Experimental demonstration of input-output indefiniteness in a single quantum device

Yu Guo,1,2,3 Zixuan Liu,4,5 Hao Tang,1,2,3 Xiao-Min Hu,1,2,3 Bi-Heng Liu,1,2,3 Yun-Feng Huang,1,2,3 Chuan-Feng Li,1,2,3 Guang-Can Guo,1,2,3 and Giulio Chiribella4,6,5,7,

1CAS Key Laboratory of Quantum Information, University of Science and Technology of China, Hefei, 230026, China
2CAS Center For Excellence in Quantum Information and Quantum Physics, University of Science and Technology of China, Hefei, 230026, China
3Hefei National Laboratory, University of Science and Technology of China, Hefei, 2300088, China
4QICI Quantum Information and Computation Initiative, Department of Computer Science, The University of Hong Kong, Pokfulam Road, Hong Kong
5HKU-Oxford Joint Laboratory for Quantum Information and Computation
6Department of Computer Science, University of Oxford, Wolfson Building, Parks Road, Oxford, UK
7Perimeter Institute for Theoretical Physics, 31 Caroline Street North, Waterloo, Ontario, Canada

Quantum theory allows information to flow through a single device in a coherent superposition of two opposite directions, resulting into situations where the input-output direction is indefinite. Here we introduce a theoretical method to witness input-output indefiniteness in a single quantum device, and we experimentally demonstrate it by constructing a photonic setup that exhibits input-output indefiniteness with a statistical significance exceeding 69 standard deviations. Our results provide a way to characterize input-output indefiniteness as a resource for quantum information and photonic quantum technologies and enable table-top simulations of hypothetical scenarios exhibiting quantum indefiniteness in the direction of time.

Introduction.— A cornerstone of quantum theory is the CPT theorem [1, 2], stating that the fundamental dynamics of quantum fields is invariant under inversion of time direction, charge, and parity. The theorem implies that, at the fundamental level, the roles of past and future are symmetric: while we normally treat systems at earlier times as the inputs and systems at later times as the outputs, the dynamical laws of quantum mechanics are indifferent to the direction of time. The time symmetry of the fundamental quantum dynamics was later extended to scenarios involving measurements by Aharonov and collaborators [3-5]. With the advent of quantum information, the role of time symmetry in quantum theory has attracted renewed attention, due to its connection with the structure of quantum protocols [6], multitime quantum states [7, 8], simulation of closed timelike curves [9, 10], inversion of unknown quantum evolutions [11-13], quantum retrodiction [14, 15], and the origin of irreversibility [16, 17]. Time-symmetric frameworks for quantum theory have been developed and analyzed.

Recently, Refs. [24, 25] extended the notion of time-reversal to a broader notion of input-output inversion, which applies whenever the roles of the input and output ports of a quantum device can be exchanged. This includes, for example, the case of linear optical devices, which can be traversed in two opposite spatial directions. Notably, all kinds of input-output inversions turned out to share the same mathematical structure. As a consequence, hypothetical scenarios involving the reversal of the time direction between two spacetime events can be simulated by real-world setups that reverse the direction of a path between two points in space. Building on the notion of input-output inversion, Ref. [24] then introduced a new type of operations that utilize quantum devices in a coherent superposition of two alternative input-output directions, giving rise to a feature called input-output indefiniteness. This feature has been found to offer advantages in information-theoretic [24, 25] and thermodynamical tasks [26, 27]. Input-output indefiniteness is also related to the notion of indefinite order [28, 29], whose applications to quantum information have been extensively investigated in the past decade, both theoretically [31-38] and experimentally [39-47]. An important difference is that, while indefinite order requires multiple devices (or multiple uses of the same device), input-output indefiniteness can already arise at the single-device level, enabling quantum protocols that could not be achieved with indefinite order (see Appendices [48] for examples in the tasks of gate transformation, estimation, and testing).

Here we develop a general method for witnessing input-output indefiniteness in the laboratory, and we use it to experimentally demonstrate a photonic setup that probes a single quantum device in a coherent superposition of two alternative directions. By optimizing the choice of witness, we demonstrate incompatibility of our setup with a definite input-output direction by more than 89 statistical deviations. Notably, our setup applies not only to reversible quantum devices, such as polarization rotators, but also to a class of irreversible devices including postselected polarization measurements. In addition to single-device indefiniteness, we experimentally demonstrate the combination of two devices in a quantum superposition of two opposite input-output directions, building a setup that achieves 99.6% winning probability in a quantum game where every strategy using both devices in the same direction fails with at least 11% probability. Our techniques enable a rigorous characterization of input-output indefiniteness as a resource for quantum information and photonic quantum technologies, and, at the same time, could be used to simulate exotic physical
models where the arrow of time is subject to quantum indefiniteness.

Witnesses of input-output indefiniteness. — For many processes in nature, the role of the input and output ports can be exchanged. An example is the transmission of a single photon through an optical crystal, schematically illustrated in Fig. 1. Quantum devices with exchangeable input-output ports, called bidirectional, can be used in two alternative ways, conventionally referred to as the “forward mode” (with the inputs entering at port A and the outputs exiting from port B) and “backward mode” (with the inputs entering at port B and the outputs exiting from port A). In the special case where ports A and B are associated with two moments of time \( t_A < t_B \), the forward mode corresponds to the standard use of the device in the forward time direction, while the backward mode corresponds to a hypothetical use of the device in the reverse time direction \([24]\).

Ref. \([24]\) showed that a device is bidirectional if and only if the corresponding transformation of density matrices is a bistochastic quantum channel \([49, 50]\), that is, a linear map \( \mathcal{C} \) of the form \( \mathcal{C}(\rho) = \sum_i C_i \rho C_i^\dagger \), where \( \rho \) is the input density matrix, and \( \{C_i\} \) are square matrices satisfying the conditions \( \sum_i C_i^\dagger C_i = \sum_i C_i C_i^\dagger = I \), \( I \) being the identity matrix. If a bistochastic channel \( \mathcal{C} \) describes the state change in the forward mode, then the state change in the backward mode is described by a (generally different) bistochastic channel \( \mathcal{\Theta}(\mathcal{C}) \) given by \( \Theta(\mathcal{C}) : \rho \mapsto \sum_i \theta(C_i) \rho \theta(C_i)^\dagger \), where the square matrix \( \theta(C_i) \) is either unitarily equivalent to \( C_i^\dagger \), the transpose of \( C_i \), or unitarily equivalent to \( C_i^\dagger \), the adjoint of \( C_i \) \([24]\).

The map \( \mathcal{\Theta} \) is called an input-output inversion. Physically, it can represent a time reversal (if the two ports of the device correspond to two moments of time), an inversion of spatial directions (as in the example of the optical crystal), or any other symmetry transformation obeying a set of general axioms specified in \([24]\). In the following, we will focus on the case where the input-output inversion is (unitarily equivalent to) the transpose. This case includes in particular the canonical time-reversal in quantum mechanics \([51, 52]\) and quantum thermodynamics \([18, 53]\) (see \([18]\) for more details.)

In principle, quantum mechanics allows for setups that coherently control the input-output direction, such as the setup shown in Fig. 1(d). We now develop a method for witnessing input-output indefiniteness in the laboratory. A witness for a given quantum resource, such as entanglement \([54]\), indefinite causal order \([55]\), and causal connection \([56]\), is an observable quantity that distinguishes between resourceful and non-resourceful setups \([57]\). In our case, the non-resourceful setups are those that use the device in a well-defined direction. Setups that use it in the forward (backward) mode are described by a suitable set of positive operators, denoted by \( S_{\text{fwd}} (S_{\text{bwd}}) \). The explicit characterization of these operators is provided in the Appendices \([18]\). For the following discussion, it will suffice to know that they act on the tensor product Hilbert space \( \mathcal{H}_A \otimes \mathcal{H}_{A_0} \otimes \mathcal{H}_{B_0} \otimes \mathcal{H}_{B_0} \), where \( \mathcal{H}_A \) (\( \mathcal{H}_{A_0} \)) is the Hilbert space of the input (output) system of the device, while \( \mathcal{H}_{B_0} \) (\( \mathcal{H}_{B_0} \)) is the Hilbert space of the input (output) system of the overall process obtained by inserting the device into the setup.

A setup that uses the device in a random mixture of the forward and backward modes corresponds to an operator of the form

\[
S = p S_{\text{fwd}} + (1 - p) S_{\text{bwd}} ,
\]

with \( S_{\text{fwd}} \in S_{\text{fwd}}, S_{\text{bwd}} \in S_{\text{bwd}}, \) and \( p \in [0, 1] \). We will denote by \( S_{\text{definite}} \) the set of all operators of the form \( S \). The setups outside \( S_{\text{definite}} \) are incompatible with the use of the given device in a definite input-output direction: in these setups, the device is not used in the forward mode, nor in the backward mode, nor in any random mixture thereof. For an operator \( S \) outside \( S_{\text{definite}} \), we define a witness of input-output indefiniteness to be a self-adjoint operator \( W \) such that

\[
\text{Tr}(WS) < 0 ,
\]

and

\[
\text{Tr}(WS') \geq 0 , \quad \forall S' \in S_{\text{definite}} .
\]

The condition \([3]\) is characterized in the following Theorem, which provides a systematic way to construct witnesses of input-output indefiniteness.

**Theorem 1** A Hermitian operator \( W \) satisfies Eq. \([3]\) if and only if there exist operators \( W_0 \) and \( W_1 \) such
that $W \geq W_0$, $W \geq W_1$, $W_0 = [\lambda_0]_0 = [\lambda_0]_0 + [\lambda_0]_0 W_0 + [\lambda_0]_0 W_0 - [\lambda_0]_0 W_0$, and $W_1 = [\lambda_0]_0 W_1 - [\lambda_0]_0 W_1 + [\lambda_0]_0 W_1 - [\lambda_0]_0 W_1 W_1$, having used the notation $|X\rangle := \text{Tr}_X [S] \otimes \frac{1}{d_X}$ for a system $X$ of dimension $d_X$.

The proof is provided in the Appendices 48, where we also show that the expectation value of any witness can be decomposed into a linear combination of outcome probabilities arising from settings in which a device is inserted in the setup and the resulting process is probed on multiple input states.

Experimental demonstration of input-output indefiniteness of a single quantum device. — Our experimental setup, illustrated in Fig. 2, is inspired by a theoretical primitive known as the quantum time flip (QTF) 21. The QTF takes in an input an arbitrary bidirectional device and adds quantum control to the direction in which the device is used. When applied to a bidirectional device that acts as channel $C$ in the forward direction, the QTF generates a new quantum channel $F(C)$, acting jointly on the target system and on a control qubit. Explicitly, the Kraus operators of the new quantum channel $F(C)$, denoted by $\{F_i\}$, are related to the Kraus operators of the original channel $C$, denoted by $\{C_i\}$, as

$$F_i = C_i \otimes |0\rangle\langle 0| + C_i^T \otimes |1\rangle\langle 1|, \quad (4)$$

where $\{|0\rangle, |1\rangle\}$ are two orthogonal states of the control qubit. When the control qubit is initialized in a coherent superposition of $|0\rangle$ and $|1\rangle$, the new channel $F(C)$ implements a superposition of channel $C$ and its input-output inversion $C^T$, in the sense of Refs. 58–65.

In our experiment, schematically illustrated in Fig. 1(d), a heralded single photon is generated through spontaneous parametric down-conversion 48. The polarization qubit, serving as the target system in the QTF, is initialized in an arbitrary fixed state, using a fiber polarizer controller, a half-wave plate (HWP) and a quarter-wave plate (QWP). The photon is sent to a 50/50 beamsplitter (BS1) to prepare the spatial qubit in the superposition state $|+\rangle = (|0\rangle + |1\rangle) / \sqrt{2}$, where $|0\rangle$ and $|1\rangle$ correspond to the two alternative paths shown green and carmine in Fig. 2. The input device for the QTF is a bistochastic measure-and-reprepare operation 21, implemented by an assemblage of two HWPs, two QWPs, and a polarizing beam splitter (PBS), shown inside the dotted rectangle in Fig. 2. The input-output inversion is realized by routing the photon through the same assemblage along a backward path sandwiched between two fixed Pauli gates $Z = |0\rangle\langle 0| - |1\rangle\langle 1|$. A coherent superposition of the forward and backward measure-and-reprepare operations is created by using the spatial qubit as a control qubit. Finally, two paths are coherently recombined on BS2, followed by a measurement on the polarization qubit.

To certify input-output indefiniteness, we derived the witness $W^{\text{obs}}$ with maximum robustness to noise (see Appendices 48). This witness can be estimated by probing the setup on a set of bistochastic measure-and-reprepare processes that measure the polarization qubit in the eigenbasis of a Pauli gate and reprepare the output.
in a state in the eigenbasis of another Pauli gate. The overall evolution induced by the setup is probed by initializing the path qubit in the maximally coherent state $|+\rangle$ and the polarization qubit in one of the states $|0\rangle$, $|1\rangle$, $|+\rangle$ and $\frac{1}{\sqrt{2}}(|0\rangle+i|1\rangle)$. Finally, the target qubit and control qubit are measured in the eigenbases of the three Pauli gates. The measured probabilities, shown in Fig. 3 are used to calculate the experimental value of the witness $\text{Tr}[W_{\text{opt}}S_{\text{QTF}}]$, which we find to be $-(0.345\pm0.005)$, corresponding to a violation of the condition of definite input-output direction by more than 69 standard deviations.

To implement the optimal witness $W_{\text{opt}}$, we performed local operations on 5 qubits, using a total of 794 settings. The complexity of the experiment is less than that of a full process tomography, which would require at least 1023 settings. To further reduce the complexity, we designed a simplified witness where the target qubit is initialized in a fixed state $|0\rangle$ and is eventually discarded. This witness involves only 3 qubits and 48 settings, which we show to be the optimal values. In the experiment, we find the value $-(0.140\pm0.004)$, which certifies incompatibility with a definite input-output direction by more than 35 standard deviations.

**Experimental demonstration of advantage in a quantum game.**—Input-output indefiniteness offers an advantage in a quantum game where a referee challenges a player to find out a hidden relation between two unknown quantum gates [24]. In this game, the referee provides the player with two devices implementing unitary gates $U$ and $V$, respectively, promising that the two gates satisfy either the relation $UV^T = U^TV$ or the relation $UVT = -U^T V$. The player’s task is to determine which of these two alternatives holds. Ref. [24] showed that a player that uses the two gates in the QTF can win the game with certainty, while every strategy that uses the two devices in the same input-output direction will fail at least 11% of the times.

In our experiment, discussed in the Appendices [48], we observe an average success probability of $99.60 \pm 0.18\%$ over a set of 21 gate pairs. The worst-case error probability is approximately $0.68 \pm 0.19\%$, which is 16 times smaller than 11%, the lower bound on the error probability for all possible strategies with definite-input-output direction. We also show that the advantage of input-output indefiniteness persists even if the player has coherent control on each of the gates $U$ and $V$: every strategy using the controlled gates $\text{ctrl} - U = I\otimes|0\rangle\langle0| + U\otimes|1\rangle\langle1|$ and $\text{ctrl} - V = I\otimes|0\rangle\langle0| + V\otimes|1\rangle\langle1|$ in the same input-output direction will necessarily have an error probability of at least 5.6%. Overall, this game can be regarded as a bipartite witness of global input-output indefiniteness. In the Appendices [48], we provide a general theory of such witnesses.

**Conclusions.**—In this paper we introduced the notion of witness of input-output indefiniteness and used it to experimentally demonstrate input-output indefiniteness in a single photonic device. Our results provide a way to rigorously characterize input-output indefiniteness in the laboratory, and represent a counterpart to recent experiments on indefinite order of quantum gates [39–47]. Overall, input-output indefiniteness provides a new resource for quantum information protocols, and could potentially lead to advantages in photonic quantum technologies. Our setup and its generalizations could also be used to simulate exotic physics in which the arrow of time is a quantum variable. These hypothetical phenomena fit into a broad framework developed by Hardy [28], who suggested that a full-fledged theory of quantum gravity would require spacetime structures to be subject to quan-
A. J. Parzygnat and J. Fullwood, PRX Quantum 4 (2023).

S. Yoshida, A. Soeda, and M. Murao, Physical Review 7 (2023).

A. J. Parzygnat and F. Buscemi, Quantum 7, 1013 (2023).

L. Maccone, Physical Review Letters 103, 080401 (2009).

Y. Aharonov, P. G. Bergmann, and J. L. Lebowitz, Physical Review 134, B1410 (1964).

D. Genkina, G. Chiribella, and L. Hardy, Physical Review Letters 125, 043017 (2021).

O. Oreshkov and N. J. Cerf, Nature Physics 11, 452 (2015).

A. Di Biagio, P. Donà, and C. Rovelli, Quantum 5, 520 (2021).

G. Chiribella, E. Aurell, and K. Życzkowski, Physical Review Research 3, 033028 (2021).

L. Hardy, arXiv preprint arXiv:2104.00071 (2021).

J. H. Selby, M. E. Stasinou, S. Gogioso, and B. Coecke, arXiv preprint arXiv:2209.07867 (2022).

G. Chiribella and Z. Liu, Communications Physics 5, 1 (2022).

G. Chiribella, G. D’Ariano, P. Perinotti, and B. Valiron, arXiv preprint arXiv:0912.0195 (2009).

G. Chiribella, G. M. D’Ariano, P. Perinotti, and B. Valiron, Physical Review Letters 113, 040403 (2019).

G. Chiribella, G. M. D’Ariano, P. Perinotti, and B. Valiron, Physical Review A 88, 022318 (2013).

G. Chiribella, Physical Review A 85, 022330 (2012).

M. T. Quintino, Q. Dong, A. Shimbo, A. Soeda, and M. Murao, Physical review letters 123, 210502 (2019).

D. Ebler, S. Salek, and G. Chiribella, Physical Review Letters 120, 120502 (2018).

Y. Shikano, S. Pirandola, L. A. Rozema, A. Darabi, Y. Soudagar, L. K. Shalm, et al., Physical Review Letters 106, 040403 (2011).

D. Genkina, G. Chiribella, and L. Hardy, Physical Review A 85, 022330 (2012).

M. T. Quintino, Q. Dong, A. Shimbo, A. Soeda, and M. Murao, Physical review letters 123, 210502 (2019).

D. Trillo, B. Dive, and M. Navascués, Physical Review Letters 130, 110203 (2023).

S. Yoshida, A. Soeda, and M. Murao, Physical Review Letters 131, 120602 (2023).

M. T. Quintino, Q. Dong, A. Shimbo, A. Soeda, and M. Murao, Physical review letters 123, 210502 (2019).

X. Zhao, Y. Yang, and G. Chiribella, Physical Review Letters 124, 070503 (2020).

D. Felce and V. Vedral, Physical Review Letters 125, 070503 (2020).

Y. Aharonov and L. Vaidman, in Time in quantum mechanics (Springer, 2002) pp. 369–412.

M. T. Quintino, Q. Dong, A. Shimbo, A. Soeda, and M. Murao, Physical Review Letters 123, 210502 (2019).

S. Yoshida, A. Soeda, and M. Murao, Physical Review Letters 131, 120602 (2023).

G. Chiribella, M. Wilson, and H. Chau, Physical Review Letters 127, 190502 (2021).

M. Procopio, A. Moqanaki, M. Araújo, and A. Di Biagio, Physical Review Research 4, 033028 (2021).
Appendix A: Relation with indefinite order

Here we discuss the relation between indefinite input-output direction and indefinite order, highlighting similarities and differences, and presenting a list of quantum
protocols that can be achieved using indefinite input-output direction but cannot be achieved using indefinite order.

Indefinite order refers to scenarios where two or more quantum devices are connected with one another in a way that is not compatible with any probabilistic mixture of well-defined orders [30]. The simplest example of this situation is the quantum SWITCH [29, 31], an operation \( \mathcal{S} \) that takes in two quantum devices \( \mathcal{A} \) and \( \mathcal{B} \), acting on the same target system, and produces in output a new quantum device \( \mathcal{S}(\mathcal{A}, \mathcal{B}) \) that adds quantum control to the order in which \( \mathcal{A} \) and \( \mathcal{B} \) are applied. Mathematically, the new device \( \mathcal{S}(\mathcal{A}, \mathcal{B}) \) is a bipartite quantum channel, acting on the target system and on a control qubit that determines the relative order of \( \mathcal{A} \) and \( \mathcal{B} \). The Kraus operators of channel \( \mathcal{S}(\mathcal{A}, \mathcal{B}) \), denoted by \( \{S_{ij}\} \) are given by

\[
S_{ij} = A_i B_j \otimes |0\rangle\langle 0| + B_j A_i \otimes |1\rangle\langle 1| , \quad (A1)
\]

where \( \{A_i\} \) and \( \{B_j\} \) are the Kraus operators of channels \( \mathcal{A} \) and \( \mathcal{B} \), respectively, and \( \{|0\rangle, |1\rangle\} \) are orthogonal states of the control qubit. When the control qubit is initialized in the state \( |0\rangle \) (|1\rangle), the target qubit undergoes processes \( \mathcal{A} \) and \( \mathcal{B} \) in the definite order \( AB \) (BA). When the control qubit is in a coherent quantum superposition of \( |0\rangle \) and \( |1\rangle \), instead, the order of \( \mathcal{A} \) and \( \mathcal{B} \) becomes indefinite. A number of information-theoretic applications of the quantum switch has been discussed over the past years [31, 68], and series of experiments inspired by the quantum switch has been performed on photonic systems [39–47] (see also Refs. [66–68] for a theoretical discussion on the interpretation of the experiments.)

Another famous example of indefinite order is a process introduced by Oreshkov, Costa, and Brukner [30]. More generally, the quantum switch, the Oreshkov-Costa-Brukner process, and other operations with indefinite causal order are represented by quantum supermaps [31, 69], that is, higher-order maps acting on quantum channels. This type of supermaps take in input two (or more) quantum channels and produce a new quantum channel as output.

For indefinite order, the input channels can be arbitrary completely positive, trace-preserving maps. Physically, this condition guarantees that the corresponding supermaps could in principle be implemented on arbitrary quantum devices. In stark contrast, operations with indefinite input-output direction can only be applied to bidirectional devices, mathematically described by bistochastic channels. In other words, operations with indefinite input-output direction are defined on a strictly smaller domain. The restriction of the domain leads to a broader set of conceivable supermaps, which can sometime lead to stronger advantages in quantum information tasks. In the following, we will present three such advantages:

1. **Unitary black box inversion/transposition.** In this task, one is given a black box implementing an unknown quantum dynamics, represented by a unitary gate \( \mathcal{U} \) acting on a \( d \)-dimensional quantum system. The goal is to produce a new black box implementing the inverse of the original dynamics, corresponding to the unitary \( \mathcal{U}^\dagger \), or the transpose of the original dynamics, corresponding to the gate \( \mathcal{U}^T \).

For qubits, the inverse and the transpose are unitarily equivalent, due to the relation \( \mathcal{U}^\dagger = \mathcal{V} \mathcal{U} \mathcal{V}^\dagger \), valid for every matrix \( \mathcal{U} \in \text{SU}(2) \). For higher dimensional systems, we will focus our attention on the transpose. If the gate \( \mathcal{U} \) is an arbitrary unitary matrix, the transformation \( \mathcal{U} \mapsto \mathcal{U}^T \) cannot be perfectly achieved by inserting the gate \( \mathcal{U} \) in a quantum circuit: every implementation of this transformation must necessarily be probabilistic, or approximate [13, 70, 71]. Furthermore, Ref. [24] showed that the transpose \( \mathcal{U}^T \) cannot be perfectly generated from two copies of the original gate \( \mathcal{U} \), even if the two copies are used in an indefinite order.

The above no-go theorems do not apply if the experimenter is able to use the black box in two opposite input-output directions, that is, if the black box can be treated by the experimenter as a bidirectional quantum device, acting as \( \mathcal{U} \) in one direction, and acting as \( \mathcal{U} \mathcal{V} \mathcal{U}^\dagger \) in the opposite direction, where \( \mathcal{V} \) is a fixed unitary. In this case, the experimenter has only to use the device in the appropriate direction and to undo the unitary \( \mathcal{V} \).

In summary, the transformation \( \mathcal{U} \mapsto \mathcal{U}^T \) provides an example of a quantum task that cannot be perfectly achieved by operations with indefinite order (using two queries of the unitary gate \( \mathcal{U} \)), but can be perfectly achieved by operations with indefinite input-output direction.

2. **Gate estimation.** In this task an experimenter is given access to a black box implementing an unitary gate of the form \( \mathcal{U}_\theta = e^{-i\theta \mathcal{H}} \), where the generator \( \mathcal{H} \) has eigenvalues \( \{0, 1, \ldots, d - 1\} \) and satisfies the condition \( \mathcal{H}^T = -\mathcal{H} \), while the shift parameter \( \theta \) is in the range \( [0, 2\pi] \). The goal is to estimate the \( \theta \) with minimum error, that is, to produce an estimate \( \hat{\theta} \) that minimizes the root mean square error (RMSE) \( \Delta \theta := \sqrt{\frac{1}{N} \sum_{n=1}^{N} (\hat{\theta}_n - \theta)^2} \), where \( p(\hat{\theta} | \theta) \) is the conditional probability of obtaining \( \hat{\theta} \) when the true value is \( \theta \).

Suppose that the \( N \) copies of the gate \( \mathcal{U}_\theta \) are available to the experimenter. When the \( N \) gates are used in parallel, they act as a single \( N \)-partite gate, and the total generator has spectrum \( \{0, \ldots, N(d - 1)\} \). If \( \theta \) is then known that the minimum RMSE scales as \( \Delta \theta \approx \pi/\sqrt{2Nd} \) at the leading order in \( Nd \), both in the worst case over \( \theta \) and on average over uniformly distributed \( \theta \) [72, 73]. This value of
the RMSE is optimal even if the $N$ gates are used in a sequence [76], or, more generally, in an indefinite order (by a simple generalization of the argument in [76]).

In contrast, the QTF can transform each gate $U_g$ into the gate $W_g = U_g \otimes |0\rangle \langle 0| + U_g^T \otimes |1\rangle \langle 1|$, whose generator $K := H \otimes Z$ has spectrum $\{-d+1, \ldots, d-1\}$.

By applying the QTF to the initial $N$ gates, the experimenter can obtain $N$ copies of the gate $W_0$, which can then be used to estimate $\theta$ with RMSE $\pi/[2 \sqrt{2}N \Delta]$ at the leading order in $N \Delta$. In other words, indefinite input-output direction can reduce the RMSE by a factor 2, equivalent to doubling the number of available uses of the unknown gate $U_g$.

**Appendix B: Relation with the notion of time reversal in quantum theory and quantum thermodynamics**

Here we briefly summarize the relation between the notion of input-output inversion and the notion of time-reversal in quantum mechanics [51, 52] and in quantum thermodynamics [53].

Ref. [24] showed that the input-output inversion of a unitary channel, corresponding to a unitary matrix $U$, is another unitary channel, corresponding to another unitary matrix $\theta(U)$. The map $\theta : U \mapsto \theta(U)$ is either unitarily equivalent to the transpose $\theta(U) = U V U^T V^T$ or the relation $U V U^T = -U V V^T$ is satisfied. A player can query each black box one time, and then has to determine which of the two alternative relations holds.

Ref. [24] showed every strategy that uses the two devices in the same input-output direction will necessarily have a nonzero probability of error. This conclusion applies even if the two black boxes are used in an indefinite order; if the input-output direction is fixed and equal for both boxes, then the error is non-zero. The nature of this advantage will be discussed in detail in Section 3 where we classify different types of witnesses of input-output indefiniteness for pairs of bistochastic channels.

**Testing properties of quantum gates.** Another advantage of input-output indefiniteness arises in the game described in the main text, originally introduced in Ref. [24]. In this game, a referee prepares a pair of black boxes implementing unitary gates $U$ and $V$, respectively, and guarantees that either the relation $U V U^T = U V$ or the relation $U V U^T = -U V V^T$ is satisfied. A player can query each black box one time, and then has to determine which of the two alternative relations holds.

Ref. [24] showed every strategy that uses the two devices in the same input-output direction will necessarily have a nonzero probability of error. This conclusion applies even if the two black boxes are used in an indefinite order; if the input-output direction is fixed and equal for both boxes, then the error is non-zero. The nature of this advantage will be discussed in detail in Section 3 where we classify different types of witnesses of input-output indefiniteness for pairs of bistochastic channels.

The choice of a time reversal symmetry at the state level induces a notion of time reversal of unitary dynamics. Suppose that the time-reversal symmetry is described by an operator $A$ (either unitary or anti-unitary).

In the non-unitary case, notions of time-reversal have been proposed in the literature. An early formulation is due to Crooks [83], who defined the time reversal of a quantum channel $\mathcal{C}$ as the Petz’ recovery map $\mathcal{C}_{\text{Petz}}$ explicitly given by $\mathcal{C}_{\text{Petz}}(\rho) := \rho_0^{1/2} \mathcal{C}^\dagger (\rho_0^{-1/2} \rho_0^{-1/2} \rho_0^{1/2})$,
where $\rho_0$ is any quantum state such that $C(\rho_0) = \rho_0$, and $C^\dagger$ is the adjoint of channel $C$. If one restricts the time-reversal to bistochastic channels and one picks the state $\rho_0$ to be maximally mixed, then Crooks’ definition coincides with input-output inversion discussed in the main text, in the special case where the input-output inversion is unitarily equivalent to the adjoint.

An extension of Crooks’ approach was recently proposed by Chiribella, Aurell, and Życzkowski [55]. In this extension, one defines a fixed reference state for every system, and defines the time-reversal on the subset of channels $C$ satisfying the condition $C(\rho_S) = \rho_S$, where $\rho_S$ and $\rho_{S2}$ are the fixed reference states of the systems $S$ and $S_2$ corresponding to the input and output of channel $C$, respectively. On this subset of channels, the time-reversal is defined as the Petz recovery map $C_{\text{Petz}}(\rho) := \rho_S^{1/2}C^\dagger(\rho_S^{-1/2}\rho\rho_S^{-1/2})\rho_S^{1/2}$, or as the variant of the Petz recovery map where the adjoint $C^\dagger$ is replaced by the transpose $C^T$ (explicitly, $C_{\text{Petz}}(\rho) := \rho_S^{1/2}C^T(\rho_S^{-1/2}\rho\rho_S^{-1/2})\rho_S^{1/2}$, where $\overline{\rho}$ denotes the complex conjugate of the matrix $\rho$ in the given basis). When the reference states are set to be maximally mixed, this notion of time reversal mathematically coincides with the notion of input output inversion discussed in the main text.

### Appendix C: Characterization of the witnesses

In this section, we provide the characterization of the witnesses of input-output indefiniteness, which is done with the Choi representation [80]. The Choi representation associates linear maps $M : L(\mathcal{H}_{in}) \rightarrow L(\mathcal{H}_{out})$ to bipartite operators $\text{Choi}(M) := (\mathcal{I}_{in} \otimes M)(|I\rangle\langle I|)_{in,in}$, where $L(\mathcal{H})$ denotes the linear operators on a Hilbert space $\mathcal{H}$, and $|I\rangle_{in,in} = \sum_m |m\rangle_\mathcal{I}_{in}|m\rangle_\mathcal{I}_{in}$ is the (unnormalized) maximally entangled quantum state on $\mathcal{H}_{in} \otimes \mathcal{H}_{in}$. The Choi representation is related to the Jamiołkowski representation [57] $\text{Jam}(M) := \sum_m |m\rangle_\mathcal{I}_{in}|m\rangle_\mathcal{I}_{in} M(|m\rangle_\mathcal{I}_{in})$ by a partial transposition on the first Hilbert space.

In the Choi representation, a setup that uses the original device in the backward mode corresponds to a positive operator $S_{\text{fwd}}$ satisfying the conditions $\text{Tr}_{B_2}[S_{\text{fwd}}] = I_{A_2} \otimes \text{Tr}_{A_2B_2}[S_{\text{fwd}}]/d_A$ and $\text{Tr}_{A_2A_1B_2}[S_{\text{fwd}}]/d_A = I_{B_1}$ [88]. Similarly, a setup that uses the original device in the backward mode corresponds to a positive operator $S_{\text{fwd}}$ satisfying the conditions $\text{Tr}_{B_2}[S_{\text{fwd}}] = I_{A_2} \otimes \text{Tr}_{A_2B_2}[S_{\text{fwd}}]/d_A$ and $\text{Tr}_{A_2A_1B_2}[S_{\text{fwd}}]/d_A = I_{B_1}$. More generally, a setup that uses the device in a random mixture of the forward and backward mode corresponds to an operator of the form $S = p S_{\text{fwd}} + (1 - p) S_{\text{bwd}}$, which, as we have mentioned in the main text, is incompatible with the use of the given device in an indefinite input-output direction. The scenarios involving indefinite input-output direction can be described by quantum supermaps [24]. Every quantum supermap $S : \mathcal{M} \rightarrow \mathcal{S}(\mathcal{M})$ induces a linear map $\hat{S}$ on the Choi operators via the relation $\hat{S}(\text{Choi}(M)) = \text{Choi}(S(M))$. Hence, we can define the Choi operator of the supermap $S$ as the Choi operator of the induced map $\hat{S}$. In turn, the Choi operator of a supermap acting on a single bidirectional device as in Fig. 1 of the main text can be described by a positive operator $S$, acting on the tensor product Hilbert space $\mathcal{H}_{A_2} \otimes \mathcal{H}_{A_2} \otimes \mathcal{H}_{B_2} \otimes \mathcal{H}_{B_2}$, where $\mathcal{H}_{A_2}$ is the Hilbert space of the input (output) system of the initial channel $C$, whereas $\mathcal{H}_{B_2}$ is the Hilbert space of the input (output) system of the final channel $\hat{S}(C)$. Since the original device $C$ transforms a given quantum system into itself, the Hilbert spaces $\mathcal{H}_{A_2}$ and $\mathcal{H}_{A_0}$ have the same dimension, hereafter denoted by $d_A$.

In the following, the set of all witnesses of input-output indefiniteness will be denoted by $\mathcal{W}_{\text{in/out}}$, which will also be characterized in the Choi representation. With this notation, we are ready to provide our characterization of the set $\mathcal{W}_{\text{in/out}}$.

#### 1. Proof of Theorem 1 of the main text

Since all witnesses have non-negative expectation values on the setups with definite input-output direction, $\mathcal{W}_{\text{in/out}}$ is the dual cone of $\mathcal{S}_{\text{finite}}$. In turn, $\mathcal{S}_{\text{finite}}$ is the convex hull of the (Choi operators of) supermaps with forward input-output direction and of those with backward input-output direction. The Choi operators of supermaps with forward input-output direction generate a closed convex cone of operators on $\mathcal{H}_{A_2} \otimes \mathcal{H}_{A_2} \otimes \mathcal{H}_{B_2} \otimes \mathcal{H}_{B_2}$. Specifically, the closed convex cone is equal to $\mathcal{P} \cap \mathcal{L}_f$, where $\mathcal{P}$ is the cone of positive operators, $\mathcal{L}_f$ is the subspace defined by the intersection of two subspaces.

$$\mathcal{L}_f := \left\{ X \mid \text{Tr}_{B_2}(X) = \text{Tr}_{A_2B_2}(X) \otimes \frac{I_{A_2}}{d_A} \right\} \quad \text{(C1)}$$

$$\cap \left\{ X \mid \text{Tr}_{A_1A_2B_2}(X) = \frac{\text{Tr}(X)}{d_{B_2}} \cdot I_{B_2} \right\}. \quad \text{(C2)}$$

Notice that the projections onto the two subspaces in Eq. (C1) commute. It follows that $\mathcal{L}_f$ can be expressed, with the notation of Theorem 1 in the main text, as below,

$$\mathcal{L}_f = \left\{ X \mid [A_{B_2}]X - [A_{B_2}]X + [A_{A_2A_2B_2}]X = 0 \right\}. \quad \text{(C3)}$$

$$= [A_{A_2A_2B_2}]X = 0. \quad \text{(C4)}$$

Similarly, the Choi operators of supermaps with backward input-output direction generate the closed convex cone $\mathcal{P} \cap \mathcal{L}_b$, with $\mathcal{L}_b$ given by

$$\mathcal{L}_b = \left\{ Y \mid [A_{B_2}]Y - [A_{B_2}]Y + [A_{A_2A_2B_2}]Y = 0 \right\}. \quad \text{(C5)}$$

$$- [A_{A_2A_2B_2}]Y = 0. \quad \text{(C6)}$$

It follows that the conic hull of $\mathcal{S}_{\text{finite}}$ is equal to

$$\text{con}(\mathcal{S}_{\text{finite}}) = \text{conv}(\mathcal{P} \cap \mathcal{L}_f) \cup (\mathcal{P} \cap \mathcal{L}_b). \quad \text{(C7)}$$
The dual cone of $S_{\text{definite}}$ can be deduced using the duality properties of closed convex cones:

\[
\begin{align*}
\con(S_{\text{definite}})^* &= (P \cap L_f)^* \cap (P \cap L_o)^* , \\
&= \text{conv}(P \cup L_f^+ ) \cap \text{conv}(P \cup L_o^+ ) , \\
&= (P + L_f^+ ) \cap (P + L_o^+ ) ,
\end{align*}
\]

where $\perp$ denotes orthogonal complement and $+$ denotes Minkowski addition.

Now we can conclude from Eq. (C8) that a Hermitian operator $W$ on $H_A \otimes H_A \otimes H_B \otimes H_B$ is a witness of input-output indefiniteness if and only if

\[
W \geq W_0 , \quad W \geq W_1 ,
\]

for some $W_0 \in L_f^+$ and $W_1 \in L_o^+$. According to the characterization of $L_f$ and $L_o$ in Eq. (C3) and (C5), respectively, $W_0$ satisfies the condition

\[
W_0 = [B_0]W_0 - [A_0]B_0W_0 + [A_1]A_0B_0W_0 - [A_1]A_0B_1B_0W_0 ,
\]

and $W_1$ satisfies the condition

\[
W_1 = [B_0]W_1 - [A_1]B_0W_1 + [A_1]A_0B_0W_1 - [A_1]A_0B_1B_0W_1 .
\]

2. Measures of input-output indefiniteness

Witnesses can be used not only to detect resources, but also to define quantitative resource measures. This is done by assessing the robustness of a given resource to the addition of noise, as it was done, e.g. for the robustness of entanglement [59], robustness of indefinite causal order [55], and robustness of causal connection [56].

Setups that use bidirectional channels are mathematically described by quantum supermaps that transform bistochastic channels into (generally non-bistochastic) channels [24]. The Choi operator of every such supermap, denoted by $S$, satisfies the conditions $\text{Tr}_{A_0A_0B_0}[S]/d_A = I_{B_1}$ and

\[
\text{Tr}_{B_0}[S] = \text{Tr}_{B_0A_0}[S] \otimes \frac{I_A}{d_A} + \text{Tr}_{B_0A_1}[S] \otimes \frac{I_A}{d_A} - \text{Tr}_{B_0A_0A_0}[S] \otimes \frac{I_A}{d_A} \otimes \frac{I_A}{d_A} .
\]

In the following, the set of all Choi operators satisfying the above constraints will be denoted by $S$.

We define the robustness of a setup $S \in S$ with respect to a witness $W$ as

\[
r(S \mid W) := \max_{T \in S} \left\{ \lambda \geq 0 \mid \text{Tr} \left( W \cdot \frac{S + \lambda T}{1 + \lambda} \right) \geq 0 \right\} ,
\]

which is equal to the amount of noise the setup $S$ can tolerate until its input-output indefiniteness stops to be detected by the witness $W$. The definition implies that the input-output indefiniteness of the setup $S$ can be detected by the witness $W$ only if the quantity $r(S \mid W)$ is larger than zero. Optimizing the quantity $r(S \mid W)$ over the witness $W$, we obtain the robustness of input-output indefiniteness of the setup $S$

\[
r(S) := \max_{W \in \text{Win/out}} r(S \mid W) .
\]

Both Eq. (C13) and Eq. (C14) can be phrased as semidefinite programming (SDP) problems and can be computed efficiently (see Section D). In particular, the quantity $r(S)$ is given by the following SDP:

\[
\begin{align*}
\text{maximize} & \quad - \text{Tr}(WS) \\
\text{subject to} & \quad W \in S_{\text{definite}}^* \\
& \quad \frac{I}{d_A d_B} - W \in S^* ,
\end{align*}
\]

where $S_{\text{definite}}^*$ and $S^*$ are the dual cones of $S_{\text{definite}}$ and $S$, respectively. The last constraint of Eq. (C15) can be interpreted as a normalization condition, noticing that, for every setup $T$, the value $\text{Tr}(WT)$ should not exceed 1.

The form of Eq. (C15) implies that the robustness is a convex in its argument, is faithful measure of input-output indefiniteness ($r(S)$ is equal to zero if and only if $S$ is compatible with a well-defined input-output direction), and cannot be increased by composing the setup $S$ with local bistochastic channels.

Appendix D: Efficient computation of the robustness of input-output indefiniteness via SDP

1. Derivation of the SDP problems

The evaluation of the robustness of input-output indefiniteness can be cast into an SDP problem that can be solved efficiently, similarly to the SDP problem for the causal robustness in Ref. [55]. Explicitly, the SDP for the robustness of input-output indefiniteness is

\[
\begin{align*}
\text{minimize} & \quad \text{Tr}(T) \\
\text{subject to} & \quad \text{Tr}(W(S + T)) \geq 0 \\
& \quad T \in \text{coni}(S) ,
\end{align*}
\]

where coni($S$) denotes the conic hull of the set $S$, consisting of the Choi operators of all deterministic supermaps on bistochastic channels. The above expression follows from the definition of robustness in Eq. (C13).
The dual of the above SDP is

\[
\begin{align*}
\text{maximize} & \quad -y \text{Tr}(WS) \\
& \quad \frac{1}{d_B d_A} \subseteq yW \in \text{coni(S)}^* \\
y & \geq 0.
\end{align*}
\] (D2)

In the next subsection, we show that the primal-dual pair \([D1] \text{ and } [D2]\) satisfies the condition for strong duality, meaning that the solutions of the primal and dual problems coincide. Hence, the robustness \(r(S | W)\) can be computed through the dual problem \([D2]\).

Furthermore, the maximum robustness over all possible witnesses, denoted by \(r(S) := \max_S r(S | W)\), can also be computed by an SDP, which follows from Eq. \([D2]\) by absorbing the variable \(y\) into the variable \(W\), thus obtaining

\[
\begin{align*}
\text{maximize} & \quad -\text{Tr}(WS) \\
& \quad W \in \text{coni(S)}^* \\
& \quad \frac{1}{d_B d_A} \subseteq W \in \text{coni(S)}^*.
\end{align*}
\] (D3)

The problem \([D3]\) is exactly the SDP problem in Eq. \([C15]\). Its dual problem is

\[
\begin{align*}
\text{minimize} & \quad \text{Tr}(T) \\
& \quad \frac{1}{d_B d_A} \\
\text{subject to} & \quad S + T \in \text{coni(S)}^* \\
& \quad T \in \text{coni(S)}.
\end{align*}
\] (D4)

2. Proof of strong duality and efficient solvability

Here we prove that the SDP problems derived in the previous subsection can be solved efficiently and satisfy the condition for strong duality. To obtain this result, we will use some general facts about conic optimization problems, a large category of optimization problems that include SDP as a special case.

Mathematically, a conic optimization problem is defined as follows:

**Definition 1** Let \(E\) be a finite-dimensional vector space, \(K\) a closed convex pointed cone in \(E\) with a nonempty interior, and \(L\) a linear subspace of \(E\). Let also \(b \in E\) and \(c \in E\). The data \(E, K, L, b, \) and \(c\) define a pair of conic problems

\[(P) : \text{minimize } \langle c, x \rangle \quad \text{subject to } \quad x \in K \cap (L + b),
\]

\[(D) : \text{minimize } \langle y, b \rangle \quad \text{subject to } \quad y \in K^* \cap (L^\perp + c),
\]

where \(K^* \subseteq E\) is the cone dual to \(K\), \(L^\perp \subseteq E\) is the orthogonal complement to \(L\), \(L + b \subseteq E\) and \(L^\perp + c \subseteq E\) are affine subspaces. \(P\) and \(D\) are called, respectively, the primal and dual problems associated with the above data.

A relation between the optimal solutions of the primal and dual problems was provided in Theorem 4.2.1 of Ref. [90]:

**Theorem 2** Let \((P), (D)\) be a primal-dual pair of conic problems as defined above, and let the pair be such that

1. The set of primal solutions \(K \cap (L + b)\) intersects \(\text{int } K\);
2. The set of dual solutions \(K^* \cap (L^\perp + c)\) intersects \(\text{int } K^*\);
3. \(\langle c, x \rangle\) is lower bounded for all \(x \in K \cap (L + b)\).

Then both the primal and the dual problems are solvable with polynomial-time interior-point methods, and the optimal solutions \(x^*\) and \(y^*\) satisfy the relation

\[
\langle c, b \rangle = \langle c, x^* \rangle + \langle y^*, b \rangle.\] (D5)

We now apply the above results to the primal-dual SDP pairs \([D1] [D2]\) and \([D4] [D3]\). Let \(L_s\) be the linear space spanned by the operators of general supermaps on bistochastic channels, which is characterized by Eq. \([C12]\):

\[
L_s := \left\{ S \mid [B_0]S - [A_0 B_0]S - [A_0 B_0]S + 2[A_0 A_1 B_0]S - [A_0 A_1 B_0]S = 0 \right\}, \] (D6)

having used the notation

\[
\langle X \rangle := \text{Tr}_X[S] \otimes \frac{I_X}{d_X}.\] (D7)

for a system \(X\) of dimension \(d_X\). Recall the subspaces \(L_f\) and \(L_b\) defined in Section \([C]\) which are the linear span of the operators of forward and backward setups, respectively. Since every operator of supermap we consider is in \(L_s\), it suffices to optimize witnesses within \(L_s\). Let us start by translating these SDPs into the language of Definition 1. To this purpose, we define the following data of a conic optimization problem:

\[
\begin{align*}
E & = L_s \times L_s, \\
L & = \{(T, T) \mid T \in L_s\}, \\
b & = (S, 0), \\
c & = \left(0, \frac{I}{d_B d_A}\right).
\end{align*}\] (D8)

For the pair \([D1] [D2]\), we define

\[
K_1 = \{T \in \text{coni(S)} \mid \text{Tr}(WT) \geq 0\} \times \text{coni(S)}.\] (D9)

For the pair \([D4] [D3]\), we define

\[
K_2 = \text{coni(S)}_{\text{definite}} \times \text{coni(S)}.\] (D10)

To show that the set of primal solutions of the conic optimization problem intersects the interior of the convex cone \(K_1(K_2)\), we prove the following lemma which shows that the sum of the identity operator and an arbitrary operator with a constrained norm from \(L_s\) is contained in the convex cone of \(S_{\text{definite}}\):
Lemma 1 Every operator $S \in \mathcal{L}$, with Hilbert-Schmidt norm bounded as $\|S\|_2 \leq 1/2$ satisfies the condition $I + S \in \mathcal{S}_{\text{definite}}$.  

Proof. This proof is similar to the proof of Lemma 7 of [55]. Let $S$ be such an operator. $I + S$ can be decomposed into $S_f + S_b$ where  

$$S_f = \frac{I}{2} + [A_0]S, \quad S_b = \frac{I}{2} + S - [A_0]S. \tag{111}$$

We can check that $S_f \in \mathcal{L}_f$ and $S_b \in \mathcal{L}_b$. Notice that $[A_0]S$ and $S - [A_0]S$ are orthogonal. By Pythagoras' theorem, it holds that  

$$\|\{A_0\}S\|_2^2 + \|S - [A_0]S\|_2^2 = 2. \tag{112}$$

Therefore,  

$$\max\{\|\{A_0\}S\|, \|S - [A_0]S\|\} \leq \max\{\|\{A_0\}S\|_2, \|S - [A_0]S\|_2\} \leq \|S\|_2 \leq \frac{1}{2}. \tag{113}$$

where $\| \cdot \|$ is the standard operator norm (i.e. the maximal singular value). Therefore, both $S_f$ and $S_b$ are positive. So, $S_f \in \mathcal{P} \cap \mathcal{L}_f$ and $S_b \in \mathcal{P} \cap \mathcal{L}_b$. It follows that $I + S$ belongs to $\mathcal{S}_{\text{definite}}$.  

Now we check that the three conditions of Theorem 2 are satisfied by the two pairs of SDP problems [D1] and [D4]. Let $S$ be the Choi operator of a setup.  

1. Since $\|S\|_2 \leq \|S\|_2 = \text{Tr}(S) = d_B d_A$ and $S \in \mathcal{L}_s$, Lemma 4 implies that  

$$S + \lambda I \in \mathcal{S}_{\text{definite}}, \quad \forall \lambda \geq 2d_B d_A. \tag{114}$$

So by choosing $\lambda_0 > 2d_B d_A$, we have $(\lambda_0 I, \lambda_0 I) \in \mathcal{L}$ and  

$$b + (\lambda_0 I, \lambda_0 I) = (S + \lambda_0 I, \lambda_0 I) \in \text{int} K_2. \tag{115}$$

2. Let $W_0 = \frac{I}{2d_B d_A}$. We have $(W_0, -W_0) \in \mathcal{L}^+$ and  

$$(W_0, -W_0) + c = (W_0, W_0). \tag{116}$$

$(W_0, W_0)$ is an interior point of the dual cone of $S \times S$ because $\text{Tr}(W_0 T) = \text{Tr}(W_0 T') = 1/2 > 0$ for every $T, T' \in S$.  

3. Both [D1] and [D4] are lower bounded by 0 because $T$ is a positive operator.
Due to the relation
\[ \mathcal{K}_2 \subseteq \mathcal{K}_1 \subseteq \text{coni}(S) \times \text{coni}(S), \]  
we have
\[ \text{int}(\mathcal{K}_2) \subseteq \text{int}(\mathcal{K}_1), \]  
and
\[ \text{int}((\text{coni}(S) \times \text{coni}(S))^*) \subseteq \text{int}(\mathcal{K}_1^+) \subseteq \text{int}(\mathcal{K}_2^+). \]

It follows that the two pairs of SDP problems \((D1, D2)\) and \((D4, D3)\) satisfy the three conditions of Theorem 2 and thus can be solved efficiently.

### Appendix E: Details on the source and measurement

All the experiments reported in this paper used a heralded single photon source based on a spontaneous parametric down-conversion (SPDC) process on a type-II cut ppKTP crystal. The crystal was pumped by focusing a 2.5 mW diode laser centered at 404 nm on it using a convex lens (focal length is 12.5 cm). Setting the polarization of the pump laser to be horizontal, we generated pairs of correlated photons centered at 808 nm in a polarization state \(|H\rangle|V\rangle\), which were then separated by a PBS. The pump laser was blocked with long pass and narrow band pass filters. After this, the photon pairs were coupled into single-mode fibers and detected with single photon detectors (photon counting module from PerkinElmer). The idler photon was used as a herald and the signal photon was sent to our experiment. When setting the coincidence window to be 1 ns, the observed coincidence rate of the photon source was about 20000 pairs per second, the counting rate of each detector was about 60000, and thus the coincidence efficiency was 0.33. The coincidence rate was attenuated to 1850 pairs per second after the signal photon passed through the whole apparatus.

An important factor in our experiments is to guarantee a high interference visibility in the Mach-Zehnder interferometer, which in the case of the game is directly related to the success probability. The coherence length of our photon source was over 1000 \(\mu m\). The length difference between the two interference paths was ensured to be within the coherence length by using a trombone-arm delay line composed of a translation stage with a precision of 10 \(\mu m\). The phase between the two interference paths was stabilized by using a piezoelectric transducer (not shown in Fig. 2 of the main text). The interference visibility was measured to be \(0.9921 \pm 0.0035\) in our experiment (see the end of Section H for more details).

It is also worth mentioning that our setup guarantees a high single-photon purity, which can be estimated from the expression \(\sqrt{1 - g^{(2)}(0)}\), where \(g^{(2)}(0)\) is the heralded idler-idler self-correlation. In the end of Section H we briefly discuss how \(g^{(2)}(0)\) can be estimated from our experimental data (including coincidence rates between de-
FIG. 6. Experimental data for the optimal witness (settings $|0\rangle\langle+|$, $|1\rangle\langle+|$, $|+\rangle\langle+|$, and $|+\rangle\langle+i|$). The figure shows the outcome probabilities of different measurements on the control and target qubits, in the setting where the target qubit is measured on the state $|+\rangle$ and re-prepared in the state $|0\rangle$ (a), $|1\rangle$ (b), $|+\rangle$ (c), and $|+\rangle$ (d). The data representation in this figure follows the same convention as in Fig. 4.

FIG. 7. Experimental data for the optimal witness (settings $|0\rangle\langle+i|$, $|1\rangle\langle+i|$, $|+\rangle\langle+i|$, and $|+\rangle\langle+i|$). The figure shows the outcome probabilities of different measurements on the control and target qubits, in the setting where the target qubit is measured on the state $|+\rangle$ and re-prepared in the state $|0\rangle$ (a), $|1\rangle$ (b), $|+\rangle$ (c), and $|+\rangle$ (d). The data representation in this figure follows the same convention as in Fig. 4.
FIG. 8. Probabilities in the decomposition of the simplified witness $W'$. The dots represents the experimentally observed probabilities. The corresponding theoretical predictions shown in colored bars (red, cyan, yellow, and green bars represent results when the target state is reprepared to $|0\rangle, |1\rangle, |+\rangle, \frac{1}{\sqrt{2}}(|0\rangle + i|1\rangle)$ respectively). The $x$ axis enumerates the outcomes of measurements on the control qubit and of the switched measurement on the target qubit, where ‘0’, ‘1’, ‘2’, ‘3’ are used to represent the states $|0\rangle, |1\rangle, |+\rangle, (|0\rangle + i|1\rangle)/\sqrt{2}$ respectively.

Simultaneously, the probabilities used in the experimental evaluation of the simplified witness $W'$ are shown in Fig. [8].

Appendix F: Witnesses for the QTF

Here we use the SDP Eq. (C15) to calculate the optimal witness of input-output indefiniteness for the quantum time flip supermap $\mathcal{F}$, with the control qubit initialized in the state $|+\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$.

Consider the case where the target system is a qubit. The input (output) qubit of a bistochastic channel $\mathcal{C}$ will be denoted by $A_1 (A_0)$, while the input (output) qubits of $\mathcal{F}(\mathcal{C})$ will be denoted by $B_1 = B_{1i}B_{1c}$ ($B_0 = B_{0i}B_{0c}$), where $B_{1i}$ ($B_{0i}$) is the input (output) target qubit and $B_{1c}$ ($B_{0c}$) is the input (output) control qubit. The Choi operator of QTF is given by

$$\text{Choi}(\mathcal{F}) = |V\rangle\langle V|,$$  

with

$$|V\rangle = |I\rangle_{A_1B_{1i}} \otimes |I\rangle_{A_0B_{0i}} \otimes |0\rangle_{B_{1c}} \otimes |0\rangle_{B_{0c}} + |I\rangle_{A_0B_{0i}} \otimes |I\rangle_{A_1B_{1i}} \otimes |1\rangle_{B_{1c}} \otimes |1\rangle_{B_{0c}}.$$  

It follows that the Choi operator including the state preparation of the control qubit in the state $|+\rangle$ is

$$S_{\text{QTF},|+\rangle\langle +|} = |V^+\rangle\langle V^+|,$$

with

$$|V^+\rangle = \frac{1}{\sqrt{2}}|I\rangle_{A_1B_{1i}} \otimes |I\rangle_{A_0B_{0i}} \otimes |0\rangle_{B_{1c}} + |I\rangle_{A_0B_{0i}} \otimes |I\rangle_{A_1B_{1i}} \otimes |1\rangle_{B_{1c}}.$$
Solving the SDP \((C15)\) for \(S = S_{\text{QTF},+}(+)\), we then obtain the robustness \(r(S_{\text{QTF},+}(+)) \approx 0.4007\) and a matrix representation of the optimal witness \(W^{\text{opt}}\). To measure the witness \(W^{\text{opt}}\) in the experiment, we decompose it into a collection of linearly independent operations, including state preparations for \(B_n\), measurements on \(A_1\), and state repreparations of \(A_0\), and measurements on \(B_{oc}\) and \(B_{oc}\). To be bidirectional, the measure-and-reprepare operations are required to be bistochastic instruments \([21]\), i.e. the operators \(\{M_j\}\) of such an instrument satisfy the condition that \(\sum_j M_j\) is a bistochastic channel. In the case of QTF, we realized bistochastic instruments by measuring the system \(A_1\) in some orthonormal basis \(\{|v_0\}, |v_1\}\), and then repreparing states from another orthonormal basis \(\{|w_0\}, |w_1\}\), with the state repreparation depending on the measurement outcome. Explicitly, the decomposition is \(W^{\text{opt}} = \sum a, b, c, d, e \alpha_{a, b, c, d, e} W_{abcde}\), where \(\alpha_{a, b, c, d, e}\) are real coefficients (obtained from the solution of the SDP),

\[
W_{abcde} = (\overline{\alpha}(\overline{\beta})_{BA} \otimes |b\rangle\langle b|_{A1} \otimes \overline{\beta}|c\rangle\langle c|_{A0}
\otimes |d\rangle\langle d|_{B_{oc}} \otimes |e\rangle\langle e|_{B_{oc}}\), \tag{F5}
\]

and the vectors \(|a\rangle, |b\rangle, |c\rangle, |d\rangle\) and \(|e\rangle\) are chosen from the set

\[
|0\rangle, |1\rangle, |0\rangle \pm |1\rangle, \frac{|0\rangle + i|1\rangle}{\sqrt{2}}, \frac{|0\rangle - i|1\rangle}{\sqrt{2}}. \tag{F6}
\]

The number of terms that contribute to the decomposition of \(W^{\text{opt}}\) is 794.

The robustness is then given by the expression

\[
r(S_{\text{QTF},+}(+)) = -\text{Tr}(W^{\text{opt}} S_{\text{QTF},+}(+)) = -\sum a, b, c, d, e \alpha_{a, b, c, d, e} p(a, b, c, d, e), \tag{F7}
\]

where \(p(a, b, c, d, e) = \text{Tr}[W_{abcde} S_{\text{QTF},+}(+)]\) is the probability of the event labeled by \((a, b, c, d, e)\). In the experiment, we estimated the probabilities \(p(a, b, c, d, e)\), and inserted the estimates into the expression of the robustness, obtaining the experimental value \(0.345 \pm 0.005\).

We also consider witnesses of input-output indefiniteness that require fewer measurement settings. For this purpose, we restrict the optimization of the robustness to a subset of witnesses the cases that can be estimated by measuring only \(A_1, A_0, B_{oc}\), fixing the state of \(B_n\) to \(|0\rangle\) and tracing out \(B_{oc}\). These witnesses can be computed by adding the following constraint to Eq. \((C15)\):

\[
W = |0\rangle\langle 0|_{B_n} \otimes I_{B_{oc}} \otimes W^{\text{reduce}}, \tag{F8}
\]

where \(W^{\text{reduce}}\) is a Hermitian matrix on \(A_1 \otimes A_0 \otimes B_{oc}\). The maximal robustness under this constraint is given by \(r'(S_{\text{QTF},+}(+)) \approx 0.1716\). The witness that achieves maximal robustness, denoted by \(W'\), can be decomposed as \(W' = \sum b, c, e \beta_{b, c, e} W'_{bce}\), where \(\beta_{b, c, e}\) are real coefficients, \(W'_{bce} = |0\rangle\langle 0|_{B_n} \otimes I_{B_{oc}} \otimes \overline{\beta}|b\rangle\langle b|_{A1} \otimes |c\rangle\langle c|_{A0} \otimes |e\rangle\langle e|_{B_{oc}}\), and the vectors \(|b\rangle, |c\rangle, |e\rangle\) are chosen from the set \((F6)\). Now, the number of terms that contribute to the decomposition is only 48. The robustness is \(r'(S_{\text{QTF},+}(+)) = \sum b, c, e \beta_{b, c, e} p(b, c, e)\), with \(p(b, c, e) = \text{Tr}[W'_{bce} S_{\text{QTF},+}(+)]\). By experimentally estimating the probabilities \(p(b, c, e)\), we then obtain the experimental value \(0.140 \pm 0.004\).

We show that 3 is the minimum number of qubits to be measured for witnesses of input-output indefiniteness of QTF. Thus the complexity of witnesses of QTF can not be further reduced.

**Lemma 2** Supermaps transforming bistochastic channels to unit probability has no indefiniteness of input-output direction.

**Proof.** Let \(S\) be the Choi operator of a supermap transforming every bistochastic channel (from system \(A_1\) to system \(A_0\)) to unit probability. Then \(S\) is a positive operator on \(\mathcal{H}_{A1} \otimes \mathcal{H}_{A0}\) satisfying \(S = [A_1]S + [A_0]S - [A_1A_0]S\). Consider the spectral decomposition

\[
[A_1]S = \sum_i \lambda_i |u_i\rangle\langle u_i| \otimes I_{A_0}, \tag{F9}
\]

\[
[A_0]S = I_{A1} \otimes \sum_i \mu_j |v_j\rangle\langle v_j| \tag{F10}
\]

Note that \([A_1]S\) and \([A_0]S\) are simultaneously diagonalizable. Let \([A_1]S\) and \([A_0]S\) be positive operators. Positivity of \(S\) implies that, for any \(i, j\), the sum of eigenvalues satisfies \(\lambda_i + \mu_j \geq m\) and thus the minimum satisfies \(\lambda_{\text{min}} + \mu_{\text{min}} \geq m\). Now we can decompose \(S\) into the sum of two positive operators: \(S = S_f + S_b\) where

\[
S_f = [A_1]S - \lambda_{\text{min}} I_{A1} \otimes I_{A_0} \tag{F11}
\]

and

\[
S_b = [A_1]S - (m - \lambda_{\text{min}}) I_{A1} \otimes I_{A_0}. \tag{F12}
\]

It holds that \(S_f = [A_1]S_f\) and \(S_b = [A_1]S_b\). Therefore the supermap \(S\) is a random mixture of a forward and backward supermaps (corresponding to the operators \(S_f\) and \(S_b\), respectively). ■

**Theorem 3** Any witness of QTF has to include variations on both system \(A_1\) and system \(A_0\), as well as one of the control systems: \(B_{oc}\) and \(B_{oc}\).

**Proof.** If there is no variation on the composite system of \(A_1\) and \(A_0\), i.e. a fixed bistochastic channel \(\mathcal{C}_0\) is applied on system \(A_1\) and \(A_0\), then the reduced process on the remaining systems is a simple channel from \(B_nB_{oc}\) to \(B_{oc}B_{oc}\), which is irrelevant to the input-output direction of \(A_1\) and \(A_0\). If there is no variation on system \(A_0\) in a witness \(W\), i.e. \(W = \mathcal{F}_{A0} W \otimes \mathcal{I}_{A1}\). Then the expectation of \(W\) on \(S_{\text{QTF}}\) satisfies

\[
\text{Tr}(WS_{\text{QTF}}) = \text{Tr}\left[W \left(\frac{1}{d_{A0}} \mathcal{F}_{A0} S_{\text{QTF}} \otimes I_{A0}\right)\right] \tag{F13}
\]
The reduced operator $\frac{1}{\mathcal{D}_{O}} \text{Tr}_{A_0} S_{\text{QTF}}$ turns out to be the
the Choi operator of a channel $\mathcal{C}$ from $B_1 B_2$ to $A_1 B_3 B_4$ since $\frac{1}{\mathcal{D}_{O}} \text{Tr}_{A_1 B_3 B_4} S_{\text{QTF}} = I_{B_3} \otimes I_{B_4}$. Hence the
operator $\frac{1}{\mathcal{D}_{O}} \text{Tr}_{A_0} S_{\text{QTF}} \otimes I_{A_0}$ corresponds to a forward
supermap consisting of the channel $\mathcal{C}$ and a discarding operation on system $A_0$. It follows that the expectation $\text{Tr}(W_{\text{QTF}})$ must be non-negative. Similarly, if there is no variation on system $A_1$, the expectation of wit-nesses must be non-negative. Thus, variations on both system $A_1$ and system $A_0$ is necessary for certification of input-output indefiniteness of QTF.

In the case that variations occur only on systems $A_1$ and $A_0$, according to Lemma 2, the reduced operator of any supermap corresponds to a forward supermap transforming bistochastic channels to unit probability and thus has no indefiniteness of input-output direction. Hence variations restricted to systems $A_1$ and $A_0$ are not sufficient.

Now we show that a variation on one of the control systems of QTF is necessary. If the control qubit $B_c$ of QTF is initialized in a fixed state $\sigma$ (or the target $B_d$ and control system $B_c$ are initialized in a joint state $\rho$), and the control qubit $B_{oc}$ is traced out in the end, then QTF becomes a random mixture of a forward supermap and backward supermap. To see this, we compute the Choi operator of the supermap after fixing the corresponding deterministic operations on systems $B_{ic}$ ($B_1 B_2$) and system $B_{oc}$:

$$
\text{Tr}_{B_{ic} B_{oc}} (S_{\text{QTF}} \sigma^T_{B_{ic}}) = \langle 0 | \sigma | 0 \rangle \langle I | \langle I |_{A_1 B_2} \otimes | I \rangle \langle I |_{A_2 B_4} + \langle 1 | \sigma | 1 \rangle \langle I | \langle I |_{A_2 B_4} \otimes | I \rangle \langle I |_{A_1 B_4}
$$

$$
\text{Tr}_{B_{ic} B_{oc}} (S_{\text{QTF}} \rho_{B_{ic} B_{oc}}) = P_0 \otimes | I \rangle \langle I |_{A_2 B_4} + P_1 \otimes | I \rangle \langle I |_{A_1 B_4}
$$

where

$$
P_0 = \text{Tr}_{B_{ic} B_{oc}} (| I \rangle \langle I |_{A_2 B_4} \otimes | 0 \rangle \langle 0 |_{B_{ic}} \rho^T_{B_{ic} B_{oc}})
$$

and

$$
P_1 = \text{Tr}_{B_{ic} B_{oc}} (| I \rangle \langle I |_{A_2 B_4} \otimes | 1 \rangle \langle 1 |_{B_{ic}} \rho^T_{B_{ic} B_{oc}})
$$

are positive operators.

In conclusion, any witness of QTF has to include variations on both system $A_1$ and system $A_0$, as well as one of the control systems: $B_{ic}$ and $B_{oc}$.

### Appendix G: Bipartite witnesses of input-output indefiniteness

Here we introduce the notion of bipartite witness of input-output indefiniteness, distinguishing between two levels of strength of this notion and showing that the quantum game discussed in the main text is an example of the weaker kind, while the single-device witnesses introduced earlier in our paper provide examples of the stronger kind.

1. **Strong vs weak witnesses**

Consider a setup that uses a pair of bidirectional devices (mathematically represented by their forward processes $A$ and $B$, respectively) to generate a new device (mathematically represented by a quantum channel; $C$). The setup can be represented by a Choi operator, acting on the Hilbert spaces of systems $A_1, A_2$ (input and output of device $A$), $B_1, B_2$ (input and output of device $B$), and $C_1, C_2$ (input and output of device $C$). Mathematically, the setup is a bilinear supermap, sending pairs of bistochastic channels $(A,B)$ into (gen-erally non-bistochastic) channels $C$. These supermaps can be naturally extended to supermaps transforming no-signaling bipartite bistochastic channels, of the form $N = \sum_i x_i A_i \otimes B_i$ where $A_i$ and $B_i$ are bistochastic and $x_i$ are real coefficients, into channels $C$. The set of Choi operators of these supermaps has been characterized in Ref. [24].

Now, consider the subset of bipartite supermaps that use both devices $A$ and $B$ in the forward direction, meaning that the action of these supermaps is well-defined even if $A$ and $B$ are not necessary (non-bistochastic) channels.

The supermaps coincide with the supermaps defined in Refs. [30] [31], where the input-output direction is fixed and the relative order of the devices $A$ and $B$ can be indefinite. The Choi operators of these supermaps satisfy the constraints [30]

$$
S \geq 0,
$$

$$
\text{Tr}_{A_1 A_2 B_1 B_2} S = d_A d_B I_{C_1},
$$

$$
[| A_1 A_2 C_1 |] S = [| A_1 A_2 B_2 C_2 |] S,
$$

$$
[| B_1 B_2 C_2 |] S = [| A_1 B_1 B_2 C_2 |] S,
$$

$$
[| C_1 |] S = [| A_1 C_1 |] S - [| B_1 C_1 |] S + [| A_2 B_2 C_2 |] S.
$$

We denote the set of Choi operators satisfying these constraints as $S_{\text{def},bw}^f$, meaning that they are Choi operators of supermaps that use both devices $A$ and $B$ in the forward direction.

Similarly, we define the set of (Choi operators) of supermaps that use both devices in the backward direction. Mathematically, the set $S_{\text{def},bw}^b$ is characterized by the constraints

$$
S \geq 0,
$$

$$
\text{Tr}_{A_1 A_2 B_1 B_2} S = d_A d_B I_{C_1},
$$

$$
[| A_1 A_2 C_1 |] S = [| A_1 A_2 B_1 C_1 |] S,
$$

$$
[| B_1 B_2 C_1 |] S = [| A_1 B_1 B_2 C_1 |] S,
$$

$$
[| C_1 |] S = [| A_1 C_1 |] S + [| B_1 C_1 |] S - [| A_2 B_2 C_2 |] S.
$$

which can be obtained from the constraints (G1) by exchanging the roles of the input and output systems.

The sets $S_{\text{def},bw}^f$ and $S_{\text{def},bw}^b$ are compatible with a global input-output direction, defined jointly for both devices $A$ and $B$. These sets are especially important in scenarios where the input-output direction coincides...
with the arrow of time. In this case, the set \( S_{\text{definite}} \) represents the largest set of operations accessible to an agent that operates in the forward time direction, while the set \( S_{\text{definite}} \) represents the largest set of operations accessible to a hypothetical agent that operates in the backward time direction.

In general, however, one can also consider setups that use device \( A \) in the forward direction and device \( B \) in the backward direction, or vice-versa with the arrow of time. In this case, the set \( S_{\text{definite}} \) and \( S_{\text{opposite}} \) are characterized by the constraints

\[
S \geq 0,
\]

\[
\text{Tr}_{A_1 A_0 B_1 B_0} S = d_A d_B I_{C_1},
\]

\[
|A_1 A_0 C_0| S = |A_1 A_0 B_1 C_0| S,
\]

\[
|B_1 B_0 C_0| S = |A_0 B_1 B_0 C_0| S,
\]

\[
|C_0| S = |A_0 C_0| S + |B_1 C_0| S - |A_1 B_0 C_0| S
\] (G3)

and

\[
S \geq 0,
\]

\[
\text{Tr}_{A_1 A_0 B_1 B_0} S = d_A d_B I_{C_1},
\]

\[
|A_1 A_0 C_0| S = |A_1 A_0 B_1 C_0| S,
\]

\[
|B_1 B_0 C_0| S = |A_0 B_1 B_0 C_0| S,
\]

\[
|C_0| S = |A_1 C_0| S + |B_0 C_0| S - |A_0 B_0 C_0| S
\] (G4)

respectively.

Now, consider the setups that use both devices in a definite input-output direction, with the same input-output direction for both devices. The corresponding set of Choi operators, denoted by \( S_{\text{definite}} \), consists of all Choi operators of the form

\[
S_{\text{same}} = p S_{\text{fwd,fwd}} + (1 - p) S_{\text{bwd,bwd}},
\] (G5)

where \( p \in [0, 1] \) is a probability, and the operators \( S_{\text{fwd,fwd}}, S_{\text{bwd,bwd}} \) belong to the sets \( S_{\text{fwd,fwd}}, S_{\text{bwd,bwd}} \), respectively.

Likewise, we can consider the setups that use the two devices in a definite input-output direction, but with opposite directions for the two devices. The corresponding set of Choi operators, denoted by \( S_{\text{definite}} \), consists of all Choi operators of the form

\[
S_{\text{opposite}} = p S_{\text{fwd,bwd}} + (1 - p) S_{\text{bwd,fwd}},
\] (G6)

where \( p \in [0, 1] \) is a probability, and the operators \( S_{\text{fwd,bwd}}, S_{\text{bwd,fwd}} \) belong to the sets \( S_{\text{fwd,bwd}}, S_{\text{bwd,fwd}} \), respectively.

Finally, we can consider the setups that use both devices in a definite input-output direction, without any restriction on how the input-output direction of the first device is related to the input-output direction of the second device. The Choi operators of these setups are of the form

\[
S_{\text{definite}} = p S_{\text{same}} + (1 - p) S_{\text{opposite}},
\] (G7)

where \( p \in [0, 1] \) is a probability, and the operators \( S_{\text{same}}, S_{\text{opposite}} \) belong to the sets \( S_{\text{definite}}, S_{\text{definite}} \), respectively. The set of Choi operators of the form (G7) will be denoted by \( S_{\text{definite}} \).

We are now ready to define three different notions of witnesses of bipartite input-output indefiniteness.

**Definition 2** An operator \( W \) acting on the composite system \( A_1 A_0 B_1 B_0 C_1 C_0 \) is a weak witness of bipartite input-output indefiniteness if \( \text{Tr}[W S] \geq 0 \) for every \( S \in S_{\text{definite}} \) and \( \text{Tr}[W S] < 0 \) for some \( S \in S \).

**Definition 3** An operator \( W \) acting on the composite system \( A_1 A_0 B_1 B_0 C_1 C_0 \) is a conjugate weak witness of bipartite input-output indefiniteness if \( \text{Tr}[W S] \geq 0 \) for every \( S \in S_{\text{opposite}} \) and \( \text{Tr}[W S] < 0 \) for some \( S \in S \).

**Definition 4** An operator \( W \) acting on the composite system \( A_1 A_0 B_1 B_0 C_1 C_0 \) is a strong witness of bipartite input-output indefiniteness if \( \text{Tr}[W S] \geq 0 \) for every \( S \in S_{\text{definite}} \) and \( \text{Tr}[W S] < 0 \) for some \( S \in S \).

The above definitions imply that an operator \( W \) is a strong witness if and only if \( W \) is both a weak witness and a conjugate weak witness.

### 2. The game as a weak witness

We now show that the game described in the main text is a weak witness, but not a strong one.

First, let us cast the game in the form of a witness. The possible strategies are described by bipartite supermaps which have two slots (corresponding to systems \( A_1 A_0 \) and \( B_1 B_0 \)) and one qubit output \( C_0 \) (compared to the general framework in the previous subsection, here we are taking the input system \( C_1 \) to be trivial). The strategy is carried out by placing the two gates \( U \) and \( V \) in the slots and then measuring the output qubit \( C_0 \) in Fourier basis \( \{ |\pm \rangle \} \) which gives the outcome that \( (U, V) \) belongs to

\[
G_+ = \{ (U, V) | UV^T = U^TV \},
\] (G8)

or

\[
G_- = \{ (U, V) | UV^T = -U^TV \}. \] (G9)

Let \( S \) be the Choi operator of the strategy. If the gates \( (U, V) \in G_+ \), then the probability of winning is

\[
\text{Tr}((\text{Choi}(U) \otimes \text{Choi}(V) \otimes |+)\langle+)^TS). \] (G10)

If the gates \( (U, V) \in G_- \), then the probability of winning is

\[
\text{Tr}((\text{Choi}(U) \otimes \text{Choi}(V) \otimes |-)\langle-)^TS). \] (G11)
In the general case, suppose that $\mu(U, V)$ is a probability measure on $G_+ \cup G_-$. We define the operators $M_+$ and $M_-$ to be

$$M_\pm := \int_{G_\pm} (\text{Choi}(U) \otimes \text{Choi}(V) \otimes |\pm\rangle\langle\pm|)^T d\mu(U, V).$$

(G12)

Then the probability of winning is

$$p_{\text{succ}} =$$

$$\int_{G_+} \text{Tr}((\text{Choi}(U) \otimes \text{Choi}(V) \otimes |+\rangle\langle+|)^T S) \cdot d\mu(U, V)$$

$$+ \int_{G_-} \text{Tr}((\text{Choi}(U) \otimes \text{Choi}(V) \otimes |-\rangle\langle-|)^T S) \cdot d\mu(U, V)$$

$$= \text{Tr}((M_+ + M_-) S).$$

(G13)

Ref. [21] provided examples of probability distributions $d\mu(U, V)$ with the property that the probability of winning is strictly smaller than 1 for every strategy that uses both gates in the forward direction, or both gates in the backward direction. For all those probability distributions, the operator

$$W_{\text{game}} := \frac{I}{d^2} - \frac{M_+ + M_-}{p_{\text{max}}},$$

is a weak witness.

The witness $W_{\text{game}}$ is not a strong witness, because there exist quantum strategies that achieve a unit probability of success while using each of the gates $U$ and $V$ in definite (opposite) input-output directions.

For example, a perfect winning strategy uses the gate $U$ in the forward direction and the gate $V$ in the backward direction, corresponding to the gate $V^T$. The strategy is to insert the gates $U$ and $V^T$ into the quantum SWITCH supermap [31], which turns them into the controlled unitary gate

$$S(U, V^T) = UV^T \otimes |0\rangle\langle0| + V^TU \otimes |1\rangle\langle1|.$$  

(G15)

Now, if the control qubit of the quantum SWITCH is initialized in the state $|+\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$ and the target system is maximally entangled with an auxiliary system $R$, the gate $S(U, V^T)$ produces the output state

$$\frac{|UV^T + V^TU\rangle}{2\sqrt{d}} \otimes |+\rangle + \frac{|UV^T - V^TU\rangle}{2\sqrt{d}} \otimes |-\rangle,$$

(G16)

where we used the double-ket notation

$$|M\rangle = (M \otimes I) |I\rangle.$$  

(G17)

We now show that a projective measurement on the above states can determine with certainty whether the gates $(U, V)$ belong to the $G_+$ or to the $G_-.$

Suppose that $(U, V)$ belongs to $G_+$. In this case, we obtain the relations

$$UV^T + V^TU = UV^T + VU^T \in H_+,$$

(G18)

and

$$UV^T - V^TU = UV^T - VU^T \in H_-,$$

(G19)

where $H_\pm$ are the symmetric and anti-symmetric subspaces of $H \otimes H$.

Instead, if $(U, V)$ belongs to $G_-$, we have

$$UV^T + V^TU = UV^T - VU^T \in H_-,$$

(G20)

and

$$UV^T - V^TU = UV^T + VU^T \in H_+.$$  

(G21)

Hence, it is possible to determine whether the pair $(U, V)$ belongs to $G_+$ or to the set $G_-$ by performing the binary projective measurement with projectors

$$Q_+ = P_+ \otimes |+\rangle\langle+| + P_- \otimes |-\rangle\langle-|$$

(G22)

and

$$Q_- = P_- \otimes |+\rangle\langle+| + P_+ \otimes |-\rangle\langle-|.$$  

(G23)

The (Choi operator of the) above winning strategy belongs to the set $S_\text{definite}.$

3. Strong bipartite witnesses from single-device witnesses

Single-device witnesses can be easily converted into strong bipartite witnesses. Intuitively, the idea is that if we can rule out a definite input-output direction for the use of an individual device, then we can also rule out a definite input-output direction for the combined use of two or more devices.

More formally, suppose that $W \in L(H_A \otimes H_{A_0} \otimes H_{C_1} \otimes H_{C_0})$ is a single-device witness for setups that transform an input bistochastic channel $A$ (with input $A_1$ and output $A_0$) into a channel $C$ (with input $C_1$ and output $C_0$), and that $B$ is the Choi operator of a bistochastic channel with input $B_1$ and output $B_0$. Then, the operator $W \otimes B$ is a strong witness for bipartite setups taking in input two bidirectional devices (with input-output pairs $(A_1, A_0)$ and $(B_1, B_0)$, respectively) and producing in output a device (with input-output pair $(C_1, C_0)$). This particular witness corresponds to plugging a fixed bidirectional device into one slot of the bipartite setup, and testing the input-output indefiniteness in the remaining slot.

Similarly, one can construct strong witnesses of the form $A \otimes W'$, where $A$ is the Choi operator of a bistochastic channel with input $A_1$ and output $A_0$, and $W' \in L(H_{B_1} \otimes H_{B_0} \otimes H_{C_1} \otimes H_{C_0})$ is a single-device witness for setups that transform an input bistochastic channel $B$ (with input $B_1$ and output $B_0$) into a channel $C$ (with input $C_1$ and output $C_0$).

By taking linear combinations of witnesses of the above form, one can easily construct examples of strong bipartite witnesses that take negative values only if both devices are used in an indefinite input-output direction. It
FIG. 9. *Theoretical scheme for the game.* A target system traverses two devices V and U in two different directions dependent on the state of a control qubit, which is prepared in the state $|+\rangle$. To be explicit, the target undergoes $U^T V$ if the control is in $|0\rangle$ (red path) and undergoes $UV^T$ if the control is in $|1\rangle$ (blue path).

FIG. 10. *Experimental setup for the game.* The source and measurements are the same as in the experiment on the witness. The only difference is that instead of the measure-and-reprepare process, we put two unitary gates $U$ and $V$ in the setup. The abbreviations used in the figure are: HWP, half-wave plate; QWP, quarter-wave plate; PBS, polarizing beam splitter; BS, beam splitter ($T/R = 50/50$); RM, reflection mirror; LC, liquid crystal variable retarder; FC, fiber coupler; SPD, single photon detector; DL, trombone-arm delay line.

is worth stressing that, of course, not all strong witnesses are of this form. However, the construction shown in this subsection is conceptually important because it shows that single-device witnesses, like the witnesses measured in our experiments, offer intrinsically stronger certificates of input-output indefiniteness compared to weak bipartite witnesses.

**Appendix H: Experimental setup for the quantum game**

The theoretical scheme and the experimental setup for the quantum game are shown in Fig. 9 and Fig. 10 respectively. Compared to the single-device witness, the game has a simpler realization. The spatial qubit is initialized in the state $|+\rangle$ by a 50/50 beamsplitter, while precise initialization of the polarization is not necessary here, since the winning strategy works equally well for every initial state of the target qubit. Moreover, there measure-and-prepare instruments used in the single-device witness are now replaced by unitary gates. Depending on the state of the spatial qubit, the polarization qubit traverses two optical devices in two alternative directions. The first (second) device rotates the polarization, implementing the unitary gate $U$ ($V$) in the forward direction, and the gate $ZU^T Z$ ($ZV^T Z$) in the backward direction (the Pauli $Z$ in the backward process is then removed by a Pauli $Z$ rotation placed on the appropriate path). Overall, the two paths used in the setup result in the two combinations $U^T V$ and $UV^T$. The two paths are then recombined by a beamsplitter, and photon detections are performed at its output. The game involves the implementation of several unitary gate pairs $(U, V)$ satisfying either the property $UV^T = U^T V$ or the property $UV^T = -U^T V$. In the experiment, we
implement the following sets of gates

\[ S_+ = \{(I, I), (I, X), (I, Z), (X, I), (X, X), (X, Z), (Z, I), (Z, X), (Z, Z), (U_1, V_1), (V_1, U_1), (U_2, V_2), (V_2, U_2)\}, \]

\[ S_- = \{(Y, I), (Y, X), (Y, Z), (I, Y), (X, Y), (Z, Y), (U_3, V_3), (V_3, U_3)\}, \]

where \( I \) is identity gate and \( X, Y, Z \) are Pauli gates. These gates are implemented using a combination of three waveplates (in a quarter-half-quarter configuration), using the angle settings shown in Table 3.

In this game, every strategy that uses the two devices in the same input-output direction (either forward direction for both devices or backward direction for both devices) will fail at least 11% of the times [24]. This lower bound applies even if the relative order between the two gates is indefinite: for example, combining the forward processes (backward processes) will fail at least 11% of the times [24]. More generally, no coherently controlled choice of quantum circuits containing a single use of the forward processes (backward processes) in a general form of indefinite order [30, 31].

In contrast, a player that uses the two gates in a superposition of the opposite input-output directions \((U, V^T)\) and \((U^T, V)\), can bring the error probability arbitrarily close to zero. This result is achieved by combining the two devices into a controlled unitary gate \( W = U V^T \otimes |0\rangle \langle 0| + V U^T \otimes |1\rangle \langle 1| \). Then, the relations \( U V T = U V^T \) and \( U^T V = -U V^T \), guarantee that one has either \( W = U V^T \otimes I \) or \( W = U V^T \otimes Z \), respectively, with \( Z = |0\rangle \langle 0| - |1\rangle \langle 1| \). If the control qubit is initialized in the state \( |+\rangle := (|0\rangle + |1\rangle) / \sqrt{2} \), the gates \( U V^T \otimes I \) or \( W = U V^T \otimes Z \) turn it to the orthogonal states \( |+\rangle \) and \( |\rangle := (|0\rangle - |1\rangle) / \sqrt{2} \), respectively. Hence, a projective measurement of the control system can determine without error which of the two alternative relations holds.

The experimental data from our implementation are provided in Fig. 3. The average success probability over all pairs of gates is 99.6% ± 0.18%, while the worst-case error probability over all trials is approximately 0.68 ± 0.19%. This value is approximately 16 times smaller than 1%, the lower bound on the error probability for all possible strategies with definite-input-output direction.

Before concluding this section, we provide a brief discussion on the role of the Mach-Zehnder visibility and single-photon purity. Imperfect Mach-Zehnder visibility is the main source of errors in our setup. Ideally, if the visibility were perfect, the photons’ output port would be fully determined by the commutation properties of the unitary gates tested in the game. With a realistic interferometer, however, some of the photons will reach the wrong port, leading the player to give a wrong answer. The probability that the player answers correctly, per detector click, is then given by \( P_{\text{right}} = n_{\text{right}}/(n_{\text{right}} + n_{\text{wrong}}) \), where \( n_{\text{right}} \) (\( n_{\text{wrong}} \)) is the number of clicks at the right (wrong) detector. The probability of success in the game is then related to the Mach-Zehnder visibility \( V = (n_{\text{right}} - n_{\text{wrong}})/(n_{\text{right}} + n_{\text{wrong}}) \) via the relation \( P_{\text{succ}} = (1 + V)/2 \). In our experiment, the visibility is \( V = 0.9921 \pm 0.0035 \), which yields the value \( P_{\text{succ}} = (1 + V)/2 = 0.9966 \pm 0.0018 \).

In the above analysis, we considered the success probability per detector click. Another relevant quantity is the success probability per photon entering our setup. Ignoring photon loss, detector efficiency, and detector noise, the total number of detector clicks \( n_{\text{click}} = n_{\text{right}} + n_{\text{wrong}} \) in our experiment is close to the number of photons \( n_{\text{photon}} \) entering into the setup, up to a small correction due to the presence of multiphotons generated by SPDC.

SPDC generates pairs of \( k \)-photons with probability \( p_k \), approximately proportional to \( (P_{\text{pump}})^k \) where \( P_{\text{pump}} \) is the pump power. A \( k \)-photon, however, is counted as a single event by the PerkinElmer detectors. This implies that, when \( n_{\text{photon}} \) is large, there will be approximately \( n_{\text{photon}} p_k/k \) clicks originating from the \( k \)-photon contribution of SPDC. Hence, the total number of clicks at the output of the Mach-Zehnder interferometer will be \( n_{\text{click}} = n_{\text{photon}} (p_1 + p_2/2 + p_3/3 + \ldots) \). In practice, in the low power regime of our experiment (\( P_{\text{pump}} \approx 2.5 \text{ mW} \)), \( p_k \) is negligible for \( k > 2 \), and therefore one can assume \( p_1 + p_2 \approx 1 \).

In fact, the correction due to the \( p_2 \) term is itself almost negligible. The value of \( p_2 \) is closely related to the single-photon purity, given by \( 1 - g^{(2)}(0) \) where \( g^{(2)}(0) \) is the heralded idler-idler self-correlation function. The latter can be estimated from experimental data according to a heralded Hanbury-Brown-Twiss configuration [31]. In this configuration, a SPDC process generates a pair of photons, with the signal photon acting as a herald (detected by a detector \( H \)) and the idler photon split into two path modes by a 50/50 beamsplitter, with detectors \( A \) and \( B \) at its output ports. The self-correlation function is then given by \( g^{(2)}(0) = (n_{\text{HA}} n_{\text{HB}})/(n_{\text{HA}} n_{\text{HA}}) \), where \( n_{\text{HA}} = \) the number of counts at detector \( H \), \( n_{\text{HB}} = n_{\text{HB}} \), and \( n_{\text{HA}} \) and \( n_{\text{HB}} \) are the coincidence rates at zero delay time at detectors \( H/A/B, H/A, \) and \( H/B \), respectively. In our experiment, \( n_{\text{HA}} \) can be estimated as \( n_{\text{HA}} = 2 \tau n_{\text{HAB}} \), where \( \tau \) is the coincidence window, \( n_{\text{HAB}} \) is the coincidence efficiency of the SPDC source, and \( n_{\text{A}} \) and \( n_{\text{B}} \) are the counts at detectors \( A \) and \( B \), respectively. The coincidence rate was 20000, \( n_{\text{H}} \) and \( n_{\text{A}} + n_{\text{B}} \) were 61000 and 58000, respectively, and \( \tau \) was set to be 1 ns. Using these data, we can estimate \( n_{\text{HAB}} = 0.33(n_{\text{HA}} n_{\text{HB}})/(n_{\text{HA}} n_{\text{HA}}) \), \( n_{\text{HAB}} = 0.33 \), and \( n_{\text{HAB}} = 0.56 \). From these values, we then obtain \( g^{(2)}(0) \approx 0.0003 \), corresponding to a single-photon purity \( \sqrt{1 - g^{(2)}(0)} \approx 0.9998 \).

From the coincidence rates, one can also directly estimate the two-photon pair event probability, which is given by \( p_2 = (2n_{\text{HAB}})/(n_{\text{HA}} + n_{\text{HB}}) \approx 6 \times 10^{-5} \).
| Unitary gate | QWP1 | HWP | QWP2 |
|-------------|------|-----|------|
| $I$         | 0°   | 0°  | 0°   |
| $X$         | 0°   | 45° | 0°   |
| $Y$         | 90°  | 45° | 0°   |
| $Z$         | 90°  | 0°  | 0°   |
| $U_1 = (X - Y)/\sqrt{2}$ | 45°  | 67.5° | 135° |
| $V_1 = (X + Y)/\sqrt{2}$ | 135° | 67.5° | 45° |
| $U_2 = (Z - Y)/\sqrt{2}$ | 0°   | 22.5° | 90° |
| $V_2 = (Z + Y)/\sqrt{2}$ | 90°  | 22.5° | 0° |
| $U_3 = (I - iY)/\sqrt{2}$ | 22.5° | 135° | 67.5° |
| $V_3 = (I + iY)/\sqrt{2}$ | 67.5° | 135° | 22.5° |

TABLE I. The chosen unitary gates and corresponding setting angles of QWPs and HWPs for preparing these gates in our experiment.

FIG. 11. Experimental data for the quantum game. The figure shows the experimentally estimated probabilities of the outcomes $+$ (red bars) and $-$ (cyan bars) of the Pauli-$X$ measurement used to distinguish two alternative properties of an unknown gate pair $(U, V)$. The measurement data refers to 21 different gate pairs, corresponding to different combinations of single-qubit gates in the set $\{I, X, Y, Z, U_1, U_2, U_3, V_1, V_2, V_3\}$. Gate pairs in the set $S_+ (S_-)$ satisfy the condition $UV^T = U^TV (UV^T = -U^TV)$. The experimental data show that the outcome $+$ ($-$) has high probability to occur for gate pairs in the $S_+$ ($S_-$), offering a near unit winning probability in the game.

This value implies that in our game the success probability per photon is $P_{\text{right}} = P_{\text{succ}} n_{\text{click}}/n_{\text{photon}} \approx P_{\text{right}} [p_1 + p_2/2] \approx P_{\text{right}} [1 - p_2/2] \approx 0.9960$. Note that the success probability per photon is equal, within the experimental error bars, to the success probability per click $P_{\text{succ}} = 0.9960 \pm 0.0018$, as a result of the high single-photon purity in our experiment.

Appendix I: A strengthened advantage in the quantum game

Here we show that the advantage demonstrated in our experiment holds even if our setup is compared with arbitrary setups that use the control-unitary gates $\text{ctrl} - U = I \otimes |0\rangle\langle 0| + U \otimes |1\rangle\langle 1|$ and $\text{ctrl} - V = I \otimes |0\rangle\langle 0| + V \otimes |1\rangle\langle 1|$ instead of the original gates $U$ and $V$. As long as these controlled gates are used in a fixed input-output direction (either the forward direction for both gates or the backward direction for both gates), the error probability in the game cannot go below 5.6%. Notably, this result holds even if the relative order between the gates $\text{ctrl} - U$ and $\text{ctrl} - V$ is indefinite.

To derive this result, we adapt the method developed in Ref. [24] to find the minimal worst-case error over all strategies using the gates $U$ and $V$ in a fixed input-output direction, equal for both gates. Here we replace the 2-by-2 matrices $U$ and $V$ with their controlled versions $\text{ctrl} - U$ and $\text{ctrl} - V$, mathematically equivalent to the block diagonal matrices $\tilde{U} = \begin{pmatrix} I & 0 \\ 0 & U \end{pmatrix}$ and $\tilde{V} = \begin{pmatrix} I & 0 \\ 0 & V \end{pmatrix}$, respectively. With this substitution, the minimum probability of error over all possible strategies with definite input-output direction can be cast in the
SDP form

\[
\begin{align*}
\text{minimize} & \quad p_{err} \\
\text{subject to} & \quad \langle \hat{U} | A_O A_I | \hat{V} \rangle_{B_O B_I} \geq 1 - p_{err} \\
& \quad \forall (U, V) \in S_+, \\
\langle \hat{U} | A_O A_I | \hat{V} \rangle_{B_O B_I} & \leq p_{err} \\
& \quad \forall (U, V) \in S_-,
\end{align*}
\]

$$0 \leq P \leq S,$$

$$S \in \text{Det}(A_I, A_O; B_I, B_O).$$

The optimization runs over the variables $p_{err}$, $P$, and $S$, where $p_{err}$ is the probability of error, $P \in L(\mathcal{H}_{A_O} \otimes \mathcal{H}_{A_I} \otimes \mathcal{H}_{B_O} \otimes \mathcal{H}_{B_I})$ is a positive operator, and $S$ is an operator in the set $\text{Det}(A_I, A_O; B_I, B_O)$, consisting of all operators $S$ satisfying the following constraints:

$$S \geq 0,$$

$$|A_I A_O|S = |A_I A_O B_O|S,$$

$$|B_I B_O|S = |A_O B_I B_O|S,$$

$$S = |A_O|S + |B_O|S - |A_O B_O|S.$$ 

Operationally, the pair $\{P, S - P\}$ represents a binary-outcome test that probes the gates $\hat{U}$ and $\hat{V}$ in a definite input-output direction and possibly in an indefinite order relative to one another [62, 91]. The above SDP yields the minimum error probability achievable by this type of setup. Numerical computation with the Python-embedded modeling language CVXPY [92, 93] yields the results that the minimal $p_{err}$ is 5.6%. The same result holds if we replace the forward gates $\text{ctrl} - U$ and $\text{ctrl} - V$ with the backward gates $\text{ctrl} - U^T$ and $\text{ctrl} - V^T$. Hence, a player that uses both gates in the same input-output direction will always have a finite error probability, even if the player has access to controlled unitary gates.