



Rubino, G., Rozema, L. A., Massa, F., Araújo, M., Zych, M., Brukner, Č., & Walther, P. (2022). Experimental entanglement of temporal order. *Quantum*, 6. <https://doi.org/10.22331/q-2022-01-11-621>

Peer reviewed version

License (if available):  
CC BY

Link to published version (if available):  
[10.22331/q-2022-01-11-621](https://doi.org/10.22331/q-2022-01-11-621)

[Link to publication record in Explore Bristol Research](#)  
PDF-document

This is the accepted author manuscript (AAM). The final published version (version of record) is available online via Verein zur Förderung at <https://doi.org/10.22331/q-2022-01-11-621>. Please refer to any applicable terms of use of the publisher

## University of Bristol - Explore Bristol Research

### General rights

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available: <http://www.bristol.ac.uk/red/research-policy/pure/user-guides/ebr-terms/>

# Experimental entanglement of temporal order

Giulia Rubino<sup>1,2</sup>, Lee A. Rozema<sup>1</sup>, Francesco Massa<sup>1</sup>, Mateus Araújo<sup>1,3</sup>, Magdalena Zych<sup>4</sup>, Časlav Brukner<sup>1,3</sup>, and Philip Walther<sup>1</sup>

<sup>1</sup>Vienna Center for Quantum Science and Technology (VCQ), Faculty of Physics, University of Vienna, Boltzmannngasse 5, Vienna A-1090, Austria

<sup>2</sup>Quantum Engineering Technology Labs, H. H. Wills Physics Laboratory and Department of Electrical & Electronic Engineering, University of Bristol, Bristol BS8 1FD, United Kingdom

<sup>3</sup>Institute for Quantum Optics & Quantum Information (IQOQI), Austrian Academy of Sciences, Boltzmannngasse 3, Vienna A-1090, Austria

<sup>4</sup>Centre for Engineered Quantum Systems, School of Mathematics and Physics, The University of Queensland, St Lucia, QLD 4072, Australia

**The study of causal relations has recently been applied to the quantum realm, leading to the discovery that not all physical processes have a definite causal structure. While indefinite causal processes have previously been experimentally shown, these proofs relied on the quantum description of the experiments. Yet, the same experimental data could also be compatible with definite causal structures within different descriptions. Here, we present the first demonstration of indefinite temporal order outside of quantum formalism. We show that our experimental outcomes are incompatible with a class of generalised probabilistic theories satisfying the assumptions of locality and definite temporal order. To this end, we derive physical constraints (in the form of a Bell-like inequality) on experimental outcomes within such a class of theories. We then experimentally invalidate these theories by violating the inequality using entangled temporal order. This provides experimental evidence that there exist correlations in nature which are incompatible with the assumptions of locality and definite temporal order.**

## 1 Introduction

Bell's theorem revolutionized the foundations of physics, leading to experiments which could

Giulia Rubino: giulia.rubino@bristol.ac.uk

demonstrate that nature cannot be described by a local-causal theory, and paving the way for modern quantum information [8, 12]. One of the strengths of Bell's theorem is that it allows one to draw conclusions about nature without referring to the underlying physical theory. This is a crucial feature, since "local causality" is conceived as a hypothesis about a fundamental property of nature, and, as such, its experimental violation reveals a statement about how nature must (or must not) be, and not just about any specific theory used to describe it.

Over the past decades, tests of Bell's theorem have been performed with many different physical systems thereby entangling various observables [such as spin [25, 28, 39], polarization [6, 19, 20, 41], position [26], and energy [27, 37]] of two or more particles. All these experimental tests involve two or more parties making measurements on entangled particles in distant laboratories, and determining their correlations. These tests prove that, in general, correlations between space-like separated events cannot be explained on the basis of a past common cause and local choices of measurements. However, since these correlations are always no-signalling (i.e., the statistics observed in one laboratory are independent of the choices of measurements made in other laboratories), Bell's theorem and related experimental tests do not address specific causal relations between events in which an experimental intervention may influence one of the events (i.e., signalling correlations). This stronger notion of causality [11] is at the heart of the modern research field of indefinite quantum causality, and is the

arXiv:1712.06884v5 [quant-ph] 29 Dec 2021

subject of the present experimental test.

Thus far, in all established physical theories, it was assumed that the order between events is well-defined. This means that, for any two causally related events  $A$  and  $B$ , either  $A$  signals to  $B$ , or  $B$  signals to  $A$ . In the language of quantum information, these two events could be seen as the input and output of a quantum channel. However, it has recently been realized that quantum mechanics also allows for the existence of processes other than quantum channels, i.e., processes that are neither causally ordered, nor a probabilistic mixture of causally ordered processes. For these processes, it is genuinely indefinite whether  $A$  signals to  $B$  or  $B$  signals to  $A$ , and so they are called processes with an indefinite causal structure. Under “processes,” we define the set of causal relations between operations performed in different local laboratories [11, 14, 35]. More precisely, a quantum process is called causally separable if it can be decomposed as a convex combination of causally ordered processes, otherwise it is causally non-separable. (Note that the term “temporal” order is used here to refer to the order among operations which cannot be used to receive signals — in particular, to unitary ones — whereas “causal” order refers to more general operations which allow both receiving and sending signals between laboratories.)

Recently, a method for certifying causal non-separability, based on “causal witnesses,” was developed [5, 9, 34], and used to experimentally demonstrate that a certain process — a quantum-switch [13] — is causally non-separable [21, 40]. In the quantum-switch, a qubit is transmitted between two parties, and the order in which the parties receive and act on it is entangled with a second system. This can result in a scenario in which operations are applied on the system in a quantum superposition of different temporal orders. The existence of such a superposition has been experimentally demonstrated [21, 36, 40]. However, the certification of this “indefiniteness” of temporal orders was theory-dependent, requiring the assumption that the system under investigation and the applied operations were described by quantum theory. In more detail, Ref. [21, 40] reported the measurement of a value for a causal witness that could not be explained by any model making the following three assumptions: (a) there was

a definite causal order between the parties, (b) each party acted only once, and (c) their operations are described by quantum theory. Nevertheless, the results of these experiments could potentially have also been explained in accordance with hypotheses (a) and (b) within a different theory (i.e., outside the quantum theory). Thus, the nature of indefinite causal order has not yet been probed without the use of quantum formalism to describe it.

In addition to theory-dependent causal witnesses, there are also device-independent ways of certifying indefinite causal order via “causal inequalities” [10, 35]. These inequalities only require measuring the probabilities of outcomes for different parties in the process under consideration without knowledge of the internal functioning of the devices. Any probabilities that show signalling in only one direction — which can be interpreted as an influence from the past to the future —, or that is a convex mixture of processes which allow signalling only in one direction (from  $A$  to  $B$  or from  $B$  to  $A$ ), satisfy causal inequalities. Nevertheless, it can be shown that the quantum-switch satisfies all such causal inequalities [see Refs. [5, 34] or the Suppl. Information for details], and, currently, it is not known whether or not it is possible to realize a process which violates a causal inequality. The question then arises if it is at all possible to prove the existence of an indefinite causal order in a manner that does not rely on the quantum description of the experiment. This is a relevant question, since such an experimental verification of indefinite causality would show that this is *not* a feature of a particular theory, but a *fundamental property* of a whole class of theories which can be used to describe nature.

In this work, we answer the above question affirmatively by presenting an experimental verification of Bell’s theorem for temporal order, which is formulated outside of the quantum framework. To this end, we generalize a Bell inequality for temporal order [43], and then experimentally violate it. The Bell inequality is shown to be fulfilled in a class of so-called “*generalized probabilistic theories*” in which the states and the laboratory operations are local, and the operations are applied in a definite order. The experimental violation of the Bell inequality presented here demonstrates, independent of quantum formalism, that

there exist correlations in nature which are incompatible with a class of theories assuming the order of events as locally pre-defined.

Finally, we notice that, while our inequality is valid for a class of generalized probabilistic theories, it does depend on the internal functionality of experimental devices, and in this sense it does not have the same “device-independent” status as the original Bell’s theorem. Thus, our work provides a proof of indefinite causality which is weaker than a violation of causal inequalities (i.e., a device-independent proof), but stronger than a measurement of a causal witnesses (a theory-dependent proof). It remains an open question, even from a theoretical viewpoint, whether it is possible to provide stronger evidence of indefinite causality than a violation of the present Bell inequality for temporal order.

## 2 No-go theorem for definite temporal order

We now introduce a *no-go theorem* for definite temporal order that applies to a class of generalized probabilistic theories (GPTs) in which the order of local events is assumed to be pre-defined. GPTs are a general framework that specifies a set of operations which can be applied on physical systems, assigns probabilities to experimental outcomes [16, 23, 24, 31], and which encompasses all operational theories – including classical probability theory and quantum theory as special cases. The no-go theorem which we present here was previously derived in the context of gravity [43]. Our derivation uses an assumption about the initial state of the systems which is weaker than that in Ref. [43] (we consider Bell-local states rather than separable states, which are a subset of Bell-local states), and a different notion of locality. (The relation between the assumptions and implications of the current work and those of Ref. [43] are analyzed in Appendix A-C.)

We first define what we mean by a causal order in a GPT. Consider a system in the state  $\omega \in \Omega$  of a GPT state space  $\Omega$  and imagine two parties, Alice and Bob, who perform some operations on this state. For example, suppose that the operation in Alice’s laboratory is given by a transformation  $\mathcal{A}$  and that in Bob’s laboratory is given by a transformation  $\mathcal{B}$ . Alice’s and Bob’s operations

are said to undergo a process that is “causally separable” in GPTs whenever Alice’s operation happens before or simultaneously to Bob’s ( $\mathcal{A} \preceq \mathcal{B}$ ), Bob’s operation happens before or simultaneously to Alice’s ( $\mathcal{B} \preceq \mathcal{A}$ ), or there is a convex mixture of these two cases:

$$\mathcal{S}(\omega) = \zeta \cdot \mathcal{B}(\mathcal{A}(\omega)) + (1 - \zeta) \cdot \mathcal{A}(\mathcal{B}(\omega)), \quad (1)$$

where  $0 \leq \zeta \leq 1$  is the probability with which one or the other order is chosen and  $\mathcal{Y}(\mathcal{X}(\cdot))$  is a composition of operations  $\mathcal{X}$  and  $\mathcal{Y}$ . (While in the current work we limit our analysis to the case of only  $N = 2$  parties, an analogue relation can be established for  $N > 2$  parties, giving rise to a classical mixture of all possible permutations among the  $N$  parties, or to a dynamical causal order, where the causal order between operations may depend on operations performed beforehand [2].) If a process cannot be written in the form of Eq. (1), it is called a “causally non-separable process.”

Within the GPT framework, we now consider  $\omega$  to be a state of the following composite system: one system (the *control* system) governing the order in which the operations  $\mathcal{A}$  and  $\mathcal{B}$  are applied, and another system (the *target* system) on which the operations are performed. We will further consider that there are two parties, S1 and S2, each possessing one such composite system. No restrictions are applied to the state of the control system (thus, for instance, the composite control state may violate a Bell inequality).

In Appendix A we prove a *no-go theorem*, stating that any two-party system obeying the following three assumptions cannot violate a Bell inequality (below we briefly summarize our theorem, saving the detailed version for Appendix A).

- I) The initial joint state of the two target subsystems is *Bell-local* (i.e., it satisfies the Bell’s “local-causality” condition for any pair of local measurements).
- II) The laboratory operations preserve Bell-locality of the state of the target subsystems (i.e., when applied to a Bell-local state of the target systems, they produce a Bell-local state.).
- III) The order of local operations on the two target subsystems is well-defined.

We will briefly comment on assumptions I and II below, and refer to Sec. 4 for an in-depth analysis of all three assumptions.

In our experiment and for the class of GPTs considered, we can demonstrate stronger conditions than I and II. Indeed, the initial joint state of the targets is separable within the GPTs; and the laboratory transformations are local maps on the target systems within the GPTs (i.e., they transform separable states into separable states). Then, I and II follow from these stronger conditions since any separable state is Bell-local. We could have formulated I and II in terms of separable states of the GPTs, but decided to leave them in the present form, since Bell’s locality is a weaker condition than separability, and the formulation does not involve theory-dependent notions.

There are two forms of violation of Assumption II: IIa) The laboratory operations can induce non-local interaction between the target subsystems of parties S1 and S2. IIb) Within a single party  $S_i$ ,  $i = 1, 2$ , the laboratory operations may “couple” the control and target system. Such a “coupling” could transfer existing non-local correlations between the pair of controls to the pair of targets, thereby enabling a violation of Bell inequalities. We will provide experimental evidence that neither case occurs in our experiment.

In the next section, we will present a quantum mechanical process that violates this no-go theorem. Thus, at least one of the assumptions must not hold for this process. In Sec. 4, we will analyse our experimental data testing a Bell-like inequality to provide evidence in support of assumption I within the framework of GPTs. Consequently, either assumption II does not hold, assumption III does not hold, or both assumptions are invalid. On the basis of the data collected for the quantum-switch of system S1 (or S2) individually, we will show that it is not possible to describe our results by violating only assumption II. Thus, the only viable conclusion is that the order of operations applied on each system  $S_i$  is indefinite (i.e., that assumption III is necessarily false).

### 3 Entangled quantum-switch

To understand a single quantum-switch, first imagine two parties, Alice and Bob, who are in

two *closed laboratories*, i.e., their only interaction with the external environment is through input and output systems. Suppose that each of the parties performs an operation on the same qubit (a “target” qubit), and that this qubit may be sent first to Alice and then to Bob, or vice versa. Now, in a quantum-switch, one governs the order of the operations on the target qubit according to the state of a second quantum system, a “control” qubit. If the control qubit is placed in a superposition, this establishes a quantum-superposition of the order of the two operations. For instance, if the control qubit is in the state  $|0\rangle^c$ , the target qubit is sent first to Alice and then to Bob, and vice versa if the control qubit is in the state  $|1\rangle^c$ . When the control qubit is prepared in the state  $(|0\rangle^c + |1\rangle^c)/\sqrt{2}$ , the resulting process has been shown to be causally non-separable within quantum mechanics [5, 13, 14, 40].

Next, consider two quantum-switches (S1 and S2), each containing an Alice and a Bob. S1 and S2 are prepared in a state where their control qubits are entangled, but their target qubits are in a product state (see Fig. 1):

$$|0\rangle_1^t \otimes |0\rangle_2^t \otimes \left( \frac{|0\rangle_1^c \otimes |0\rangle_2^c - |1\rangle_1^c \otimes |1\rangle_2^c}{\sqrt{2}} \right). \quad (2)$$

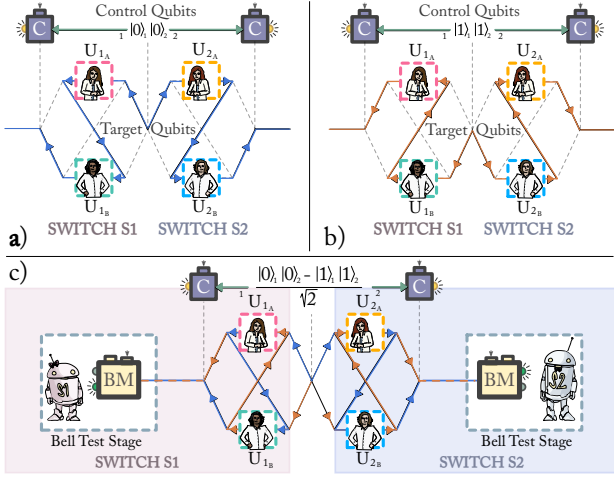
The superscripts  $c$  and  $t$  refer to the control and target qubits within one quantum-switch, respectively, while the subscripts 1 and 2 refer to quantum-switch S1 and S2. Since we will attempt to observe a Bell violation with the target qubits, which are in a separable state, this initial condition satisfies assumption I in quantum theory.

Given this input state and the action of an individual quantum-switch, it is straightforward to calculate the output of the entangled quantum-switch system

$$\begin{aligned} & \frac{1}{\sqrt{2}} \left( U_{1B} U_{1A} |0\rangle_1^t \right) \otimes |0\rangle_1^c \otimes \left( U_{2B} U_{2A} |0\rangle_2^t \right) \otimes |0\rangle_2^c \\ & - \frac{1}{\sqrt{2}} \left( U_{1A} U_{1B} |0\rangle_1^t \right) \otimes |1\rangle_1^c \otimes \left( U_{2A} U_{2B} |0\rangle_2^t \right) \otimes |1\rangle_2^c, \end{aligned} \quad (3)$$

where  $U_{i_A}$  and  $U_{i_B}$  ( $i = 1, 2$ ) are the unitaries performed by the two parties Alice and Bob inside each quantum-switch  $S_i$ .

Next, we measure the two control qubits in the basis  $\{|+\rangle, |-\rangle\}$ . If we observe both of the control qubits in the same state (either  $|+\rangle_1^c |+\rangle_2^c$  or  $|-\rangle_1^c |-\rangle_2^c$ ), the target qubits will be in the (in



**Figure 1 Entangled quantum-switch.** Our work is based on two quantum-switches (S1 and S2). In each quantum-switch, there are two parties, Alice ( $U_{iA}$ ) and Bob ( $U_{iB}$ ). A target qubit is first sent to one party, and then to the other. The order in which the qubit is sent to the two parties is governed by the state of an additional qubit: if the state of the control qubit is  $|0\rangle_i^c$ , the target qubit is sent first to Alice and then Bob (Panel a), and vice versa if the control qubit is in the state  $|1\rangle_i^c$  (Panel b). In our work, we entangle the control qubits (Panel c). In this case, the order in which the target qubit in quantum-switch S1 passes through  $U_{1A}$  and  $U_{1B}$  is entangled with the order in which the target qubit in quantum-switch S2 passes through  $U_{2A}$  and  $U_{2B}$ . The control qubits are measured in the basis  $\{|+\rangle_i^c, |-\rangle_i^c\}$ . If the orders inside the two quantum-switches are entangled, it will be possible to violate a Bell inequality by measuring the target qubits after the quantum-switches (BM). This is possible even if the target qubits start in a separable state and only local operations are applied within each quantum-switch.

general) unnormalised state

$$\frac{1}{\sqrt{2}}(U_{1B}U_{1A}|0\rangle_1^t \otimes U_{2B}U_{2A}|0\rangle_2^t - U_{1A}U_{1B}|0\rangle_1^t \otimes U_{2A}U_{2B}|0\rangle_2^t), \quad (4)$$

while, if we find the control qubits in orthogonal states (either  $|+\rangle_1^c |-\rangle_2^c$  or  $|-\rangle_1^c |+\rangle_2^c$ ), the sign between the two terms in the superposition in the equation above is “+.” In general, depending on the choice of the unitaries in the two quantum-switches, the target qubits will be left either in a separable or in an entangled state. In particular, if we choose the gates

$$U_{1A} = U_{2A} = \sigma_z, \quad (5a)$$

$$U_{1B} = U_{2B} = \frac{\mathbb{1} + i\sigma_x}{\sqrt{2}}, \quad (5b)$$

where  $\sigma_x$  and  $\sigma_z$  are the Pauli operators, the state of the target qubits (upon finding the control

qubits in  $|\pm\rangle_1^c |\pm\rangle_2^c$ ) becomes

$$\frac{1}{\sqrt{2}}(|l\rangle_1^t |l\rangle_2^t - |r\rangle_1^t |r\rangle_2^t), \quad (6)$$

where  $|r\rangle = (|0\rangle - i|1\rangle)/\sqrt{2}$  and  $|l\rangle = (|0\rangle + i|1\rangle)/\sqrt{2}$  (analogously, one gets  $(|l\rangle_1^t |l\rangle_2^t + |r\rangle_1^t |r\rangle_2^t)/\sqrt{2}$  when the control qubit is in  $|\pm\rangle_1^c | \mp \rangle_2^c$ ). This is a *maximally entangled state* and, as a result, one can now violate a Bell inequality on the target qubits.

Within quantum theory, the entanglement between the targets and the resulting violation of the Bell inequality can be explained in terms of the indefiniteness of the temporal orders in the two quantum-switches. In other words, such entanglement is not “generated,” but rather “transferred” from the control qubits by means of the indefinite temporal order of the unitaries applied. A related interpretation of the violation in quantum mechanics is in terms of *time-delocalized quantum operations* [33] and *causal reference frames* [3], according to which a frame can be chosen such that while Bob’s operation acts at a fixed time, Alice’s operation is in a superposition of being implemented before and after Bob’s operation, thus resulting in an indefinite causal order between them.

In the class of GPTs considered here, the presence of non-classical correlations can be determined through a violation of a Bell inequality. In our case, the violation of a Bell inequality with the target subsystems implies the violation of the no-go theorem for temporal order, thereby proving that no underlying GPTs where assumptions I, II and III hold can explain the experimental data. We will experimentally confirm that I holds both in quantum mechanics, and in our class of GPTs (as detailed in Appendix D). Then, we will show, both within quantum mechanics and in our class of GPTs, that one cannot describe our results if only assumption II is invalid. We will thus conclude that either assumption III is wrong or both assumptions II and III are false, hence proving the presence of indefinite causal order beyond the quantum framework.

### 3.1 Experimental scheme

We create a quantum-switch with entangled control qubits using a photonic set-up. Let us first consider a single quantum-switch. Each

quantum-switch applies gates on a target qubit, where the gates' order depends on the state of a control qubit. Experimentally, we encode the control qubit in a path degree of freedom (DOF), and the target qubit in the polarization DOF of a single photon. The photon is initially placed in a superposition of two paths (as explained in Fig. 2 and Appendix E). These paths are labeled  $0_1$  and  $1_1$  for quantum-switch S1 and  $0_2$  and  $1_2$  for quantum-switch S2 in Fig. 2. The two paths are then routed through a two-loop *Mach-Zehnder interferometer* [36, 40]. The  $0_i$  paths lead the photons through a set of gates acting on the polarization DOF in the order  $U_{i_A} \preceq U_{i_B}$ . While the paths  $1_i$  guide the photons through the gates in the opposite order  $U_{i_B} \preceq U_{i_A}$ . To generate the maximally entangled state between the target qubits in Eq. (6), we need to implement the non-commuting gates  $U_{i_A} = \sigma_z$  and  $U_{i_B} = (\mathbb{1} + i\sigma_x)/\sqrt{2}$ , which we do with waveplates. In particular, a half-waveplate (HWP) at  $0^\circ$  for  $\sigma_z$  and a sequence of quarter-waveplate (QWP) and HWP both at  $45^\circ$  for  $(\mathbb{1} + i\sigma_x)/\sqrt{2}$ . After this, the two paths are recombined on a 50/50 beamsplitter (BS) — which projects the path DOF in the basis  $\{|+\rangle\langle+|, |-\rangle\langle-|\}$ . The path lengths and the relative phases are set by means of a piezo-driven trombone-arm delay line. At the two outputs of each interferometer, QWPs, HWPs and polarizing beam splitters (PBSs) are used to perform arbitrary polarization measurements on the target qubits.

To entangle the two quantum-switches, we first entangle the path DOFs of the two photons. As explained in Appendix E, we generate path-entangled photon pairs that are separable in their polarization DOF:

$$|\Phi^-\rangle_{1,2}^{\text{path}} \otimes (|H\rangle_1 |H\rangle_2)^{\text{polar.}} = \left( \frac{|0\rangle_1 |0\rangle_2 - |1\rangle_1 |1\rangle_2}{\sqrt{2}} \right)^{\text{path}} \otimes (|H\rangle_1 |H\rangle_2)^{\text{polar.}}. \quad (7)$$

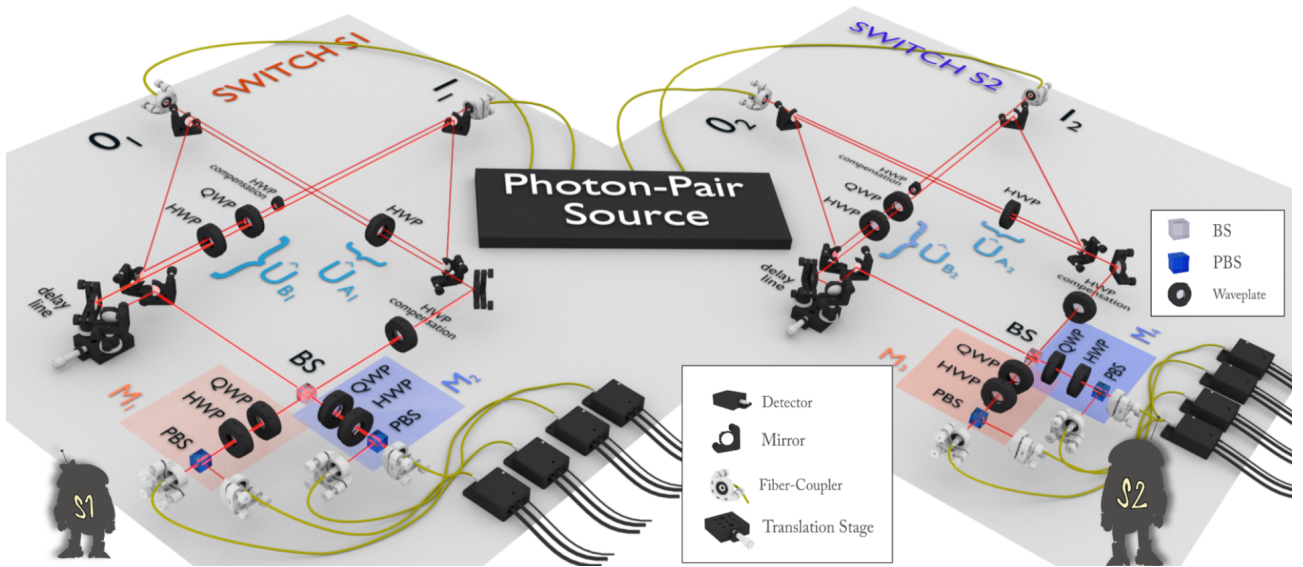
Each photon is thus delocalized over two paths. The two photons are then sent to their respective quantum-switches, and, since the control qubits began in an entangled state, the order in which the gates act on the two target qubits becomes entangled.

Through this scheme, we engineered a situation wherein the only way entanglement can be transferred from one pair of systems to another

is by means of causally non-separable processes. In our experiment, this transfer takes place between different DOFs of photon pairs. Although it is often easy to transfer the entanglement from one DOF to another, this is typically done with a device that directly couples the two DOFs; e.g., in the case of path-polarization transfer, a PBS could be used. In our experiment, we used an entangled quantum-switch to accomplish this interchange. Our quantum-switches do not contain any device which directly couples these DOFs (only waveplates, which act solely on the polarization state, and 50/50 BSs, which act solely on the path state). Rather, here the interchange occurs because the control qubit (the path) governs the order of the application of gates on the target qubit (the polarization). Then, since we begin with an entangled state of the control qubits, this state is transferred to the target qubits via an indefinite order of the application of the gates. In other words, by choosing a specific set of operations, the temporal superposition of the application of these operations is mapped onto a superposition of orthogonal states. As a result, this transfer of entanglement is the signature of an indefinite temporal order.

## 4 Results

Our goal is to demonstrate that the order of application of the gates within the two quantum-switches is genuinely indefinite without assuming that the laboratory operations and the states of the systems are described by quantum theory. We can arrive at this conclusion in three steps. We will first show experimental data that violate a Bell inequality. From this we can assert that *at least one* of the three assumptions must be false. We will then prove that assumption I is satisfied in our experiment using a class of GPTs. Thus, one of the remaining assumptions (i.e., assumptions II and/or III) must not hold. We will analyse the case in which only assumption II does not hold within the set of GPTs. By acquiring additional data on a single quantum-switch, we will show that such scenario cannot reproduce the results of our experiment. Consequently, the only two possible explanations are that either assumption III does not hold, or that both assumptions II and III are false. In either case, assumption



**Figure 2** Experimental implementation of an entangled quantum-switch. Each quantum-switch is composed of a two-loop Mach-Zehnder interferometer. The interferometers start in the photon-pair source, wherein photon 1 and photon 2 are placed in superposition of the paths  $0_1$  and  $1_1$ , and  $0_2$  and  $1_2$ , respectively (see Appendix E). (For simplicity, we have drawn these paths as fibers, however the photons are transmitted via free-space from the source to the experiment.) These paths are routed such that path  $0_i$  sees gate  $U_{i_A}$  and then gate  $U_{i_B}$ , and vice versa for the path  $1_i$ . Each gate, acting on the polarization degree of freedom, is made up of waveplates (as described in the main text). The paths  $0_i$  and  $1_i$  are then combined on a beam splitter (BS). In quantum-switch S1 (S2), the photon is detected after the polarization measurement at  $M_1$  or  $M_2$  ( $M_3$  or  $M_4$ ). Both quantum-switches contain two compensation half-waveplate at the beginning and at the end of the reflected arm, so as to compensate for the phase shifts due to the reflection from the polarizing beam splitters composing the source (shown in Fig. 3). Together with the BS (which applies a Hadamard gate to the qubit encoded in the path DOF), detecting the photon at  $M_1$  or  $M_2$  ( $M_3$  or  $M_4$ ) projects the path qubit on  $|+\rangle$  or  $|-\rangle$ , respectively. Furthermore, within each measurement  $M_i$ , the polarization qubit can be measured in any basis by a combination of a quarter-waveplate (QWP), half-waveplate (HWP), and polarizing beam splitter (PBS).

III must be false, and therefore the local operations within the two quantum-switches have been applied in an indefinite temporal order.

#### 4.1 Violation of the no-go theorem

We begin by performing a Bell test between the target states at the output of the apparatus. This allows us to experimentally probe a conjunction of all three assumptions. We realize the Bell test (more specifically, we measure a Clauser-Horne-Shimony-Holt (CHSH) inequality [15]) on the polarization DOF, using four equivalent measurement set-ups (orange and blue boxes in Fig. 2). Since the 50/50 BSs apply a Hadamard gate on the path qubits, we post-select the control qubits in the same state (either  $|+\rangle_1^c |+\rangle_2^c$  or  $|-\rangle_1^c |-\rangle_2^c$ ) by grouping the results of  $M_1$  with  $M_3$  (orange boxes) and  $M_2$  with  $M_4$  (blue boxes). We obtain  $S_{\text{target}} = 2.55 \pm 0.08$ . This violates the inequality, and thus also the no-go theorem, by almost 7 standard deviations. Therefore, in our class of GPTs, no theory satisfying assumptions I, II and

III is compatible with the experimental data.

#### 4.2 Verification of assumption I

We now proceed to test the validity of assumption I, which says that the joint target state (shared between system S1 and S2) is Bell-local. We will show experimentally that, within the limits of our experimental precision, the target systems are compatible with a product state in the class of GPTs under consideration, and they are therefore Bell-local.

To test assumption I within a class of GPT, we assume that the set of “fiducial measurements” of the class of GPTs contains “quantum fiducial measurements” as a subset. Moreover, we assume that pure states in quantum theory are also pure states in GPTs. In this sense, quantum theory is embeddable in the GPT. This is similar to how classical theory can be embedded in quantum theory (i.e., classical theory has one fiducial measurement in the “computational basis”). In particular, we consider a class of GPTs wherein the state



space of a single two-level system is described by a  $d$ -dimensional Bloch ball [16, 30] (i.e., there are three quantum fiducial measurements, and  $d - 3$  non-quantum fiducial measurements), with  $d > 3$  in general. The methodological advantage of considering such theories is that, in certain cases, the structure of the state can be inferred despite the fact that only quantum measurements can be made (i.e., only measurements in the three-dimensional quantum subspace). For this class of theories, it was shown that a single system is in a pure state if there exists a measurement for which the system returns a given result with probability one. Similarly, a bipartite system is in a pure product state, if the above statement applies to each individual system. In more detail, in GPTs a state is pure if it cannot be written as a non-trivial mixture of other states. Moreover, a bipartite system is in a product state if, for all local measurements, the probabilities for outcome pairs on a bipartite-state are equal to the product of the two marginal probabilities of each subsystem. Such a state has perfect correlations only for a pair of fiducial measurements, it exhibits no further correlations in any other pair of fiducial measurements, and it cannot violate a Bell inequality [16, 23, 24, 31].

For our target photon pair we demonstrated that both photons return value  $H$  with certainty. This means that, already from a pair of quantum fiducial measurements, one can conclude that, up to experimental imperfections, the state is a pure product state, and therefore it cannot violate a Bell inequality. This supports the validity of assumption I in the special case of Bloch-vector theories (see Appendix A). In Table S1, we compare the probabilities for outcome pairs on a bipartite-state to the product of the two marginal probabilities of each subsystem. The excellent agreement between the two probability distributions indicates that the joint target state is indeed a pure product state, and cannot violate a Bell inequality. This proves that assumption I of our no-go theorem holds for the class of GPTs under consideration. We quantify to what extent the two distributions agree by calculating the *root-mean-square* (RMS) difference between the two distributions, resulting in an average difference of  $0.6 \cdot 10^{-2} \pm 2.7 \cdot 10^{-2}$ . Although this value is consistent with zero, one could imagine that this small difference is in fact caused by correlations

between the two target systems. In Appendix D we show, however, that such small correlations can only give rise to a vanishingly small violation of Bell’s inequality. The possible level of violation from this amount of potential coupling is insufficient to explain our experimentally observed violation. Therefore, we have confirmed that the joint target system starts in an (approximately) separable state. Additionally, in Appendix E, we experimentally show that the joint control system is initially entangled. We then send this joint state into our two quantum-switches and perform measurements on the output state. Furthermore, Appendix G reports the results of the same studies presented above when a quantum-mechanical description of the experimental set-up is assumed.

Having proven, both in quantum theory and within the class of GPTs, that our no-go theorem is violated and that assumption I is justified, we can conclude that either assumption II, or assumption III, or both must be false. We will now consider the case in which only assumption II is false.

### 4.3 Verification of assumption II

The second assumption of our no-go theorem says that the laboratory operations performed in the two quantum-switches cannot transform the joint state of the target systems of S1 and S2 from a local state to a non-local one. We will again experimentally prove a stronger condition, namely that laboratory operations are local maps on the targets in the considered GPTs. As discussed earlier, this condition can be violated for two reasons. First, the laboratory operations (IIa) may induce a direct interaction between two targets. To avoid this, one could perform the operations with a space-like separation in which case the condition would be guaranteed in any theory obeying relativistic locality. However, in our experiment, we make the (well-justified) device-dependent assumption that the laboratory operations are local transformations within S1 and S2 in GPTs, since the transformations of the systems take place at spatially separated parts of the optical table. As a consequence, the first form of violation (IIa) of assumption II arguably does not take place in our experiment.

Let us consider now the second, more substantial, way in which assumption II could be violated. Since we are analysing the case in which

only assumption II is false while assumption III holds, we will now study the scenario in which the laboratory operations occur only in a causally-ordered manner.

The second form of violation of assumption II is that the laboratory operations of one party  $S_i$  could “couple” the control and the target systems (IIb). Such a coupling would make it possible to transfer non-local correlations from the control systems of the two parties  $S_1$  and  $S_2$  to their target systems, and therefore a violation of Bell’s inequalities would be possible.

We can experimentally prove that, within the class of GPTs considered, the laboratory operations transform product states of the control and the target subsystems into product states within each quantum-switch. We will do so using a similar technique as for assumption I. We start by placing bounds on the degree of coupling between the target and the control qubits within a single quantum-switch in the presence of *only one* of the two operations (i.e., only  $U_A$  or only  $U_B$ ) inside the quantum-switch. We perform the full set of quantum fiducial measurements. With this, we show that the joint probabilities of the target and the control subsystems are factorizable into the products of the two marginal probabilities of each subsystem in the case where either only  $U_A$  or  $U_B$  is inserted inside a single quantum-switch. From this we can conclude that the joint probabilities must also remain factorizable even when both  $U_A$  and  $U_B$  are inserted if the order is well-defined (or a classical mixture of the two well-defined orders) [1]. In other words, in our GPTs, the optical elements do not couple the control and the target subsystem in the “quantum subspace” of the GPT state space. Moreover, the marginal probabilities measured on the control and the target subsystem correspond to a pure (product) state. From this we also conclude that there can be no coupling in the “non-quantum subspace” either. We analyse this by performing a set of measurements on the joint control-target system, and by showing that the joint probabilities can be described by the product of the marginal probabilities (see Tables S2-S3, and Appendix F for more details). The RMS difference between the two distributions is, on average,  $0.02 \pm 0.03$ . This value is within one standard deviation of zero, confirming that the probability distribution is consistent with that of a product state. As we discussed for assumption

1, this small discrepancy could be caused by correlations between the control and the target systems. However, as we show in Appendix F, these correlations are too weak to explain our experimentally observed violation of Bell’s inequality.

#### 4.4 Implications for the temporal order

In the previous subsection we proved that the laboratory operations do not couple the control and target systems when they are applied in a well-defined order. This means that, for our experiment, whenever assumption III holds, assumption II must also hold. Because the statement  $a \Rightarrow b$  is logically equivalent to *not*  $b \Rightarrow$  *not*  $a$ , this further implies that if assumption II is invalid, then assumption III must also be invalid. It follows that the only two possible scenarios are that (1) assumption II is wrong, and therefore III is also wrong, or (2) assumption III is false, independently of assumption II. In either case, it is not possible to explain our experimental data unless assumption III is discarded. We thus conclude that the local operations in our experiment were applied in an indefinite order. (Notice that it is not a logical necessity that every time III holds, II must hold in turn. In general, one could find a system with a defined causal order (III) in which the operations are non-local (not-II). This would be the case, for instance, of circuits where control and target systems undergo entangling gates in a definite causal order (e.g., CNOT). Nevertheless, this was shown not to be the case in our experiment, as operations  $U_A$  and  $U_B$  were proven not to couple control and target systems when applied individually.)

## 5 Discussion

In this work, we entangled the temporal orders between operations applied by two parties and experimentally showed that the resulting temporal order is indefinite, by violating a Bell inequality using the joint target system after the quantum-switches. We thus verified that the data collected by entangling temporal orders in the quantum-switches cannot be described by a class of (generalized probabilistic) theories under the assumption that the initial joint target state is Bell-local, the operations on the target states are local, and they have a pre-defined order. This

did not require the assumption that the systems and the operations are described by the quantum formalism. Clearly, for our demonstration to be loophole-free (as proposed in Ref. [43]), the standard Bell loopholes (fair-sampling and locality) would need to be closed. Further loopholes can arise related to the implementation of the quantum-switch. In fact, it is known that experimental data produced by the quantum-switch can be simulated by a causally-separable process if at least one of the operations (either  $A$  or  $B$ ) is performed two or more times. In relation to the experimental implementations of the quantum-switch, this is the so-called “single-usage loophole.” Closing this loophole would require an operational verification that each operation in the quantum-switch is performed only once. For example, this could include implementation of a “counter” that would estimate the number of times an operation is performed, or a process tomography on time-delocalized quantum systems [33]. However, our experiment is immune to the single-usage loophole. Indeed, a multiple usage of operations in a defined causal order on either side of the Bell test cannot be simultaneously shown to be local and violate the Bell inequality for temporal order.

All previous studies involving quantum processes with indefinite temporal order achieved their goal by superposing the order of operations, rather than entangling them. The first proposal to entangle the temporal order was made only recently [43]. Here we show that the basis of this theoretical concept is in fact experimentally accessible. Moreover, we exploit this resource as a new means to validate indefinite causal structures. Techniques to characterize these structures are becoming increasingly relevant, as it is known that these processes can lead to logarithmic [38] advantages in query complexity, and exponential advantages in quantum communication tasks [4, 14, 18, 22].

**Acknowledgments:** We thank C. Branciard, F. Costa, B. Dakić, M. Jacquet, A. Moqanaki and T. Strömberg for useful discussions. **Funding:** acknowledges financial support from the Royal Society through the Newton International Fellowship No. NIF\R1\202512. L.A.R. acknowledges support from the Templeton World Charity Foundation (fellowship no. TWCF0194). M.A. acknowledges support from the Excel-

lence Initiative of the German Federal and State Governments (Grant ZUK 81). M.Z. acknowledges support through an ARC DECRA grant DE180101443, and ARC Centre EQuS (CE170100009). Č.B. acknowledges support from the John Templeton Foundation, Foundational Questions Institute (FQXi), Austrian Science Fund (FWF) through BeyondC (F7103-N38), the projects no. I-2562-N27 and I-2906, as well as support from the European Commission via Testing the Large-Scale Limit of Quantum Mechanics (TEQ) (No. 766900) project. Č.B. and P.W. acknowledge support from the Austrian Science Fund (FWF) via Doctoral Programme CoQuS (no. W1210-4), and the research platform TURIS. P.W. also acknowledges support from the Austrian Science Fund (FWF) through BeyondC (F7113-N38), the GIPSS (P30817-N36) and NaMuG (P30067-N36), United States Air Force Office of Scientific Research via QAT4SECOMP (FA2386-17-1-4011) and Red Bull GmbH. **Author contributions:** G.R., L.A.R., Č.B. and P.W. designed the experiment. G.R. built the set-up and carried out data collection. G.R. and L.A.R. performed data analysis. F.M. built the single-photon source. M.Z., M.A. and Č.B. developed the theoretical idea. L.A.R., P.W. and Č.B. supervised the project. All authors contributed to writing the paper. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Information. Additional data related to this paper may be requested from the authors.

## A Proof of no-go theorem for temporal order

All previous experimental studies of causally non-separable processes [36, 40] were dependent on the validity of the quantum theory (i.e., they were *theory-dependent*), and all known physically realizable processes satisfy all causal inequalities (see the Suppl. Information) [5, 34]. The latter means that experimental data taken from a given causally non-separable quantum process could still be understood as arising in causal manner, for example from a process with a definite causal order in an underlying generalized probabilistic theory (GPT). Therefore, it is unknown whether a fully theory-independent experimental

proof of indefinite causal order is possible.

In our current work, we relate a violation of a Bell inequality to the violation of a no-go theorem for temporal order, as proposed in Ref. [43]. This results in a proof of causal indefiniteness outside of the quantum framework as it is valid for a large class of generalized probabilistic theories. In this section, we provide a rigorous introduction to such no-go theorem for temporal order.

We will begin by giving a brief introduction to the basic elements of GPTs which are necessary for our no-go theorem. A more detailed discussion of the GPT framework can be found in Ref. [7, 23, 31].

In a GPT, a system is described by a state  $\omega$  that specifies outcome probabilities for all measurements that can be performed on it. A complete representation of the state is given by specifying the outcome probabilities of a so-called “fiducial set.” The smallest such set defines the number  $d$  of degrees of freedom of the system. We restrict our consideration here to binary systems that have two perfectly distinguishable states and no more. For example, the fiducial set for a two-level system in quantum theory consists of the (three) probability outcomes of spin projections along  $x$ ,  $y$  and  $z$ . The state space is a compact and convex set  $\Omega$  embedded in a vector space. The extremal states of  $\Omega$  that cannot be decomposed as a convex mixture of other states are called “pure states.” An effect  $e$  is defined as a linear functional on  $\Omega$  that maps each state onto a probability, i.e.,  $e : \Omega \rightarrow [0, 1]$ , where  $e(\omega)$  is the probability to obtain an outcome on the state  $\omega$ . The linearity is required to preserve the convex structure of the state space.

A transformation  $U$  is a linear map from a state to a state, i.e.,  $U : \Omega \rightarrow \Omega$ . The transformation is linear for the same reason that probabilities have to be linear maps of states. The sequence of transformations  $U_1, \dots, U_n$ , in which transformation  $U_1$  “precedes” transformation  $U_2$ , which “precedes”  $U_3$ , etc., is represented by a composition of maps:  $U_n \circ \dots \circ U_1$ . This defines a *definite order of transformations*, which we denote as  $U_1 \preceq \dots \preceq U_n$ .

We will now introduce a generalization of the no-go theorem for temporal order, which was originally proposed in Ref. [43].

In the framework of a GPT, the state of a composite system shared between two parties S1 and

S2 is given by  $\omega_{1,2} \in \Omega_{1,2}$ , where  $\Omega_{1,2}$  is the state space of a composite system. The state of a composite system is given by a multiplet consisting of the *local states*  $\omega_1 \in \Omega_1$  and  $\omega_2 \in \Omega_2$  of individual systems, the *correlation tensor*  $\hat{T}$  and a potential *global parameter*  $\xi$  [16, 23, 24, 31]:

$$\omega_{1,2} = \omega_{1,2}(\omega_1, \omega_2, \hat{T}, \xi). \quad (8)$$

The fact that subsystems are themselves systems implies that each has a well-defined reduced state  $\omega_1, \omega_2$  which does not depend on which transformations and measurements are performed on the other subsystem; this is often referred to as “no-signaling.” We also assume that transformations and measurements performed on subsystems commute with each other, so that *one* correlation tensor is enough to describe correlations between them. If this were not the case, we would need to introduce *two* correlation tensors, one when S1 applies operations before S2, and the other when S2 performs operations before S1. Finally, the states in GPT need not to satisfy the *local tomography* condition (stating that reduced states and correlation tensor completely describe the systems’ state), but may include a global parameter  $\xi$ .

For the present case of binary systems, the components of the state in Eq. (8) are given by

$$\omega_1^{(i)} = p^{(i)}(o_1 = 1) - p^{(i)}(o_1 = -1), \quad (9a)$$

$$\omega_2^{(j)} = p^{(j)}(o_2 = 1) - p^{(j)}(o_2 = -1), \quad (9b)$$

$$T^{(i,j)} = p^{(i,j)}(o_1 o_2 = 1) - p^{(i,j)}(o_1 o_2 = -1), \quad (9c)$$

where  $i, j = 1, \dots, d$ . Here, for example,  $p^{(i)}(o_1 = 1)$  is the probability to obtain the outcome  $o_1 = 1$  when the  $i$ -th measurement is performed on the first subsystem, and  $p^{(i,j)}(o_1 o_2 = 1)$  is the joint probability to obtain correlated results (i.e., either  $o_1 = o_2 = +1$  or  $o_1 = o_2 = -1$ ) when the  $i$ -th measurement is performed on the first subsystem and the  $j$ -th measurement on the second one.

An effect  $e_{12}$  that maps a state onto a probability for a pair of *local* measurements is given by  $e_{12} = e_{12}(r_1, r_2, r_1 r_2^T)$ , where  $r_i$  is the effect on the state of  $i$ -th system, and  $r^T$  denotes transposition of  $r$ . (Note that the global parameter does not contribute to the probability for a pair of local measurements). The probability to obtain the effect  $e_{12}$  when the system is prepared in

the state  $\omega_{12}$  is given by

$$p(e_{12}|\omega_{12}) = \frac{1}{4} \left[ 1 + (\omega_1 \cdot r_1) + (\omega_2 \cdot r_2) + (r_2 \cdot \hat{T}r_1) \right], \quad (10)$$

where  $(x \cdot y)$  is the Euclidean scalar product between two  $d$ -dimensional real vectors  $x$  and  $y$ .

The product state is represented by  $\omega_p = \omega_p(\eta_1, \eta_2, \eta_1\eta_2^T, \xi_p)$ , where the correlation tensor is of a product form. If we perform a pair of local measurements on the arbitrary product state, the outcome probability factorizes into the product of the local outcome probabilities.

We next introduce a pair of *local (reversible) transformations*  $(U_1, U_2) : \Omega_{12} \rightarrow \Omega_{12}$  as a linear map from the space of states of a composite system to itself:

$$(U_1, U_2)(\omega_{12}) = (U_1\omega_1, U_2\omega_2, U_1\hat{T}U_2^T, \xi'), \quad (11)$$

where the global parameter  $\xi'$  is, in general, changed under the transformations  $(U_1, U_2)$ . Since testing our Bell inequality involves only local transformations and measurements, it is sufficient to specify effects for those measurements.

In our experiment,  $\omega_1$  and  $\omega_2$  themselves are states of composite systems each consisting of a “control” and a “target” subsystem. Thus, the entire system under investigation consists of four subsystems, a control and a target subsystems of S1 and a control and a target subsystems of S2. The overall state is

$$\omega_{1,2,3,4} = \omega_{1,2,3,4}(\omega_1^t, \omega_1^c, \omega_2^t, \omega_2^c, \dots, \hat{T}^{ij}, \dots, \hat{T}^{ijk}, \dots, \hat{T}^{1234}, \Xi), \quad (12)$$

where  $c$  and  $t$  refer to the terms “control” and “target” subsystems,  $\hat{T}^{ij}$ ,  $\hat{T}^{ijk}$  and  $\hat{T}^{1234}$  are correlation (sub)tensors describing correlations between pairs  $\{i, j\}$ , triple  $\{i, j, k\}$  and quadruple  $\{1, 2, 3, 4\}$  of subsystems, respectively, and  $\Xi$  is the set of all global parameters.

The no-go theorem concerns the reduced state of the two target systems as given by

$$\omega_{1,2}^t = \omega_{1,2}^t(\omega_1^t, \omega_2^t, \hat{T}^{tt}, \xi^t), \quad (13)$$

where  $\omega_1^t$  and  $\omega_2^t$  are states of the target subsystems of S1 and S2,  $\hat{T}^{tt}$  is their correlation tensor, and  $\xi^t$  is the corresponding global parameter.

Leveraging these definitions, we now present three assumptions, which are the fulcrum of our no-go theorem for a definite local causal order.

### 1. The initial joint state of the target system $\omega_{1,2}^t$ is Bell-local (i.e., it satisfies Bell’s local-causality condition for all pairs of measurements)

Suppose that the two observers can each perform a measurement  $\mathcal{O}_1$  and  $\mathcal{O}_2$ , respectively. We label  $m_1$  and  $m_2$  as arbitrary measurement choices of S1 and S2, and  $o_1$  and  $o_2$  as the corresponding outcomes. Under these conditions, we suppose that probabilities obtained from any such measurements can be described through a local hidden variables theory (i.e., in Bell’s terms, a theory that satisfy “local causality”), and therefore it is associated to the probability distribution

$$p(o_1, o_2 | m_1, m_2, \omega_{1,2}^t) = \int \rho(\lambda) p(o_1 | m_1, \lambda, \omega_{1,2}^t) p(o_2 | m_2, \lambda, \omega_{1,2}^t) d\lambda, \quad (14)$$

where  $\lambda$  is often referred to as a “hidden variable.” We implicitly assume the “freedom of choice” condition — the assumption that the choices of the measurement settings are independent of  $\lambda$  — is fulfilled.

### 2. The laboratory operations are represented by local transformations $U_i^t$ on the target subsystems. If they are applied to a Bell-local state of the targets, the state remains Bell-local.

This is satisfied by definition, since in the GPTs the action of a transformation followed by a (fixed) measurement can be understood as another measurement, and the initial state is assumed to satisfy Bell’s condition of local causality for *all* pairs of local measurements. Therefore, the action of a pair of local transformations cannot produce a non-local Bell state from a local state.

### 3. The order of S1’s and S2’s operations on the target system is well defined.

Suppose first that the order of application of the local operations performed inside quantum-switch S1 ( $U_{1A}^t \preceq U_{1B}^t \preceq \dots$ ) and those performed inside quantum-switch S2 ( $U_{2A}^t \preceq U_{2B}^t \preceq \dots$ ) are fixed. Since an ordered sequence of local transformations is still a local transformation, if a Bell-local state undergoes such a transformation on S1’s and S2’s sides, it remains Bell-local. The state remains Bell-local even if the order of operations is chosen with a given probability distribution due to convexity. The mutual order between S1’s and S2’s operations is irrelevant, since we have assumed the two classes of operations to commute.

**Theorem.** *No states, set of transformations and measurements which obey the assumptions I-III can result in violation of a Bell inequality.*

*Proof.* Following I, suppose that the initial target state  $\omega_{1,2}^t$  is Bell-local. This means that Eq. (14) is fulfilled for an arbitrary pair of local measurements. Because of III, operations in S1's and in S2's laboratories are applied in a definite order, say  $U_{1A}^t \preceq U_{1B}^t \preceq \dots$  in S1's side, and  $U_{2A}^t \preceq U_{2B}^t \preceq \dots$  in S2's side. The state evolves, therefore, under a composition of the local operations as

$$\dots (U_{1B}^t, U_{2B}^t) \circ (U_{1A}^t, U_{2A}^t)(\omega_{1,2}^t).$$

Let us restrict ourselves to the case of only two transformations per quantum-switch ( $U_A^t$  and  $U_B^t$ ). After the pairs of operations are applied in order  $U_{1A}^t \preceq U_{1B}^t$  and  $U_{2A}^t \preceq U_{2B}^t$  on the two sides, the state becomes

$$\begin{aligned} \omega_{1,2}^{t'} &= (U_{1B}^t, U_{2B}^t) \circ (U_{1A}^t, U_{2A}^t)(\omega_{1,2}^t) \\ &= (U_{1B}^t \circ U_{1A}^t, U_{2B}^t \circ U_{2A}^t)(\omega_{1,2}^t) \end{aligned} \quad (15)$$

which is still local due to I - III. Hence

$$\begin{aligned} p(o_1, o_2 | m_1, m_2, \omega_{1,2}^{t'}) &= \\ \int \rho(\lambda) p(o_1 | m_1, \lambda, \omega_{1,2}^{t'}) \cdot p(o_2 | m_2, \lambda, \omega_{1,2}^{t'}) d\lambda. \end{aligned} \quad (16)$$

In general, the order of operations does not need to be fixed, but can be specified probabilistically by a further hidden variable  $\nu$ , whose different values correspond to different permutations of the order of operations. We obtain

$$\begin{aligned} p(o_1, o_2 | m_1, m_2) &= \iint \rho(\lambda, \nu) p(o_1 | m_1, \lambda, \omega_{1,2}^{t,\nu}) \\ &\cdot p(o_2 | m_2, \lambda, \omega_{1,2}^{t,\nu}) d\lambda d\nu, \end{aligned} \quad (17)$$

where  $\rho(\lambda, \nu)$  is the joint probability distribution over the two types of variables, and  $\omega_{1,2}^{t,\nu}$  is the final state of the target systems upon application of the transformations in the order given by  $\nu$ .

Thus, we conclude that a local target state subjected to the action of a set of local operations applied in a pre-defined order can by no means lead to the violation of Bell inequalities, even if the order is chosen probabilistically in each run of the experiment. This concludes the proof.  $\square$

## B Device-independency and theory-independency

Causal witnesses, violation of Bell inequalities for temporal order, and violation of causal inequalities build a hierarchy of the notion of indefinite causality. The weakest notion of indefinite causality is that of causal non-separability, which is formulated using quantum theory. A violation of a causal inequality is the strongest notion as it is formulated solely in terms of observable probabilities  $p(a, b | x, y)$  without any assumption about the internal function of experimental devices — it is therefore device-independent. The violation of a Bell inequality for temporal order should be considered, in our view, a stronger proof of indefinite causality than the measurement of a causal witness, but a weaker proof than a violation of a causal inequality. The reason why it is weaker than a causal inequality violation is that, although it too is formulated in terms of the probabilities  $p(a, b | x, y, \omega)$ , it also involves the notion of state  $\omega$  and the assumption how laboratory operations act on it (see Appendix A) — this causes the proof to be device-dependent. However, it can be defined for a class of generalized probabilistic theories, and therefore it does not rely on the quantum formalism. It is thus more general than the notion of a causal witness. Although the quantum-switch violates a weaker notion of causality, shaped for quantum theory, it cannot violate the stronger (device-independent) notion of causal inequalities. The open question addressed in our work is then whether we can still use the quantum-switch to perform a proof of indefinite causal order independent of quantum formalism. The present experimental study provides an affirmative answer to this question.

## C Relation between the present work and Ref. [43]

In Ref. [43], the position of a massive object serves as a “control” quantum system and a quantum system (e.g., a photon) that is exchanged between Alice's and Bob's laboratory as a “target” system. By putting the massive object in a macroscopic superposition of two positions, one closer to Alice's and the other closer to Bob's position, one induces a relative time dilation between Alice's and Bob's laboratory. The super-

position of massive objects can effectively lead to “entanglement” of the temporal order between local operations, enabling the violation of a Bell-type inequality. In the conceptual framework of general relativity, the resource for the violation is a “non-classical space-time” created by macroscopic superposition of large masses. In the second-quantized picture, the superposition can be seen as entanglement in the Fock basis, and the scheme enables one to “swap” this entanglement to the final entanglement of the target systems. Unfortunately, the physical demands of the proposal make that experiment infeasible. However, quantum control of the order of events can also be achieved without the use of gravitational interaction. This can be done, for example, in an extended quantum circuit model, wherein the order of applied quantum gates is coherently controlled by an ancillary system (the quantum-switch). The difference between the two schemes is that in the gravitational scheme, the spatio-temporal distance of *any* pair of events in a space-time region is influenced by a superposition state of the mass, whereas in the linear optical implementation, only the order of the gates applied on the propagating system (e.g., photons) is indefinite.

A more detailed analysis of the differences and similarities between the gravitational quantum-switch and its photonic counterpart here presented is given in the Suppl. Information - Sec. III.

## D Experimental proof of assumption I in GPTs

Recall that assumption I says that the initial target states do not violate a Bell inequality. In the notation introduced above, the initial target state is  $\omega_{1,2}^t$ . Our demonstration of assumption I presented here is based solely on experimental data, and can be shown to be valid for a class of GPTs. Our goal is to prove that the input state is a product state, and thus it is local.

Let us denote the probabilities for measurement outcomes as measured on reduced states of the target system of S1 and S2 as  $p(o_1|m_1, \omega_1^t)$  and  $p(o_2|m_2, \omega_2^t)$ , respectively. If the state is a local product state then the probability for joint outcomes, as measured on the composite system of the two target subsystems in the initial state

$\omega_{1,2}^t$ , is factorisable, i.e., it can be expressed as

$$p(o_1, o_2|m_1, m_2, \omega_{1,2}^t) = p(o_1|m_1, \omega_1^t) \cdot p(o_2|m_2, \omega_2^t). \quad (18)$$

We experimentally performed a large set of measurements on the input target states, and checked for this property. The measurements we made are tomographically complete in quantum theory. Nevertheless, in a GPT this might not be the case, as a GPT system may have more degrees of freedom than a quantum system. We thus restrict our considerations to a class of GPTs for which we assume that polarization measurements in three unbiased bases for each photon constitute a subset of the full set of “fiducial measurements”. For example, in the case of GPTs whose systems are described by Bloch vectors of dimension  $d$ , three components of the vectors correspond to “quantum fiducial measurements” of a single system. Similarly, in the GPTs, the correlation tensor is given by  $d^2$  elements of which 9 elements (i.e., 3 fiducial measurements for the first times 3 fiducial measurements for the second system) correspond to the “quantum subspace” of the correlations that are accessible through quantum measurements.

Our measurements confirm that the joint probabilities for “quantum fiducial measurements” are factorized for the two targets. In general, however, it might be possible that within the subset of quantum fiducial measurements for a bipartite system the joint probabilities are factorized into a product of marginal probabilities although the overall state is not a product one. This is because non-zero correlations could exist between non-quantum fiducial measurements. It would then be possible to transfer correlations from the “non-quantum subspace” into the correlations within the “quantum subspace” by applying some “exotic” (i.e., non-quantum) local transformations. Nevertheless, for our class of GPTs, where subsystems are represented by Bloch vectors of general dimension  $d$  and pure quantum states are pure states of the GPTs, we know that if both subsystems individually return probability one for some measurement outcome, the state is a pure product one and it cannot violate assumption I (*Lemma 1* of Ref. [16]). More precisely, the state would have the form  $\omega_{1,2}^t = \omega_{1,2}^t(\omega_1^t, \omega_2^t, \omega_1^t(\omega_2^t)^T, \xi^t)$ , with  $\omega_1^t$  and  $\omega_2^t$  being in pure states, i.e.,  $\|\omega_1^t\| = \|\omega_2^t\| = 1$ . In our experiment, we obtain outcomes with probability

one for a pair of quantum fiducial measurements on the two target subsystems, and therefore the two subsystems cannot exhibit any further correlations within the non-quantum subspace.

Table S1 shows the values of the probabilities  $p(o_1, o_2 | m_1, m_2, \omega_{1,2}^t)$  (which, for brevity, is indicated as  $p_{1,2}$  in the Tables) in the first four columns, and the marginal probability products  $p(o_1 | m_1, \omega_1^t) \cdot p(o_2 | m_2, \omega_2^t)$  (denoted as  $p_1 \cdot p_2$  in the Table) in the last four columns, with almost perfect correlations in the  $\{H, V\}$  basis. Moreover, the joint and the two marginal probabilities are all almost one for the HH outcome, confirming the high purity of the bipartite state. It can be seen that the two sets of probabilities agree well. More quantitatively, let us define the *root-mean-square* (RMS) distance between the two sets of probabilities as

$$d = \sqrt{\frac{1}{N} \sum_{o_1, o_2} \sum_{m_1, m_2} \Delta p_{o_1, o_2, m_1, m_2}^2}, \quad (19)$$

with

$$\begin{aligned} \Delta p_{o_1, o_2, m_1, m_2} &= p(o_1, o_2 | m_1, m_2, \omega_{1,2}^T) \\ &- p(o_1 | m_1, \omega_1^T) \cdot p(o_2 | m_2, \omega_2^T), \end{aligned} \quad (20)$$

and where  $N$  is the number of data points. Evaluating this over our results, we obtain a RMS distance of  $(0.6 \pm 2.7) \cdot 10^{-2}$ , indicating that the two distributions are equal within error.

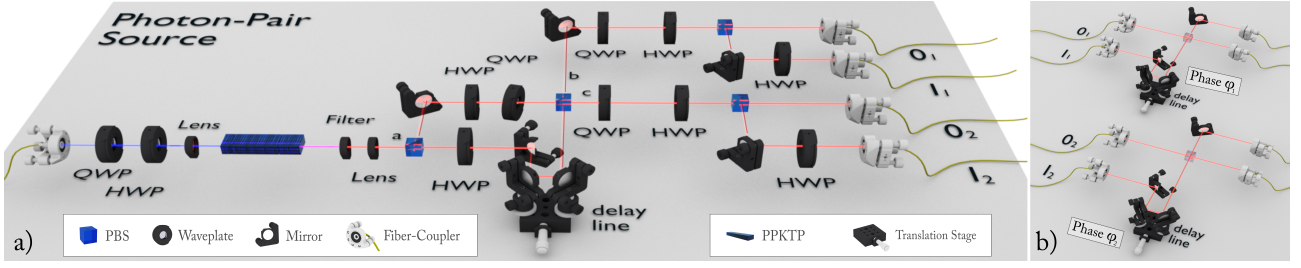
Although the two target systems are approximately in a product state, the small discrepancy between the two distributions allows for some correlations between the systems. Following the Peres-Horodecki criterion [17, 42], the maximal value of the CHSH inequality in quantum mechanics is given in terms of the two largest absolute values of the correlation tensor singular values, say  $t_1$  and  $t_2$ , as  $2\sqrt{t_1^2 + t_2^2}$ . From the full set of the fiducial measurements in the quantum subspace, we can estimate the maximal possible amount of violation of the CHSH inequality for the two target systems to be  $2\sqrt{t_1^2 + t_2^2} = 2.12 \pm 0.04$ . This is more than ten standard deviations lower than the observed violation of the inequality. Thus, this digression from assumption I cannot explain the violation of the inequality.

## E Entangled photon source

A *periodically-poled potassium titanyl phosphate* (PPKTP) crystal, phase-matched for collinear type-II *spontaneous parametric down-conversion* (SPDC), converts one photon at 426 nm into two photons at 852 nm. The photonic state after the crystal can be approximated to a Fock state of two photons in two orthogonal polarization modes  $|H, a\rangle|V, a\rangle$ , where  $a$  indicates the common spatial mode of the two photons defined in Fig. 3. Two PBSs are used to separate and then recombine the two photons. Each photon passes through a HWP set at  $\pm 45^\circ$ . The state after the second PBS is therefore:  $(|H, b\rangle|H, c\rangle - |H, b\rangle|V, b\rangle + |H, c\rangle|V, c\rangle - |V, b\rangle|V, c\rangle) / 2$ , where  $b$  and  $c$  indicate the two output spatial modes of the second PBS. By post-selecting on coincidences, only the part of the state with the photons in two different spatial modes is kept, resulting in the polarization-entangled state  $(|H, b\rangle|H, c\rangle - |V, b\rangle|V, c\rangle) / \sqrt{2}$ . We then use two PBSs and two HWPs (Fig. 3) to convert this state into a path-entangled state:  $(|0\rangle_1|0\rangle_2 - |1\rangle_1|1\rangle_2) / \sqrt{2}$ , where the notation is the same as specified in Fig. 3. A trombone delay line in between the two PBSs is used to compensate temporal delay between the two photons, and a multi-order QWP in one mode is tilted to compensate for undesired phases between the two components of the final quantum state. The delay line and the QWP can be also used to modify the final output state in a controllable way. In particular, by unbalancing the two paths by the coherence length of the down-converted photons, the entangled state can be converted into a statistical mixture of the states  $|0\rangle_1|0\rangle_2$  and  $|1\rangle_1|1\rangle_2$ .

For our experiment, both the path and the polarization states of the photon pairs are important. To characterize the polarization state, we can perform two-qubit polarization state tomography using a QWP, a HWP and a PBS for each photon (Fig. 3, Panel a). To characterize the path entanglement, we perform a Bell measurement on the path qubits using the apparatus shown in Fig. 3, Panel b, which essentially consists of one Mach-Zehnder interferometer for each photon. The phase of the interferometers sets the measurement bases  $\{\frac{1}{\sqrt{2}}(|0\rangle + e^{-i\phi_i}|1\rangle), \frac{1}{\sqrt{2}}(|0\rangle - e^{-i\phi_i}|1\rangle)\}$ . Using these two interferometers we can measure all what is required for a CHSH pa-





**Figure 3 Entangled photon-pair source.** **a)** *The source* — The beam from a Toptica DL Pro HP 426 laser is focused on a 30-mm-long PPKTP crystal, phase-matched for degenerate collinear type-II SPDC from 426 nm to 852 nm. The phase-matching is finely tuned by controlling the temperature of the crystal with a precision greater than 0.01K. The emitted photons have a bandwidth of approximately 0.2 nm. After the crystal, the residual pump beam is filtered, the photons are then collimated and sent to a set-up to create entanglement by post-selection (as explained in the main text). The entanglement is first produced in polarization and then converted into path using polarizing beam splitters. The source produces  $\approx 30,000$  path-entangled photon pairs per second with a pump power of 8 mW. **b)** *Set-up used to measure a Bell Inequality on the path qubits* — The two paths composing each qubit are interfered on a beam splitter (BS) projecting each qubit onto a basis on the equator of the Bloch sphere (see main text for more details).

parameter:

$$S = \left| C(o_1, o_2) + C(o'_1, o_2) + C(o_1, o'_2) - C(o'_1, o'_2) \right|, \quad (21)$$

where

$$C(o_1, o_2) = \frac{N_{++} - N_{+-} - N_{-+} + N_{--}}{N_{++} + N_{+-} + N_{-+} + N_{--}}. \quad (22)$$

Here,  $N_{++}$  is the number of coincidence events between detectors labelled + for each photon in Fig. 3, Panel **b**,  $N_{+-}$  the number of coincidence events between detectors + and - for each photon, and so on.

Fig. 4 shows the characterization of the entanglement of the joint input control state, where we verified the initial entanglement by performing a Bell measurement on the joint control system before the quantum-switch, obtaining a CHSH parameter of  $2.58 \pm 0.09$ .

## F Experimental proof that IIb does not hold in GPTs under the assumption of III

In this section, we experimentally prove that the form of violation IIb does not hold for the case in which assumption III holds. Form of violation IIb states that the laboratory operations may couple (i.e., may generate non-local correlations between) the control and the target subsystems, within a given party  $S_i$ ,  $i = 1, 2$ . To test it, we first prepare the control and target subsystems in a tomographically-complete set of product states within quantum theory, i.e.,

$$p(o^c, o^t | m^c, m^t, \omega_i = \omega^c \omega^t)$$

$$= p(o^c | m^c, \omega^c) \cdot p(o^t | m^t, \omega^t) \quad (23)$$

for all states  $\omega_i = \omega^c \omega^t$  from a tomographically-complete set of product states. In the GPTs, Eq. (23) shows that, for a product state from the quantum subspace, the probability for a joint outcome factorizes into the product of the probabilities for individual outcomes. We then set a single quantum-switch to have only one operation inserted, either  $U_{i_A}$  or  $U_{i_B}$ . We finally verify that, for the full set of preparations, the control and target subsystems are still in a product state after the quantum-switch, when this contains only  $U_{i_A}$  or only  $U_{i_B}$ . More precisely, we verify that

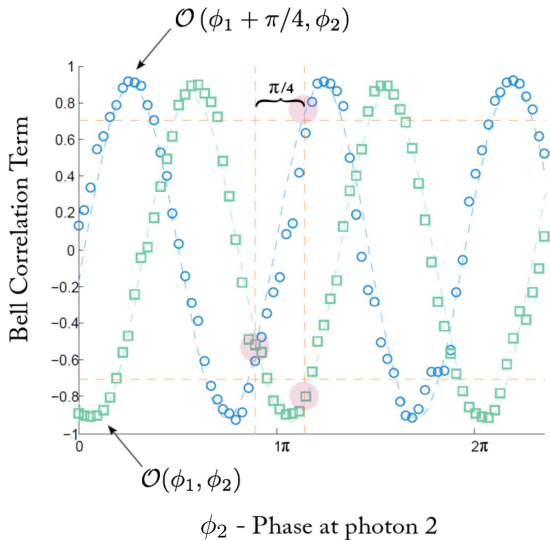
$$p(o^c, o^t | m^c, m^t, U_{i_A} \omega_i) = p(o^c | m^c, U_{i_A} \omega_c) p(o^t | m^t, U_{i_A} \omega_t), \quad (24a)$$

$$p(o^c, o^t | m^c, m^t, U_{i_B} \omega_i) = p(o^c | m^c, U_{i_B} \omega_c) p(o^t | m^t, U_{i_B} \omega_t), \quad (24b)$$

for any state from a complete set of product states, and by linear extension to an arbitrary product state  $\omega_i$ . We do this using the same technique we used to verify that the target qubits began in an input state (Appendix D). Finally, we make use of the following property: if neither operation  $U_{i_A}$  nor  $U_{i_B}$  alone couple the two subsystems, then also a sequence of the two operations cannot couple them as long as they are performed in a definite causal order. This conclusion follows directly from Eqs. (23)-(24b):

$$p(o_c, o_t | m_c, m_t, U_{i_B} \circ U_{i_A} \omega_i) = p(o_c | m_c, U_{i_B} \circ U_{i_A} \omega_c) p(o_t | m_t, U_{i_B} \circ U_{i_A} \omega_t). \quad (25)$$

Note that even under a multiple usage of  $U_{i_A}$  and  $U_{i_B}$  there can be no coupling when the operations



**Figure 4 Input control state characterization: Bell measurement on the order qubits.** Each curve is a measurement of a Bell correlation term  $C(\mathcal{O}_1(\phi_1), \mathcal{O}_2(\phi_2))$  on the control qubits, wherein the phase of  $\phi_1$  is fixed, and the phase  $\phi_2$  is scanned. As described in Eq. (21) of Appendix E, we test the Clauser-Horne-Shimony-Holt (CHSH) inequality [15] achieving a violation of  $2.59 \pm 0.09$ . For the data in the green curve, the phase  $\phi_1$  was nominally shifted by  $\pi/4$  rad with respect to the blue curve. The red shaded areas represent the regions where values of  $\phi_1$  and  $\phi_2$  correspond with those used to construct our CHSH parameter (Eq. (21) of Appendix E). In particular,  $\mathcal{O} = (\mathcal{O}_1, \mathcal{O}_2)$  where  $\mathcal{O}_i(\phi_i) = \cos(\phi_i) \sigma_x + \sin(\phi_i) \sigma_z$ . These data confirm that the two photons start in a path-entangled state, and the polarization state is initially separable.

are performed in a definite causal order. This finalizes the proof of assumption IIb.

Tables S2-S3 report the values of the probabilities  $p(o_c, o_t | m_c, m_t, \omega_1)$  (which, for brevity, are indicated as  $p_{c,t}$  in the Tables) compared with the marginal probability products  $p(o_c | m_c, \omega_1^c) \cdot p(o_t | m_t, \omega_1^t)$  (denoted as  $p_c \cdot p_t$  in the Tables).

The tomographically-complete sets of fiducial quantum measurements reported in Tables S2-S3 were performed as follows. In order to vary the input state of the control system among  $|+\rangle^c$ ,  $|-\rangle^c$ ,  $|R\rangle^c = (|0\rangle^c - i|1\rangle^c)/\sqrt{2}$ , and  $|L\rangle^c = (|0\rangle^c + i|1\rangle^c)/\sqrt{2}$ , we set the relative phase between the two trajectories after the first beamsplitter by means of a delay stage mounted on a calibrated piezo-actuator. Instead, by blocking either path, we prepared  $|0\rangle^c$  and  $|1\rangle^c$ . Likewise, we measure the path qubit in the following way. To measure in  $\{|+\rangle^c, |-\rangle^c\}$ , or  $\{|R\rangle^c, |L\rangle^c\}$ , we suitably set the relative phase between the

two paths before recombining them at the second beamsplitter. This can be done by adding the required phase for state preparation and subtracting the phase for state measurement. Such a phase is then converted into a path delay and sent to the piezo-actuated delay stage. (We use the same delay stage to both set the phase of the path state, and to measure it in  $\{|+\rangle^c, |-\rangle^c\}$ , or  $\{|R\rangle^c, |L\rangle^c\}$ .) To measure in the  $\{|0\rangle, |1\rangle\}$  basis, we block either path before the 50/50 beamsplitter, and we then sum the counts from the two paths after the beamsplitter.

The displayed output probabilities  $p(o_c, o_t | m_c, m_t, \omega_1)$  are very close to those corresponding to a product state. This is indicated by the fact that the RMS distance [Eq. (19)] between these two sets of probabilities (the measured joint probabilities  $p(o_c, o_t | m_c, m_t, \omega_1)$ , and that given by the product  $p(o_c | m_c, \omega_1^c) \cdot p(o_t | m_t, \omega_1^t)$ ) is  $(2 \pm 3) \cdot 10^{-2}$  when only operation  $U_{i_A}$  was acting on the system, and  $(3 \pm 3) \cdot 10^{-2}$  when only operation  $U_{i_B}$  was present. Moreover, the displayed output probabilities  $p(o_c, o_t | m_c, m_t, \omega_1)$  are very close to those of a pure (product) state, which means that there cannot be any correlations in the non-quantum subspace. More precisely, the non-vanishing discrepancy between the two probability distributions could be caused by a weak coupling between the control and the target system. Using the same technique as in Appendix D, we can estimate that the correlations established through this coupling are too weak to violate the CHSH inequality ( $2\sqrt{t_1^2 + t_2^2} = 1.76 \pm 0.04$ ). The coupling between each pair of the control and the target systems can “swap” the correlations from the bipartite state of the control system to that of the target system. However, assuming that the transferred amount of correlations to the target system cannot be larger than the amount produced through the coupling between each pair of the target and control system, we conclude that the coupling cannot result in the target systems violating the Bell’s inequality.

This confirms within experimental error, under the hypothesis that assumption III is valid, that the form of violation IIb does not occur in our experiment. Furthermore, this proof holds not only within quantum theory but also for our class of GPTs. In other words, our measurements imply that in both of the two quantum-switches, indi-

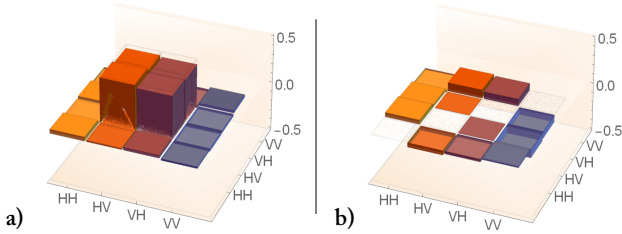
vidually, the laboratory operations do not couple the target and the control subsystems in GPTs when these operations are executed in a definite causal order. From this experimental test, we thus conclude that in our experiment assumption II cannot be false unless assumption III is also violated.

## G Experimental Tests assuming a Quantum Mechanical Description

Although our results are not based on the assumption of a quantum description of the experiment, below we report the analysis of the results of our experiment within this description for completeness.

### G.1 Entanglement of the output target state

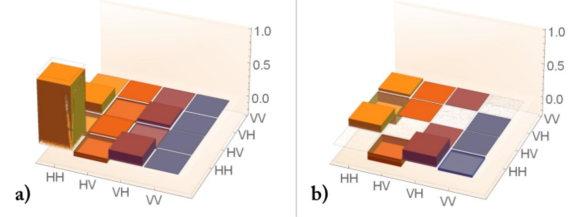
In this subsection, we report the results of the polarization-state tomography on the two-qubit output target state after the quantum-switches. The resulting density matrix is presented in Fig. 5, and it shows a clear presence of entanglement. The reconstructed state has a fidelity of  $0.922 \pm 0.005$  with the ideal one [Eq. (6)], and a concurrence of  $0.95 \pm 0.01$ .



**Figure 5 Output state characterization.** Panels **a** and **b** show the real and imaginary parts, respectively, of the two-photon polarization state measured after the photons leave the quantum-switches. For the data shown here, the two control qubits were found to be in the same state (either  $|+\rangle_1^c |+\rangle_2^c$  or  $|-\rangle_1^c |-\rangle_2^c$ ). This state has a fidelity of  $0.922 \pm 0.005$  with the target state  $(|HV\rangle + |VH\rangle)/\sqrt{2}$ , and a concurrence of  $0.95 \pm 0.01$ . Performing a Bell measurement directly using this state results in a CHSH parameter of  $2.55 \pm 0.08$ .

### G.2 Verification of assumption I

Within quantum theory, one can show the validity of assumption I by demonstrating that the state is separable; this can be done using quantum state tomography, for example. To this end,



**Figure 6 Input control state characterization: State tomography of the target qubits.** The real (Panel **a**) and imaginary (Panel **b**) parts of the two-photon polarization state are measured before the two photons enter the quantum-switches. This state has a fidelity  $0.935 \pm 0.004$  with the ideal state  $|HH\rangle$ , and a concurrence of  $0.001 \pm 0.010$ .

we performed tomography on the target states before the quantum-switches. The resulting density matrix is shown in Fig. 6, Panels **a** and **b**. For our experiment, the target state was nominally prepared in  $|HH\rangle$ ; our measured state has a fidelity of  $0.935 \pm 0.004$  with  $|HH\rangle$ . Furthermore, the concurrence of the estimated state is  $0.001 \pm 0.010$ , indicating that, within experimental error, the initial target state is separable, in agreement with assumption I. The error bars are computed using a Monte Carlo simulation of our experiment; the dominant contribution comes from errors in setting the WPs, and cross-talk in the polarizing BSs.

## H Data analysis

In order to convert the coincidence counts into probabilities, we weight each measured count rate by the net detection efficiency of the corresponding detector pair. We estimate these efficiencies in two parts. First, we measure the relative coupling efficiencies between the output ports  $M_1$  and  $M_2$  of quantum-switch S1, and  $M_3$  and  $M_4$  of quantum-switch S2. Then, within each output port, we measure the relative efficiency of the detector in the transmitted port and the reflected port. We find relative efficiencies between  $\approx 0.85$  and 1. For more details, see the Methods of our previous work [40].

The main source of error in our experiment was phase fluctuations. In the Bell measurement, this dephasing is mainly due to two contributions. (1) Undesired phase-shifts in the interferometer (which we estimated to be about  $0.97^\circ \pm 0.02^\circ$ ). (2) Fluctuations of the source,

which produces time varying phase between the  $|HH\rangle$  and  $|VV\rangle$  terms. In our source, we estimate this to be approximately  $1.9^\circ \pm 0.1^\circ$ , which is caused by a combination of fluctuations in the pump laser wavelength, and the phase-matching temperature. We convert these errors into an error in the Bell parameter using Gaussian error propagation. To calculate the error for the Bell measurements on the polarization qubits after the quantum-switches, we consider the same error sources as above (where now the phase shifts in the measurement interferometer are replaced by phase shifts in the quantum-switches). However, we also consider errors arising from setting the polarization measurements. Finally, to estimate the errors in the results extracted from tomography (i.e., fidelity and concurrence), we performed a Monte Carlo simulation considering the phase fluctuations discussed above.

## I Additional consistency tests: the insertion of noise

In this section, we present two further tests of consistency of our experimental proof.

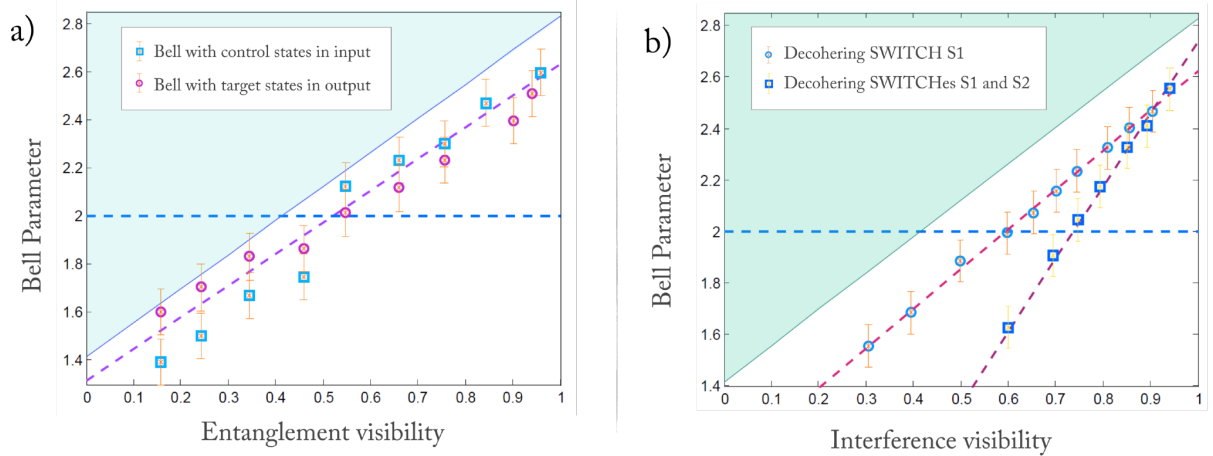
First, we decreased the entanglement of the joint control system by increasing the delay of the interferometer inside the source (see Appendix E). The more mixed the state of the control system becomes, the smaller is the amount of violation of a Bell inequality with the target systems which we can achieve, up until reaching the threshold of non-violation. The Bell parameter versus the “source visibility” (i.e., the two-photon visibility in its anti-correlated basis) is plotted in Fig. 7a. The dashed line is a calculation of the expected Bell parameter, including the imperfect visibility of the two interferometers. All the data points agree with the expected trend within error. The small step at an entanglement visibility of around 0.5 was caused by a lower fringe visibility which increased the systematic error in setting the phases  $\phi_1$  and  $\phi_1 + \pi/4$  (see Fig. 6).

As a second test, we decreased the degree of causal non-separability of the two processes. To do this, we introduced distinguishing information between the paths corresponding to the orders  $U_{i_A} \preceq U_{i_B}$  and  $U_{i_B} \preceq U_{i_A}$  (in only one quantum-switch, squares in Fig. 7b, and in both simultaneously, circles in Fig. 7b) by lengthening one of the paths with respect to the other, effectively

reducing the visibility of the interferometers comprising the quantum-switches. As this occurs, we transition from a superposition of temporal orders to a mixture of them (in other words, to a causally-separable process, which satisfies assumption III). If all three assumptions are met, one cannot violate a Bell inequality between the two systems. Indeed, we experimentally observe that as the visibility is decreased, the Bell parameter also decreases (Fig. 7, Panel b).

## References

- [1] We test hypothesis IIb by placing the two operations individually in a single quantum switch, and we show that each operation is local. Under the assumption of a definite causal order, this test implies that the sequential action of the two operations is also local. We could not have directly tested the combined action of the two operations because such a test could only show that control and target are coupled, but would not provide any information about whether this is due to the violation of IIb, or due to the indefinite causal order, or both.
- [2] Alastair A. Abbott, Julian Wechs, Fabio Costa, and Cyril Branciard. Genuinely multipartite noncausality. *Quantum*, 1:39, December 2017. ISSN 2521-327X. DOI: 10.22331/q-2017-12-14-39. URL <https://doi.org/10.22331/q-2017-12-14-39>.
- [3] Philippe Allard Guérin and Časlav Brukner. Observer-dependent locality of quantum events. *New Journal of Physics*, 20(10):103031, oct 2018. DOI: 10.1088/1367-2630/aae742.
- [4] Mateus Araújo, Fabio Costa, and Časlav Brukner. Computational advantage from quantum-controlled ordering of gates. *Phys. Rev. Lett.*, 113:250402, Dec 2014. DOI: 10.1103/PhysRevLett.113.250402.
- [5] Mateus Araújo, Cyril Branciard, Fabio Costa, Adrien Feix, Christina Giarmatzi, and Časlav Brukner. Witnessing causal nonseparability. *New Journal of Physics*, 17(10):102001, 2015. DOI: 10.1088/1367-2630/17/10/102001.
- [6] Alain Aspect, Jean Dalibard, and Gérard Roger. Experimental test of bell’s inequalities using time-varying analyzers.



**Figure 7 Bell parameter in the presence of various decoherence sources.** **a)** For these data, the initial entanglement of control qubits is decreased passing from the entangled state  $\frac{1}{\sqrt{2}}(|0,0\rangle - |1,1\rangle)$  to a mixture of  $|0,0\rangle$  and  $|1,1\rangle$ . We measure the Bell parameter both on the input path qubits (squares) and output polarization qubits (circles) as the source is decohered. Here, the Bell parameter is plotted versus the visibility of the entangled state in its anti-correlated basis. The dashed line is a simulation of the experimental results. **b)** For these data, the coherence of the superposition of the orders of operations inside the quantum-switches is decreased, leading to a classical mixture of orders. To control this transition, we decrease the visibility of either only one of the two interferometers (circles), or of both interferometers at the same time (squares). Each graph shows the Bell parameter plotted versus the visibility of one interferometer. The dashed lines are linear fits to the data. The horizontal dashed blue line, in both plots, is the classical limit for a Bell violation. When the state of the control qubit is too decohered, we can no longer violate a Bell inequality.

*Phys. Rev. Lett.*, 49:1804–1807, Dec 1982. DOI: 10.1103/PhysRevLett.49.1804. URL <https://link.aps.org/doi/10.1103/PhysRevLett.49.1804>.

- [7] Jonathan Barrett. Information processing in generalized probabilistic theories. *Phys. Rev. A*, 75:032304, Mar 2007. DOI: 10.1103/PhysRevA.75.032304. URL <https://link.aps.org/doi/10.1103/PhysRevA.75.032304>.
- [8] J. S. Bell. On the einstein-podolsky-rosen paradox. *Physics*, 1(3):195–200, 1964.
- [9] Cyril Branciard. Witnesses of causal non-separability: an introduction and a few case studies. *Scientific Reports*, 6:26018, May 2016. DOI: 10.1038/srep26018.
- [10] Cyril Branciard, Mateus Araújo, Adrien Feix, Fabio Costa, and Časlav Brukner. The simplest causal inequalities and their violation. *New Journal of Physics*, 18(1):013008, 2016. DOI: 10.1088/1367-2630/18/1/013008. URL <http://stacks.iop.org/1367-2630/18/i=1/a=013008>.
- [11] Časlav Brukner. Quantum causality. *Nat. Phys.*, 10:259–263, Apr 2014. DOI: 10.1038/nphys2930.
- [12] Nicolas Brunner, Daniel Cavalcanti, Stefano Pironio, Valerio Scarani, and Stephanie Wehner. Bell nonlocality. *Rev. Mod. Phys.*, 86:419–478, Apr 2014. DOI: 10.1103/RevModPhys.86.419. URL <https://link.aps.org/doi/10.1103/RevModPhys.86.419>.
- [13] Giulio Chiribella. Perfect discrimination of no-signalling channels via quantum superposition of causal structures. *Phys. Rev. A*, 86:040301, Oct 2012. DOI: 10.1103/PhysRevA.86.040301.
- [14] Giulio Chiribella, Giacomo Mauro D’Ariano, Paolo Perinotti, and Benoit Valiron. Quantum computations without definite causal structure. *Phys. Rev. A*, 88:022318, Aug 2013. DOI: 10.1103/PhysRevA.88.022318.
- [15] John F. Clauser, Michael A. Horne, Abner Shimony, and Richard A. Holt. Proposed experiment to test local hidden-variable theories. *Phys. Rev. Lett.*, 23:880–884, Oct 1969. DOI: 10.1103/PhysRevLett.23.880. URL <https://link.aps.org/doi/10.1103/PhysRevLett.23.880>.
- [16] Borivoje Dakic and Časlav Brukner. *Quantum Theory and Beyond: Is Entanglement Special?* Contribution to “Deep beauty”, Editor Hans Halvorson (Cam-

- bridge University Press, 2010. DOI: 10.1017/CBO9780511976971.011.
- [17] Lu-Ming Duan, G. Giedke, J. I. Cirac, and P. Zoller. Inseparability criterion for continuous variable systems. *Phys. Rev. Lett.*, 84:2722–2725, Mar 2000. DOI: 10.1103/PhysRevLett.84.2722. URL <https://link.aps.org/doi/10.1103/PhysRevLett.84.2722>.
- [18] Adrien Feix, Mateus Araújo, and Časlav Brukner. Quantum superposition of the order of parties as a communication resource. *Phys. Rev. A*, 92:052326, Nov 2015. DOI: 10.1103/PhysRevA.92.052326.
- [19] Stuart J. Freedman and John F. Clauser. Experimental test of local hidden-variable theories. *Phys. Rev. Lett.*, 28:938–941, Apr 1972. DOI: 10.1103/PhysRevLett.28.938. URL <https://link.aps.org/doi/10.1103/PhysRevLett.28.938>.
- [20] Marissa Giustina, Marijn A. M. Versteegh, Sören Wengerowsky, Johannes Handsteiner, Armin Hochrainer, Kevin Phelan, Fabian Steinlechner, Johannes Kofler, Jan-Åke Larsson, Carlos Abellán, Waldimar Amaya, Valerio Pruneri, Morgan W. Mitchell, Jörn Beyer, Thomas Gerrits, Adriana E. Lita, Lynden K. Shalm, Sae Woo Nam, Thomas Scheidl, Rupert Ursin, Bernhard Wittmann, and Anton Zeilinger. Significant-loophole-free test of bell’s theorem with entangled photons. *Phys. Rev. Lett.*, 115:250401, Dec 2015. DOI: 10.1103/PhysRevLett.115.250401. URL <https://link.aps.org/doi/10.1103/PhysRevLett.115.250401>.
- [21] K. Goswami, C. Giarmatzi, M. Kewming, F. Costa, C. Branciard, J. Romero, and A. G. White. Indefinite causal order in a quantum switch. *Phys. Rev. Lett.*, 121:090503, Aug 2018. DOI: 10.1103/PhysRevLett.121.090503. URL <https://link.aps.org/doi/10.1103/PhysRevLett.121.090503>.
- [22] Philippe Allard Guérin, Adrien Feix, Mateus Araújo, and Časlav Brukner. Exponential communication complexity advantage from quantum superposition of the direction of communication. *Phys. Rev. Lett.*, 117:100502, Sep 2016. DOI: 10.1103/PhysRevLett.117.100502.
- [23] Lucien Hardy. Quantum theory from five reasonable axioms. *Preprint at arXiv:0101012 [quant-ph]*, 2001.
- [24] Lucien Hardy. *Foliable Operational Structures for General Probabilistic Theories*, page 409–442. Cambridge University Press, 2011. DOI: 10.1017/CBO9780511976971.013.
- [25] B. Hensen, H. Bernien, A. E. Dreau, A. Reiserer, N. Kalb, M. S. Blok, J. Ruitenberg, R. F. L. Vermeulen, R. N. Schouten, C. Abellan, W. Amaya, V. Pruneri, M. W. Mitchell, M. Markham, D. J. Twitchen, D. Elkouss, S. Wehner, T. H. Taminau, and R. Hanson. Loophole-free bell inequality violation using electron spins separated by 1.3 kilometres. *Nature*, 526:682–686, Oct 2015. DOI: 10.1038/nature15759. URL <http://dx.doi.org/10.1038/nature15759>.
- [26] John C. Howell, Ryan S. Bennink, Sean J. Bentley, and R. W. Boyd. Realization of the einstein-podolsky-rosen paradox using momentum- and position-entangled photons from spontaneous parametric down conversion. *Phys. Rev. Lett.*, 92:210403, May 2004. DOI: 10.1103/PhysRevLett.92.210403. URL <https://link.aps.org/doi/10.1103/PhysRevLett.92.210403>.
- [27] P. G. Kwiat, A. M. Steinberg, and R. Y. Chiao. High-visibility interference in a bell-inequality experiment for energy and time. *Phys. Rev. A*, 47:R2472–R2475, Apr 1993. DOI: 10.1103/PhysRevA.47.R2472. URL <https://link.aps.org/doi/10.1103/PhysRevA.47.R2472>.
- [28] M. Lamehi-Rachti and W. Mittig. Quantum mechanics and hidden variables: A test of bell’s inequality by the measurement of the spin correlation in low-energy proton-proton scattering. *Phys. Rev. D*, 14:2543–2555, Nov 1976. DOI: 10.1103/PhysRevD.14.2543. URL <https://link.aps.org/doi/10.1103/PhysRevD.14.2543>.
- [29] Jean-Philippe W. MacLean, Katja Ried, Robert W. Spekkens, and Kevin J. Resch. Quantum-coherent mixtures of causal relations. *Nature Communications*, 8:15149, May 2017. DOI: 10.1038/ncomms15149.
- [30] Ll. Masanes, M. P. Müller, D. Pérez-García, and R. Augusiak. Entanglement and the three-dimensionality of the bloch ball. *Journal of Mathematical Physics*, 55(12):122203,

2014. DOI: 10.1063/1.4903510. URL <https://doi.org/10.1063/1.4903510>.
- [31] Lluís Masanes and Markus P. Müller. A derivation of quantum theory from physical requirements. *New Journal of Physics*, 13(6):063001, 2011. DOI: 10.1088/1367-2630/13/6/063001. URL <http://stacks.iop.org/1367-2630/13/i=6/a=063001>.
- [32] Nikolai Miklin, Alastair A Abbott, Cyril Branciard, Rafael Chaves, and Costantino Budroni. The entropic approach to causal correlations. *New Journal of Physics*, 19(11):113041, 2017. DOI: 10.1088/1367-2630/aa8f9f. URL <http://stacks.iop.org/1367-2630/19/i=11/a=113041>.
- [33] Ognjan Oreshkov. Time-delocalized quantum subsystems and operations: on the existence of processes with indefinite causal structure in quantum mechanics. *Quantum*, 3:206, December 2019. ISSN 2521-327X. DOI: 10.22331/q-2019-12-02-206. URL <https://doi.org/10.22331/q-2019-12-02-206>.
- [34] Ognjan Oreshkov and Christina Giarmatzi. Causal and causally separable processes. *New Journal of Physics*, 18(9):093020, 2016. DOI: 10.1088/1367-2630/18/9/093020.
- [35] Ognjan Oreshkov, Fabio Costa, and Časlav Brukner. Quantum correlations with no causal order. *Nat. Commun.*, 3:1092, Oct 2012. DOI: 10.1038/ncomms2076.
- [36] Lorenzo M. Procopio, Amir Moqanaki, Mateus Araújo, Fabio Costa, Irati Alonso Calafell, Emma G. Dowd, Deny R. Hamel, Lee A. Rozema, Časlav Brukner, and Philip Walther. Experimental superposition of orders of quantum gates. *Nat. Commun.*, 6(7913), Aug 2015. DOI: 10.1038/ncomms8913.
- [37] J. G. Rarity and P. R. Tapster. Experimental violation of bell’s inequality based on phase and momentum. *Phys. Rev. Lett.*, 64:2495–2498, May 1990. DOI: 10.1103/PhysRevLett.64.2495. URL <https://link.aps.org/doi/10.1103/PhysRevLett.64.2495>.
- [38] Martin J. Renner and Časlav Brukner. Reassessing the computational advantage of quantum-controlled ordering of gates. *Phys. Rev. Research*, 3:043012, Oct 2021. DOI: 10.1103/PhysRevResearch.3.043012. URL <https://link.aps.org/doi/10.1103/PhysRevResearch.3.043012>.
- [39] M. A. Rowe, D. Kielpinski, V. Meyer, C. A. Sackett, W. M. Itano, C. Monroe, and D. J. Wineland. Experimental violation of a bell’s inequality with efficient detection. *Nature*, 409:791, Feb 2001. DOI: 10.1038/35057215. URL <http://dx.doi.org/10.1038/35057215>.
- [40] Giulia Rubino, Lee A. Rozema, Adrien Feix, Mateus Araújo, Jonas M. Zeuner, Lorenzo M. Procopio, Časlav Brukner, and Philip Walther. Experimental verification of an indefinite causal order. *Science Advances*, 3(3), 2017. DOI: 10.1126/sciadv.1602589. URL <http://advances.sciencemag.org/content/3/3/e1602589>.
- [41] Lynden K. Shalm, Evan Meyer-Scott, Bradley G. Christensen, Peter Bierhorst, Michael A. Wayne, Martin J. Stevens, Thomas Gerrits, Scott Glancy, Deny R. Hamel, Michael S. Allman, Kevin J. Coakley, Shellee D. Dyer, Carson Hodge, Adriana E. Lita, Varun B. Verma, Camilla LAMBROCCO, Edward Tortorici, Alan L. Migdall, Yanbao Zhang, Daniel R. Kumor, William H. Farr, Francesco Marsili, Matthew D. Shaw, Jeffrey A. Stern, Carlos Abellán, Waldimar Amaya, Valerio Pruneri, Thomas Jennewein, Morgan W. Mitchell, Paul G. Kwiat, Joshua C. Bienfang, Richard P. Mirin, Emanuel Knill, and Sae Woo Nam. Strong loophole-free test of local realism. *Phys. Rev. Lett.*, 115:250402, Dec 2015. DOI: 10.1103/PhysRevLett.115.250402. URL <https://link.aps.org/doi/10.1103/PhysRevLett.115.250402>.
- [42] R. Simon. Peres-horodecki separability criterion for continuous variable systems. *Phys. Rev. Lett.*, 84:2726–2729, Mar 2000. DOI: 10.1103/PhysRevLett.84.2726. URL <https://link.aps.org/doi/10.1103/PhysRevLett.84.2726>.
- [43] Magdalena Zych, Fabio Costa, Igor Pikovski, and Časlav Brukner. Bell’s theorem for temporal order. *Nature Communications*, 10:3772, 08 2019. DOI: 10.1038/s41467-019-11579-x. URL <https://doi.org/10.1038/s41467-019-11579-x>.

# Supplementary Information

## I. Quantum-switch and causal inequalities

The quantum-switch [36, 40] has been shown not to violate causal inequalities, making it impossible to use such a violation as a theory-independent proof that the causal order of the operations in the quantum-switch is indefinite. Here, we briefly re-examine such reasoning following Refs. [5, 34].

We introduce the  $x$ ,  $y$  and  $z$  indices to refer, respectively, to the measurements choices of Alice, Bob and Charlie. We call  $a$ ,  $b$  and  $c$  their respective measurement results. It is always possible to re-write  $p(a, b, c|x, y, z)$  as

$$p^{\text{switch}}(a, b, c|x, y, z) = p(c|a, b, x, y, z) p(a, b|x, y, z). \quad (26)$$

It should be noticed that, regardless of the causal order between operations in Alice's and Bob's laboratory, the operation in Charlie's laboratory always occurs after them. In other words, his operation is in the future light cone of both Alice's and Bob's operations. Thus,  $a$  and  $b$  cannot depend on  $z$ , so

$$p(a, b|x, y, z) = p(a, b|x, y). \quad (27)$$

As previously observed, after tracing out Charlie's laboratory in the quantum-switch, the process of Alice and Bob is causally separable. Thus, one can rewrite  $p(a, b|x, y)$  in the form of a convex mixture, obtaining

$$p^{\text{switch}}(a, b, c|x, y, z) = p(c|a, b, x, y, z) [\zeta \cdot p^{A \preceq B}(a, b|x, y) + (1 - \zeta) \cdot p^{B \preceq A}(a, b|x, y)], \quad (28)$$

with  $\zeta \geq 0$ . We can combine the probabilities  $p^{A \preceq B}(a, b|x, y)$  ( $p^{B \preceq A}(a, b|x, y)$ ) and  $p(c|a, b, x, y, z)$  as a product of the probability respecting the order  $A \preceq B$  ( $B \preceq A$ ) with the probability respecting the order  $\{A, B\} \preceq C$

$$p^{\text{switch}}(a, b, c|x, y, z) = \zeta \cdot p^{A \preceq B \preceq C}(a, b, c|x, y, z) + (1 - \zeta) \cdot p^{B \preceq A \preceq C}(a, b, c|x, y, z). \quad (29)$$

Therefore, the quantum-switch is a process whose probabilities have a ‘‘causal model’’, i.e., it can always be understood as arising from events that are causally ordered, or from a convex mixture of causally ordered events. Thus, it satisfies all causal inequalities.

## II. Hidden local definite causal order

In general, while experimental tests can be used to prove that the conjunction of the assumptions underlying a given no-go theorem does not describe the phenomenology observed within quantum mechanics, they do not provide information on which of the assumptions is to be discarded. In this experiment, the application of the Bell's theorem to temporal order allowed us to test a conjunction of all our assumptions; yet, in order to verify which assumptions are valid, additional tests on a single quantum-switch were necessary. This notwithstanding, it is worth noting that testing only one single quantum-switch would not have provided an as stringent information. In fact, as we showed above the experimental data taken from a single quantum-switch cannot violate causal inequalities, and thus can be understood as arising from an underlying causal model, in the spirit of simulation of quantum statistics by hidden variables. Such a model generates statistics compatible with operations performed on a system in a definite order, or in a convex mixture therefrom. In terms of probabilities, the statistics in the quantum-switch  $p^{\text{switch}}(a, b, c|x, y, z) = \int d\lambda \rho(\lambda) p^{\text{causal}}(a, b, c|x, y, z, \lambda)$ , where  $p^{\text{causal}}(a, b, c|x, y, z, \lambda) = p^{A \preceq B \preceq C}(a, b, c|x, y, z)$  or  $p^{B \preceq A \preceq C}(a, b, c|x, y, z)$  or a classical mixture therefrom, could therefore be mimicked by an underlying causal hidden variable model. The statistics obtained measuring the double quantum-switch with entangled temporal orders rules out a *local* causal hidden variable model that allows for this description, i.e., its statistics is incompatible with

$$p^{2\text{-switches}}(a_1, b_1, c_1, a_2, b_2, c_2|x_1, y_1, z_1, x_2, y_2, z_2) =$$



$$= \int d\lambda \rho(\lambda) p^{\text{causal}}(a_1, b_1, c_1 | x_1, y_1, z_1, \lambda) p^{\text{causal}}(a_2, b_2, c_2 | x_2, y_2, z_2, \lambda). \quad (30)$$

In other words, the causal model is called “local” if the statistics is compatible with the assumptions that: (a) the order of events in each of the laboratories is definite but may be correlated by a common cause in the past, and (b) measurement choices may have only local influences. Our experimental data rule out the models in Eq. (30) for the special case where Alice and Bob both apply a single operation with a single outcome (unitary).

### III. On the physical implementability of the quantum-switch

Skepticism has been expressed about whether a tabletop experiment can demonstrate indefinite causal structures. In Ref. [29], the authors claim that it is not possible to implement the quantum-switch without “exotic physical scenarios”. In particular, they argue that one would need a closed time-like curve, and even then such an implementation would be inconsistent, being able to generate logical contradictions such as the grandfather paradox. These criticisms are based on the assumption that causal structures must be represented via directed graphs. In this representation, the quantum-switch becomes a directed graph with a cycle, which could indeed be inconsistent and could generate logical contradictions.

The tension between directed acyclic graphs (DAG) and causal structures in the quantum-switch is akin to the tension between classical “hidden” variable theories and quantum theory. For example, in order to describe an interferometric experiment in terms of classical variables, one is forced to say that the interfering system follows some exotic trajectories or in some non-local manner follows two classical trajectories “at once.” However, within quantum theory one interprets interferometric tests as demonstrating that the very assumption that a system does follow a definite path is violated.

The formal sense in which the causal order of applying operations in a quantum-switch is non-classical has been recently studied in Ref. [33]. The motivation of that work was to understand where and when the operations happen in the quantum-switch, which is precisely the question brought up in the context of a DAG representation. The author shows that the operations applied on systems in a quantum-switch act on subsystems that are not localised in time, i.e., on ‘time-delocalised’ subsystems. It is further shown that standard quantum theory, without exotic closed timelike-curves, is compatible with such time-delocalised operations and that they indeed realise genuine non-separable quantum processes. The work also concludes that experimental realisations of the quantum-switch, including specifically its entangled version described in this work, are genuine realisations of such time-delocalised processes. In other words, there is a well-defined fashion in which temporal relations between the application of operations in a quantum-switch cannot be represented with DAGs. In fact, the Bell theorem for temporal order [43] and its version in this work can be interpreted as a limitation on achievable correlations when operations acting on a quantum system can be embedded in a causal structure compatible with an underlying DAG (or a probabilistic mixture thereof).

Therefore, the suitable conclusion to draw is that the causal structure in the quantum-switch cannot be represented by a DAG since the latter can only represent what are called *definite* causal structures [32]. What our work demonstrates experimentally is that the quantum-switch represents an *indefinite* causal structure incompatible with any DAG, just like experimental violations of Bell’s inequalities show that there exist correlations incompatible with local hidden variables.

The authors of Ref. [29] further argue that, in a genuine quantum-switch, operations must be performed in the same spatio-temporal regions in each term of the superposition, so that only their order is swapped. As mentioned above, Ref. [33] showed that this is not necessary: in the quantum-switch a single operation can be “time-delocalised” over two (or more) spatio-temporal regions. In the originally proposed implementation of the quantum-switch [14], as well as in ours, one could in principle register the time at which the signal passes through each box, which would decohere the superposition and make the interference between the causal orders vanish. However, since we do observe coherence, such information does not exist (i.e., it is not stored in any physical degree of freedom, as this would alter the results of the experiment). The above requirement of “the same spatio-temporal regions” whose order

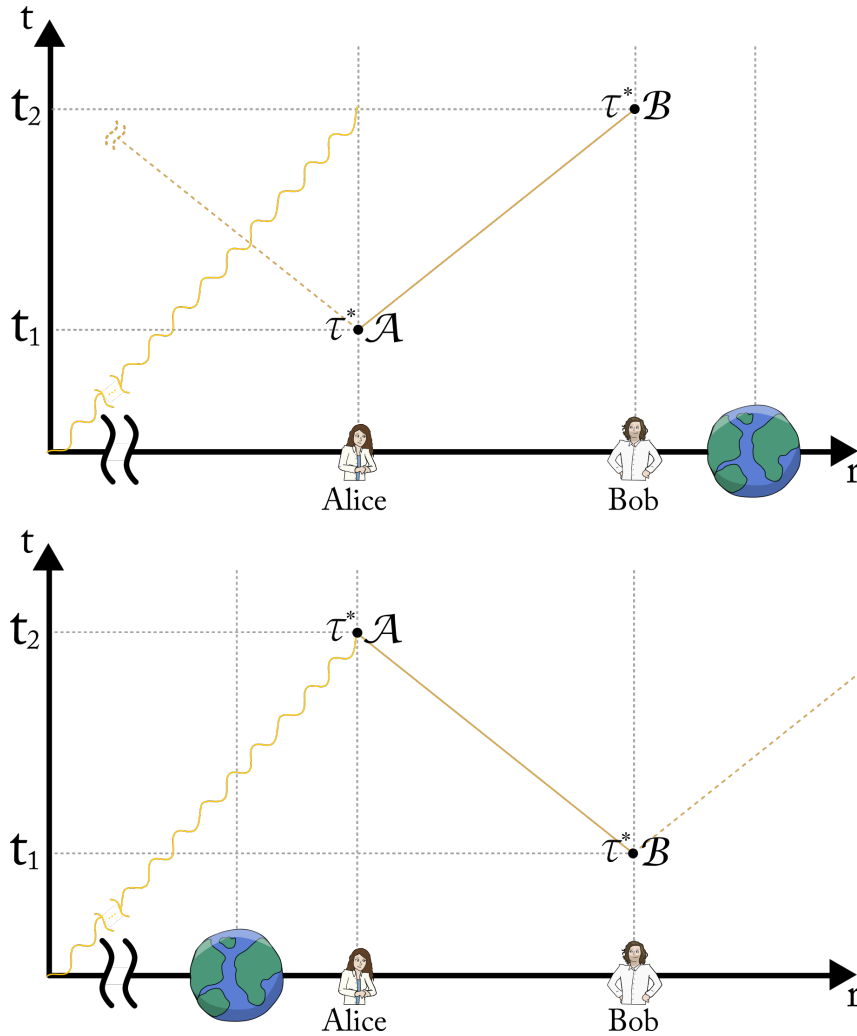
is simply swapped is in principle realisable in a gravitational implementation of the quantum-switch for a certain choice of coordinates. There, a massive object is prepared in a spatial superposition, which results in the causal order between two events being opposite in the two superposed terms. More precisely, a choice of coordinates can be found such that gate  $\mathcal{A}$  is performed at a single time (the proper time of a local clock). However, an alternative choice of coordinates may as well be done such that the gate  $\mathcal{A}$  is performed in a superposition of different times (according to the coordinate time or the time of a distant observer) in different superposed terms, before and after the gate  $\mathcal{B}$ . Therefore, even in this gravitational case, it is always possible to make a choice of coordinates where the operations appear to be performed at different times in the different superposed terms. Thus, it is in fact insubstantial to argue whether the operations are “really performed at the same times, and their order is swapped”, or they are “merely performed in superposition of different times”, as this depends on the choice of coordinates in which one wishes to describe the scenario (see [3, 43] for further discussions). Furthermore, note that, contrary to the arguments of [29], this proposal does not allow for the information to travel back in time, nor does it require closed time-like curves, and it does not give rise to any logical paradoxes.

In Ref. [29], another criticism follows from the observation that the quantum-switch can easily be “simulated” by using additional copies of the boxes  $\mathcal{A}$  and  $\mathcal{B}$ , as was already noted in the original proposal of the quantum-switch [14]. In particular, one could use an unfolded Mach-Zehnder interferometer, with gates  $\mathcal{A}_1$  and  $\mathcal{B}_1$  in one arm and gates  $\mathcal{B}_2$  and  $\mathcal{A}_2$  on the other. The straightforward response is that we do not use an unfolded Mach-Zehnder interferometer, but rather a folded one, and therefore that we use a single copy of each box instead of two. The number of applications of a box can operationally be determined by a counter (i.e., a “flag”) that is raised each time the operation is applied on the system. The very fact that an unfolded Mach-Zehnder interferometer requires two copies of each box to “simulate” the statistics of a folded one is a signature that the latter exhibits an indefinite causal order. Moreover, we also note that, in the unfolded version of the interferometer, it would be necessary to actively make  $\mathcal{A}_1$  precisely equivalent to  $\mathcal{A}_2$ , whereas in our case this clearly follows from the implementation itself, as the gates are physically the same. It should also be emphasized that the present Bell-type proof of an indefinite causal order is valid even if the local gates are used more than once, as clarified in the main text.

Furthermore, following the reasoning in Ref. [29], one could say that it is in principle possible to make the gate  $\mathcal{A}$  ( $\mathcal{B}$ ) act differently when it comes before or after  $\mathcal{B}$  ( $\mathcal{A}$ ), as in each case the photon passes through  $\mathcal{A}$  ( $\mathcal{B}$ ) at different times. This is true, but also applies to the originally proposed implementation. To make it *locally* impossible, one could use the above mentioned superposition of a massive object to control the order of operations in two space-time regions. In such a scenario, indeed, Alice (Bob) in her (his) local laboratory cannot make the gate  $\mathcal{A}$  ( $\mathcal{B}$ ) act differently in case the operation  $\mathcal{A}$  ( $\mathcal{B}$ ) is performed before or after  $\mathcal{B}$  ( $\mathcal{A}$ ). Nevertheless, a distant observer for whom Alice’s (Bob’s) operation happens in a superposition of two coordinate-times could make such contingency occur with a cleverly designed set-up (e.g., by sending a signal which triggers Alice’s operation to change once it is received, as depicted in Fig. 1). As a consequence, as much as in the case of a table-top experiment the operation  $\mathcal{A}$  can be made to act differently depending on whether it happens before or after  $\mathcal{B}$ , in its gravitational counterpart this can be achieved by a distant observer who triggers some change for certain time-coordinates. In conclusion, the requirement that operations  $\mathcal{A}$  and  $\mathcal{B}$  must *even in principle* be forbidden to change depending on the order has no absolute meaning (i.e., it cannot be realized in all reference frames). Moreover, if the operations differed or their time was revealed (or stored in any degree of freedom), the results of the experiment would differ.

Finally, because of the differences highlighted above, one may object that the physics that describes a photonic quantum-switch is not equivalent to the one which is behind the gravitational quantum-switch for all possible observers. This is indeed correct. In fact, in the first case, the physics is described by Maxwell equations on Minkowski space-time, whereas in the latter case it is non-classical space-time that determines the dynamics. However, although a local as well as a global observer could tell the difference between the gravitational and the photonic quantum-switch, it is not the case for a quantum

particle which travels along the two superposed paths in either versions of the quantum-switch. In fact, in both cases the particle experiences a genuine quantum superposition of causal orders. And this is precisely the purpose of this experimental work: we do not aim to draw conclusions concerning global/local observers, but on the system undergoing the quantum process. Therefore, neither of the schemes (i.e., the gravitational and the photonic quantum-switch) is a “simulation” of one another. They are rather two equivalent representations of the dynamics experienced by a quantum particle in presence of a quantum superposition of causal orders, i.e., two representations of a quantum-switch.



**S 1 Schematic of a gravitational quantum-switch.** A quantum system is exchanged between Alice’s and Bob’s laboratories. The order in which such “target” system is exchanged is governed by a second system, a “control” system, which is encoded in the position of a massive object. By putting the massive object in a macroscopic superposition of two positions, one closer to Alice’s and the other closer to Bob’s position, one induces a relative time dilation between Alice’s and Bob’s laboratories. If an outside observer sends some system at a suitably chosen time, let us call it  $t_{A \succ B}$ , the observer could influence the functioning of the device that implements  $\mathcal{A}$ , e.g., when it acts second but not when it acts first, making the operation of Alice act different depending on the order.

**S 1 Comparison between the two-states probabilities  $p(o_1, o_2 | m_1, m_2, \omega_{1,2}^t)$  and the products of marginal single-state probabilities  $p(o_1 | m_1, \omega_1^t) \cdot p(o_2 | m_2, \omega_2^t)$  for the input target states. - Part I.**

The compatibility between the two sets of probabilities shows the separability of the input target state  $\omega_{1,2}^t$ . We indicate with H and V the states of horizontal and vertical polarization, with D and A the diagonal and anti-diagonal states, with R and L the circular polarization states right- and left-handed. The experimental error associated to each of these probabilities is  $\pm 0.01$ .

Measur. Basis	$p_{1,2}$	$p_{1,2^\perp}$	$p_{1^\perp,2}$	$p_{1^\perp,2^\perp}$	$p_1 \cdot p_2$	$p_1 \cdot p_{2^\perp}$	$p_{1^\perp} \cdot p_2$	$p_{1^\perp} \cdot p_{2^\perp}$
H, H	0.97	0.03	0.00	0.00	0.97	0.03	0.00	0.00
H, V	0.01	0.99	0.00	0.00	0.01	0.99	0.00	0.00
H, A	0.58	0.41	0.00	0.00	0.59	0.41	0.00	0.00
H, D	0.42	0.58	0.00	0.00	0.42	0.58	0.00	0.00
H, R	0.39	0.61	0.00	0.00	0.39	0.61	0.00	0.00
H, L	0.61	0.38	0.00	0.00	0.62	0.38	0.00	0.00
V, H	0.00	0.00	0.96	0.04	0.00	0.00	0.96	0.04
V, V	0.00	0.00	0.03	0.97	0.00	0.00	0.03	0.97
V, A	0.00	0.00	0.61	0.39	0.00	0.00	0.61	0.39
V, D	0.00	0.00	0.38	0.61	0.00	0.00	0.38	0.61
V, R	0.00	0.00	0.35	0.64	0.00	0.00	0.35	0.64
V, L	0.00	0.00	0.64	0.36	0.00	0.00	0.64	0.36
A, H	0.39	0.01	0.53	0.02	0.41	0.01	0.54	0.02
A, V	0.01	0.37	0.02	0.54	0.02	0.39	0.02	0.54
A, A	0.24	0.16	0.34	0.21	0.26	0.16	0.34	0.21
A, D	0.18	0.22	0.21	0.33	0.18	0.24	0.23	0.31
A, R	0.16	0.25	0.18	0.35	0.16	0.27	0.19	0.34
A, L	0.26	0.14	0.36	0.19	0.27	0.14	0.36	0.19
D, H	0.55	0.02	0.45	0.02	0.53	0.02	0.45	0.02
D, V	0.01	0.57	0.02	0.45	0.01	0.54	0.01	0.46
D, A	0.32	0.26	0.29	0.18	0.32	0.24	0.28	0.20
D, D	0.23	0.35	0.18	0.29	0.22	0.34	0.18	0.29
D, R	0.21	0.37	0.16	0.31	0.19	0.37	0.16	0.31
D, L	0.35	0.22	0.32	0.17	0.34	0.20	0.31	0.18
R, H	0.65	0.02	0.33	0.01	0.64	0.02	0.33	0.01
R, V	0.01	0.66	0.01	0.30	0.02	0.66	0.01	0.31
R, A	0.39	0.27	0.22	0.13	0.40	0.26	0.21	0.14
R, D	0.29	0.39	0.12	0.19	0.28	0.40	0.13	0.19
R, R	0.27	0.41	0.11	0.20	0.26	0.42	0.12	0.19
R, L	0.41	0.25	0.22	0.12	0.41	0.24	0.22	0.13
L, H	0.32	0.01	0.63	0.04	0.32	0.02	0.63	0.03
L, V	0.01	0.32	0.03	0.64	0.01	0.32	0.02	0.65
L, A	0.18	0.14	0.42	0.27	0.19	0.13	0.41	0.28
L, D	0.14	0.21	0.23	0.41	0.13	0.22	0.24	0.40
L, R	0.12	0.23	0.22	0.43	0.12	0.23	0.22	0.43
L, L	0.21	0.12	0.44	0.24	0.21	0.12	0.43	0.25

**S 2 Comparison between the two-states probabilities  $p(o_c, o_t | m_c, m_t, \omega_1)$  and the products of marginal single-state probabilities  $p(o_c | m_c, \omega_1^c) \cdot p(o_t | m_t, \omega_1^t)$  for the control and the target states when only operation  $U_{i_A}$  is acting on the input state.** We denoted as 0, 1, +, -,  $l$  and  $r$  the analogue of the polarization states H, V, D, A, L, R in the path degree of freedom. The two sets of probabilities associated to the control and the target states in output are compatible within experimental errors. The experimental error associated to each of these probabilities is  $\pm 0.01$ .

Meas. Basis (target)	Prep.-Meas. Basis (control)	$p_{c,t}$	$p_{c,t^\perp}$	$p_{c^\perp,t}$	$p_{c^\perp,t^\perp}$	$p_c \cdot p_t$	$p_c \cdot p_{t^\perp}$	$p_{c^\perp} \cdot p_t$	$p_{c^\perp} \cdot p_{t^\perp}$
H	+	0.95	0.00	0.04	0.00	0.95	0.00	0.04	0.00
D	+	0.47	0.48	0.01	0.03	0.47	0.49	0.02	0.02
R	+	0.48	0.47	0.01	0.03	0.47	0.48	0.02	0.02
H	-	0.07	0.00	0.92	0.01	0.07	0.00	0.91	0.01
D	-	0.04	0.04	0.48	0.44	0.04	0.04	0.48	0.44
R	-	0.04	0.04	0.41	0.51	0.04	0.04	0.41	0.51
H	$r$	0.55	0.00	0.44	0.01	0.55	0.01	0.44	0.00
D	$r$	0.20	0.26	0.28	0.26	0.22	0.24	0.26	0.28
R	$r$	0.28	0.24	0.18	0.30	0.24	0.28	0.22	0.26
H	$l$	0.50	0.00	0.50	0.00	0.50	0.00	0.50	0.00
D	$l$	0.30	0.28	0.21	0.21	0.30	0.28	0.22	0.21
R	$l$	0.27	0.30	0.20	0.23	0.27	0.30	0.20	0.23
H	0	0.51	0.00	0.49	0.00	0.50	0.00	0.49	0.00
D	0	0.26	0.30	0.26	0.17	0.30	0.27	0.23	0.21
R	0	0.28	0.26	0.23	0.23	0.27	0.26	0.24	0.23
H	1	0.56	0.00	0.43	0.01	0.56	0.00	0.44	0.00
D	1	0.27	0.29	0.22	0.23	0.27	0.29	0.21	0.23
R	1	0.28	0.28	0.17	0.27	0.25	0.31	0.20	0.24

**S 3 Comparison between the two-states probabilities  $p(o_c, o_t | m_c, m_t, \omega_1)$  and the products of marginal single-state probabilities  $p(o_c | m_c, \omega_1^c) \cdot p(o_t | m_t, \omega_1^t)$  for the control and the target states when only operation  $U_{i_B}$  is acting on the input state.** The two sets of probabilities associated to the control and the target states in output are compatible within experimental errors. The experimental error associated to each of these probabilities is  $\pm 0.01$ .

Meas. Basis (target)	Prep.-Meas. Basis (control)	$p_{c,t}$	$p_{c,t^\perp}$	$p_{c^\perp,t}$	$p_{c^\perp,t^\perp}$	$p_c \cdot p_t$	$p_c \cdot p_{t^\perp}$	$p_{c^\perp} \cdot p_t$	$p_{c^\perp} \cdot p_{t^\perp}$
H	+	0.47	0.33	0.11	0.09	0.47	0.33	0.12	0.08
D	+	0.50	0.27	0.18	0.06	0.51	0.25	0.16	0.08
R	+	0.75	0.02	0.23	0.01	0.75	0.02	0.23	0.00
H	-	0.11	0.15	0.49	0.25	0.16	0.11	0.44	0.29
D	-	0.12	0.12	0.60	0.16	0.17	0.07	0.55	0.21
R	-	0.25	0.01	0.67	0.07	0.24	0.02	0.68	0.06
H	$r$	0.43	0.44	0.10	0.03	0.46	0.41	0.07	0.06
D	$r$	0.54	0.32	0.09	0.05	0.54	0.32	0.09	0.05
R	$r$	0.86	0.01	0.11	0.02	0.84	0.03	0.13	0.00
H	$l$	0.16	0.06	0.49	0.29	0.14	0.08	0.51	0.27
D	$l$	0.13	0.09	0.62	0.15	0.17	0.05	0.59	0.19
R	$l$	0.20	0.01	0.73	0.05	0.20	0.01	0.73	0.05
H	0	0.26	0.28	0.26	0.19	0.29	0.26	0.24	0.22
D	0	0.40	0.14	0.41	0.05	0.44	0.10	0.37	0.09
R	0	0.48	0.04	0.42	0.06	0.47	0.05	0.44	0.05
H	1	0.32	0.23	0.29	0.15	0.34	0.22	0.27	0.17
D	1	0.32	0.24	0.32	0.12	0.36	0.20	0.28	0.16
R	1	0.56	0.00	0.41	0.03	0.54	0.02	0.43	0.01