand the so-called counterfactual information processing protocols. The connection to counterfactual communication is immediate. Imagine the set-up described in our paper, with Alice next to the left-side of the box and Bob next to the right-side. Bob could communicate information to Alice by inserting or not inserting the spin-dependent wall. Say, when Bob wants to communicate a "0" he could leave the right-end of the box open, while when he wants to communicate a "1" he could insert the spin-dependent wall. In her turn, Alice could perform a weak measurement of σ_x at the start of the protocol and a strong measurement of σ_x at the end. In both cases (i.e., when Bob sends a "0" or a "1") the particle remains at all time next to Alice (except an infinitesimally small number of cases). Yet, by detecting if σ_x flipped or not, Alice can learn the bit sent by Bob. Hence this constitutes counterfactual communication. All of our analysis and insight can now be seen to apply to counterfactual communication. That is, even though the particle which was supposed to be the carrier is not sent through the channel, our results show that the disembodied information carrier, the spin, is in fact sent.

True, the above particular protocol, while exhibiting counterfactual communication, is quite inefficient. Indeed, in one round of the protocol, the weak measurement provides only very little information about the initial value of σ_x; to decide whether σ_x flips or not requires the protocol to be repeated many times, so the communication rate is very small. But this protocol is just a simple, direct, and unsophisticated application of our set-up, while efficient counterfactual communication protocols are far more complex. Analyzing those more complex examples in the light of the present insights is left for future work.

It is also interesting to compare the effect present in this paper with the so-called "negative-result measurements", a class of effects in which in some sense—different from ours—a particle is influenced by actions in a region where the particle is not found[36, 37]. The famous example by Dicke[37] is that of a particle in the ground state of a box. A measurement of position is made by sending a beam of light through the right half of the box. There is a probability of 50% of finding the particle there. The interesting case is when the light is not scattered, i.e., the particle is not found in the right half of the box. However, although the light is not scattered, the wavefunction of the particle is changed. Before the measurement the wavefunction was the ground state, which is spread on both sides of the box; after the measurement the wavefunction it is collapsed to the left side only. So although the particle was not found on the right side and the light beam was not scattered, the particle gained energy (the post-measurement state is a superposition of ground plus excited states, while initially it was ground state only). The difference between this effect and ours is that in the case of the "negative-result measurements" a significant part of the wavefunction is initially in the interaction region, (i.e., the right side of the box, where the beam of light was sent, in Dicke's example). Crucially, the size of the effect depends on the overlap of the wavefunction with the interaction region. Had the particle in Dicke's example been prepared in a state with small overlap with the right side of the box, the state would have been only a little perturbed by the beam of light; in the limit of zero initial overlap, the particle would have not been perturbed at all and the effect vanishes. In our case, the overlap of the wavefunction with the interaction region can be

made as close to zero as we want, and the effect remains. This shows the different nature of the two effects.

Finally, in this work, we present a dynamic Cheshire Cat effect. We would like to emphasize that our analysis is not instead of solving the Schrödinger equation, rather, it shows more details of the phenomena involved. In doing so we hope to achieve a better understanding of the true nature of quantum mechanics.

Methods

Weak measurements. The weakly disturbing measurement discussed here was first introduced in ref. [35]. We give the main result here, for convenience.

In order to measure an arbitrary observable $\hat{A}$ at time $t = t_0$ we follow the von Neumann measuring formalism. We consider a measuring device with a pointer $\hat{q}$ prepared in the initial state (up to normalization) $\phi(q) = e^{-\frac{q^2}{4\Delta^2}}$ where Δ is the uncertainty in the initial position of the pointer, and couple the measuring device to the measured system via the interaction Hamiltonian

$$\hat{H} = \lambda\delta(t - \tau)\hat{A}\hat{p} \qquad (27)$$

where $\hat{p}$ is the momentum canonically conjugate to the pointer position $\hat{q}$. We will be interested in the case in which the coupling constant λ is small. In the rest of the time, the measuring device is supposed to remain unchanged, i.e., to have zero Hamiltonian. The unitary evolution corresponding to the interaction is

$$\hat{V} = e^{-i\lambda\hat{A}\hat{p}} \qquad (28)$$

which is the shift operator, which shifts the pointer by the value of $\hat{A}$.

Here we are interested in performing the measurement on a system that was prepared at the initial time $t = t_i$ in the initial state $|\Psi_i\rangle$ and which, except for $t = t_0$ when the measurement is performed evolves under the time evolution operator $\hat{W}$. Furthermore, we are interested in the result of this weak measurement in the case in which a second, later measurement, of some other operator B, taking place at $t = t_f$ happened to yield the eigenvalue corresponding to its eigenstate $|\Psi_f\rangle$.

The system evolves from $t = t_i$ to $t = t_0$ under the evolution operator $\hat{W}(t_0, t_i)$. Then, the state of the system and measuring device immediately after the measurement is given by

$$e^{-i\lambda\hat{A}\hat{p}}\hat{W}(t_0, t_i)|\Psi_i\rangle\left|e^{-\frac{q^2}{4\Delta^2}}\right\rangle \qquad (29)$$

After the further evolution from t_0 to t_f and projecting on $|\Psi_f\rangle$, to select the cases in which the second measurement yielded the result of interest for us, we find the first measuring device to be left in the state

$$\left\langle\Psi_f\left|\hat{W}(t_f, t_0)e^{-i\lambda\hat{A}\hat{p}}\hat{W}(t_0, t_i)\right|\Psi_i\right\rangle\left|e^{-\frac{q^2}{4\Delta^2}}\right\rangle$$
$$\approx \left\langle\Psi_f\left|\hat{W}(t_f, t_0)(1 - i\lambda\hat{A}\hat{p})\hat{W}(t_0, t_i)\right|\Psi_i\right\rangle\left|e^{-\frac{q^2}{4\Delta^2}}\right\rangle$$
$$= \left\langle\Psi_f\left|\hat{W}(t_f, t_i)\right|\Psi_i\right\rangle(1 - i\lambda A^w\hat{p})\left|e^{-\frac{q^2}{4\Delta^2}}\right\rangle \qquad (30)$$
$$\approx \left\langle\Psi_f\left|\hat{W}(t_f, t_i)\right|\Psi_i\right\rangle e^{-i\lambda A^w\hat{p}}\left|e^{-\frac{q^2}{4\Delta^2}}\right\rangle$$
$$= \left\langle\Psi_f\left|\hat{W}(t_f, t_i)\right|\Psi_i\right\rangle\left|e^{-\frac{(q-\lambda A^w)^2}{4\Delta^2}}\right\rangle$$

where in the second line we have approximated the shift operator $e^{-i\lambda\hat{A}\hat{p}}$ to the first order in λ, in the third line we have factored out $\langle\Psi_f|\hat{W}(t_f, t_0)\hat{W}(t_0, t_i)|\Psi_i\rangle$, and we have used the fact that $\hat{W}(t_f, t_0)\hat{W}(t_0, t_i) = \hat{W}(t_f, t_i)$ and where

$$A^w = \frac{\left\langle\Psi_f\left|\hat{W}(t_f, t_0)\hat{A}\hat{W}(t_0, t_i)\right|\Psi_i\right\rangle}{\left\langle\Psi_f\left|\hat{W}(t_f, t_i)\right|\Psi_i\right\rangle}, \qquad (31)$$

in the fourth line, we used again the first-order approximation of the shift operator and in the last line, we applied the shift operator to the state of the pointer. We have this way proved that, for small enough λ, the measuring device indicates the value A^w proving thus (20). (Note that the pre-factor $\langle\Psi_f|\hat{W}(t_f, t_i)|\Psi_i\rangle$ is simply a normalization factor, corresponding to the probability of the second measurement to yield $|\Psi_f\rangle$.)

Spin measurements. Let us evaluate $(\hat{P}_R\sigma_x)^w$, Eq. (24). To evaluate the numerator, we first note that the projector $\hat{P}_R$ measured at time $nT + \tau$ selects the wavepacket that is at this time in the right side of the box. This wavepacket originated from

tunnelling in the $n + 1$ collision. Hence

$$\hat{P}_R\hat{U}(\tau)\hat{U}^n|\uparrow_z\rangle|L\rangle = \cos^n\epsilon\sin\epsilon|\tau\rangle \qquad (32)$$

where $|\tau\rangle$ represents this wavepacket (see Fig. 2).

Next, the operator σ_x applied to the $|\uparrow_z\rangle$ spin flips it to $|\downarrow_z\rangle$, which sees the wall. The time evolution operator $\hat{U}(T - \tau)$ further propagates $|\tau\rangle$ up to the right wall, where, due to the spin being $|\downarrow_z\rangle$, it gets reflected: $\hat{U}(T - \tau)|\tau\rangle|\downarrow_z\rangle = |R\rangle|\downarrow_z\rangle$ leading to

$$\hat{U}(T - \tau)\sigma_x\hat{P}_R\hat{U}(\tau)\hat{U}^n|\uparrow_z\rangle|L\rangle = \cos^n\epsilon\sin\epsilon|R\rangle|\downarrow_z\rangle. \qquad (33)$$

The evolution for the next $(N - n - 1)T$, given by $\hat{U}^{2N-n-1}$ can be computed immediately, using (9), and we obtain

$$\hat{U}^{2N-n-1}\hat{U}(T - \tau)\hat{P}_R\sigma_x\hat{U}(\tau)\hat{U}^n|\uparrow_z\rangle|L\rangle$$
$$= \cos^n\epsilon\sin\epsilon(\cos(2N - n - 1)\epsilon|R\rangle - \sin(2N - n - 1)\epsilon|L\rangle)|\downarrow_z\rangle \qquad (34)$$

Finally, projecting on $|L\rangle|\uparrow_x\rangle$ yields

$$\langle\uparrow_x|\langle L|\hat{U}^{2N-n-1}\hat{U}(T - \tau)\hat{P}_R\sigma_x\hat{U}(\tau)\hat{U}^n|L\rangle|\uparrow_z\rangle$$
$$= -\frac{1}{\sqrt{2}}\cos^n\epsilon\sin\epsilon\sin(2N - n - 1)\epsilon \qquad (35)$$

The denominator is, using (3),

$$\langle L|\langle\uparrow_x|\hat{U}^{2N}|\uparrow_z\rangle|L\rangle$$
$$= \langle L|\langle\uparrow_x|\cos^{2N}\epsilon|L\rangle|\uparrow_z\rangle$$
$$+ \langle L|\langle\uparrow_x|\sum_{k=0}^{2N-1}\sin\epsilon\cos^k\epsilon|2N - 1 - k\rangle|\uparrow_z\rangle \qquad (36)$$
$$= \frac{1}{\sqrt{2}}\cos^{2N}\epsilon$$

From (35) and (36) we get (24):

$$(\hat{P}_R\sigma_x)^w = \frac{\langle L|\langle\uparrow_x|\hat{U}^{2N-n-1}\hat{U}(T - \tau)\hat{P}_R\sigma_x\hat{U}(\tau)\hat{U}^n|\uparrow_z\rangle|L\rangle}{\langle L|\langle\uparrow_x|\hat{U}^{2N}|\uparrow_z\rangle|L\rangle}$$
$$= -\frac{\cos^n\epsilon\sin\epsilon\sin(2N - n - 1)\epsilon}{\cos^{2N}\epsilon} \qquad (37)$$

Spin flux. Here we calculate the spin flux, Eq. (26):

$$\sum_{n=0}^{2N-1} -\frac{\cos^n\epsilon\sin\epsilon\sin(2N - n - 1)\epsilon}{\cos^{2N}\epsilon}$$
$$= -\sin\epsilon\sum_{n=0}^{2N-1}\sin(2N - n - 1)\epsilon + \mathcal{O}\left(\frac{1}{N}\right) \qquad (38)$$

where we approximated $\cos^n\epsilon \approx 1$ for $0 \le n \le 2N$. Given that $\epsilon = \frac{\pi}{2N}$ we get

$$-\sin\left(\frac{\pi}{2N}\right)\sum_{n=0}^{2N-1}\sin\left((2N - n - 1)\frac{\pi}{2N}\right)$$
$$= -\sin\left(\frac{\pi}{2N}\right)\sum_{n=0}^{2N-1}\sin\left((n + 1)\frac{\pi}{2N}\right)$$
$$= -\sin\left(\frac{\pi}{2N}\right)\cot\left(\frac{\pi}{4N}\right) = -2\cos^2\left(\frac{\pi}{4N}\right) \qquad (39)$$
$$= -2 + \mathcal{O}\left(\frac{1}{N}\right)$$

where we again approximated $\cos^2(\frac{\pi}{4N})$ to 1 up to order $1/N$.

Data availability

No data sets were generated or analyzed during the current study.

References

1. Aharonov, Y., Popescu, S., Rohrlich, D. & Skrzypczyk, P. Quantum Cheshire Cats. *N. J. Phys.* **15**, 113015 (2013).
2. Jozsa, R. in *Lecture Notes in Computer Science,* Vol. 1509 (ed Williams, C. P.) 103 (Springer, 1998).
3. Mitchison, G. & Jozsa, R. Counterfactual computation. *Proc. R. Soc. Lond. A* **457**, 1175 (2001).
4. Hosten, O. et al. Counterfactual quantum computation through quantum interrogation. *Nature* **439**, 949 (2006).

5. Vaidman, L. The Impossibility of the counterfactual computation for all possible outcomes. *Phys. Rev. Lett.* **98**, 160403 (2007).

6. Mitchison, G. and Jozsa, R. The limits of counterfactual computation. Preprint at https://arxiv.org/abs/quant-ph/0606092 (2006).

7. Guo, G.-C. & Shi, B.-S. Quantum cryptography based on interaction-free measurement. *Phys. Lett. A* **256**, 109 (1999).

8. Noh, T.-G. Counterfactual quantum cryptography. *Phys. Rev. Lett.* **103**, 230501 (2009).

9. Salih, H., Li, Z.-H., Al-Amri, M. & Zubairy, M. S. Protocol for direct counterfactual quantum communication. *Phys. Rev. Lett.* **110**, 170502 (2013).

10. Salih, H. Protocol for counterfactually transporting an unknown qubit. *Front. Phys.* **3**, 94 (2016).

11. Griffiths, R. B. Particle path through a nested Mach−Zehnder interferometer. *Phys. Rev. A* **94**, 032115 (2016).

12. Vaidman, L. Comment on "Particle path through a nested Mach−Zehnder interferometer". *Phys. Rev. A* **95**, 066101 (2017).

13. Griffiths, R. B. Reply to "Comment on 'Particle path through a nested Mach−Zehnder interferometer' ". *Phys. Rev. A* **95**, 066102 (2017).

14. Vaidman, L. Counterfactuality of "counterfactual" communication. *J. Phys. A* **48**, 465303 (2015).

15. Vaidman, L. Comment on "Protocol for direct counterfactual quantum communication". *Phys. Rev. Lett.* **112**, 208901 (2014).

16. Salih, H. et al. Reply. *Phys. Rev. Lett.* **112**, 208902 (2014).

17. Danan, A., Farfurnik, D., Bar-Ad, S. & Vaidman, L. Asking photons where they have been. *Phys. Rev. Lett.* **111**, 240402 (2013).

18. Arvidsson-Shukur, D. R. M., Gottfries, A. N. O. & Barnes, C. H. W. Evaluation of counterfactuality in counterfactual communication protocols. *Phys. Rev. A* **96**, 062316 (2017).

19. Vaidman, L. Past of a quantum particle. *Phys. Rev. A* **87**, 052104 (2013).

20. Salih, H. Comment on "Particle path through a nested Mach−Zehnder interferometer". *Phys. Rev. A* **97**, 026101 (2018).

21. Griffiths, R. B. Reply to "Comment on 'Particle path through a nested Mach−Zehnder interferometer' ". *Phys. Rev. A* **97**, 026102 (2018).

22. Aharonov, Y. & Vaidman, L. Modification of counterfactual communication protocols that eliminates weak particle traces. *Phys. Rev. A* **99**, 010103 (2019).

23. Aharonov, Y. & Rohrlich, D. What is nonlocal in counterfactual quantum communication? *Phys. Rev. Lett.* **125**, 260401 (2020).

24. Vaidman, L. Analysis of counterfactuality of counterfactual communication protocols. *Phys. Rev. A* **99**, 052127 (2019).

25. Elitzur, A. C. & Vaidman, L. Quantum mechanical interaction-free measurements. *Found. Phys.* **23**, 987 (1993).

26. Kwiat, P. et al. Interaction-free measurement. *Phys. Rev. Lett.* **74**, 4763 (1995).

27. Aharonov, Y. & Albert, D. Z. Is the usual approach to time evolution adequate: I. Nonrelativistic consideration. *Phys. Rev. D* **29**, 223 (1984). Aharonov, Y., Albert, D. Z., & D'Amato, S. S. Multiple-time properties of quantum mechanical systems. *Phys. Rev. D* **32**, 1975 (1985).

28. Aharonov, Y., Popescu, S., Tollaksen, S. J. & Vaidman, L. Multiple-time states and multiple-time measurements in quantum mechanics. *Phys. Rev. A* **79**, 052110 (2009).

29. Aharonov, Y. & Popescu, S., Tollaksen. In *Quantum Theory: A Two-Time Success Story* (eds Struppa, D. & Tollaksen, J.) Each Instant of Time a New Universe. (Springer, 2014).

30. Ozawa, M. Uncertainty relations for noise and disturbance in generalized quantum measurements. *Ann. Phys.* **311**, 350 (2004).

31. Ozawa, M. S672 Universal uncertainty principle in the measurement operator formalism. *J. Opt. B: Quantum Semiclass. Opt.* **7**, S672 (2005).

32. Sulyok, G. et al. Violation of Heisenberg's error-disturbance uncertainty relation in neutron-spin measurements. *Phys. Rev. A* **88**, 022110 (2013).

33. Lund, A. P. & Wiseman, H. M. Measuring measurement−disturbance relationships with weak values. *N. J. Phys.* **12**, 093011 (2010).

34. Rozema, L. A. et al. Violation of Heisenberg's measurement−disturbance relationship by weak measurements. *Phys. Rev. Lett.* **109**, 100404 (2012).

35. 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. D* **60**, 1351 (1988). Aharonov, Y. & L. Vaidman, L. Complete description of a quantum system at a given time. *J. Phys. A: Math. Gen.* **24**, 2315 (1991).

36. Renninger, M. Messungen ohne Störung des Messobjekts. *Z. Phys.* **158**, 417 (1960).

37. Dicke, R. H. Interaction-free measurement—a paradox. *Am. J. Phys.* **49**, 925 (1981).

Acknowledgements

Author contributions

Y.A. conceived the main idea, which was further developed by the three co-authors, Y.A, E.C, and S.P. The paper was written by S.P. with contributions from Y.A. and E.C.

Competing interests

The authors declare no competing interests.

Additional information