

# Perfect discrimination of no-signalling channels via quantum superposition of causal structures

Giulio Chiribella

Center for Quantum Information, Institute for Interdisciplinary Information Sciences, Tsinghua University, Beijing 100084, China

(Received 23 September 2011; published 10 October 2012)

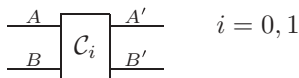
A no-signalling channel transforming quantum systems in Alice’s and Bob’s local laboratories is compatible with two different causal structures: ( $A \leq B$ ) Alice’s output causally precedes Bob’s input and ( $B \leq A$ ) Bob’s output causally precedes Alice’s input. Here I prove that two no-signalling channels that are not perfectly distinguishable in any ordinary quantum circuit can become perfectly distinguishable through the quantum superposition of circuits with different causal structures.

DOI: [10.1103/PhysRevA.86.040301](https://doi.org/10.1103/PhysRevA.86.040301)

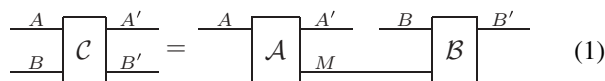
PACS number(s): 03.67.Hk, 03.65.Ud, 03.67.Ac

Distinguishing between two objects is one of the most fundamental tasks in information theory. In quantum information, an instance of the problem is the discrimination between two quantum channels [1–12]. In this scenario one has access to a black box implementing a transformation of quantum systems, which is promised to be either  $C_0$  or  $C_1$ , and the goal is to identify such a transformation with maximum probability of success using a given number of queries to the black box.

Many surprising features of quantum channel discrimination have been discovered so far. For example, two unitary channels that are not perfectly distinguishable with a single query become perfectly distinguishable when a finite number of queries is allowed [1,2]. Other remarkable phenomena arise when the two channels  $C_0$  and  $C_1$  have a bipartite structure, as in the following diagram:

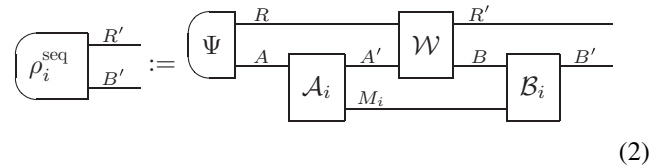


which represents a quantum channel (i.e., a completely positive trace preserving map) sending quantum states on the Hilbert space  $A \otimes B$  to quantum states on the Hilbert space  $A' \otimes B'$ . We can imagine that the channels  $C_0$  and  $C_1$  transform quantum states provided by two users, Alice and Bob. If the state of Alice’s output  $A'$  does not depend on the state of Bob’s input  $B$ , then we say that  $C_0$  and  $C_1$  are *no-signalling from Bob to Alice* (*B-no-signalling*, for short). Eggeling, Schlingemann, and Werner [13] showed that every B-no-signalling channel  $C$  can be realized as the concatenation of a channel  $A$  on Alice’s side followed by a channel  $B$  on Bob’s side, with some information transferred from Alice to Bob via a quantum memory  $M$ , as in the diagram



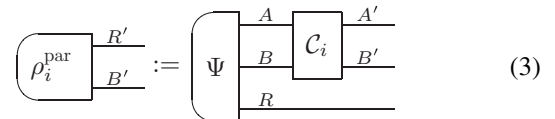
In other words, if a channel does not signal from Bob to Alice, then it is compatible with a causal structure where Alice’s output precedes Bob’s input, denoted by  $A \leq B$ . In this structure the black box provides Alice’s output before Bob supplies his input, as illustrated in the right-hand side of Eq. (1). To discriminate between two B-no-signalling channels we can then use a *sequential strategy* [7], where the channel  $C_i$  (either with  $i = 0$  or with  $i = 1$ ) is inserted in a quantum circuit with causal structure  $A \leq B$ , thus producing the output

state  $\rho_i^{\text{seq}}$  given by



Here  $R$  and  $R'$  are suitable ancillary systems,  $M_i$  is the quantum memory needed for the realization of channel  $C_i$ ,  $\Psi$  is a pure state on  $R \otimes A$ , and  $\mathcal{W}(\rho) = W\rho W^\dagger$ ,  $W^\dagger W = I_{R \otimes A'}$  is an isometry sending states on  $R \otimes A'$  to states on  $R' \otimes B$  [14].

Reference [7] showed that sequential strategies offer an advantage over *parallel strategies*, where the channel  $C_i$  is applied on one side of an entangled input state  $\Psi \in A \otimes B \otimes R$ , producing the output state  $\rho_i^{\text{par}}$  given by



In particular, Ref. [7] exhibited two B-no-signalling channels that can be perfectly distinguished by a sequential strategy, whereas every parallel strategy has a nonzero probability of error. Later, Harrow *et al.* [12] demonstrated the same phenomenon in the absence of a quantum memory, i.e., for two channels  $C_0$  and  $C_1$  of the product form  $C_i = A_i \otimes B_i$ , with  $A_i$  ( $B_i$ ) transforming states on  $A$  ( $B$ ) into states on  $A'$  ( $B'$ ). Channels of this form are a particular example of *no-signalling channels* [15,16], namely, channels that are both B-no-signalling and A-no-signalling [17].

In principle, no-signalling channels can be used in two different causal structures:  $A \leq B$  (the black box processes Alice’s input first and Bob’s input later) and  $B \leq A$  (the black box processes Bob’s input first and Alice’s input later). Usually, when there are two possible alternatives, in quantum theory one can conceive a superposition of them. Can we apply this idea also to the choice of causal order? Recently, Ref. [18] introduced the notion of *quantum superposition of causal structures*, arguing that this new primitive could be achieved in a quantum network where the connections among devices are controlled by the quantum state of a control qubit. A no-signalling channel inserted into such a network would be in a quantum superposition of being used in a circuit with causal structure  $A \leq B$  and of being used in a circuit with causal structure  $B \leq A$ .

Such a network can be thought of as a toy model for a quantum gravity scenario where the causal structure is not defined *a priori*, a scenario originally considered by Hardy [19], who posed the question as to whether indefinite causal structure can be used as a computational resource.

This paper gives a positive answer to Hardy's question, showing that the superposition of causal structures enables completely new schemes for quantum channel discrimination. The advantage of such schemes is demonstrated by exhibiting a concrete example of two no-signalling channels that cannot be perfectly distinguished by any sequential strategy using a single query to the black box, but become perfectly distinguishable through a quantum superposition of sequential strategies with different causal structures [18]. The example involves two-qubit channels  $\mathcal{C}_0$  and  $\mathcal{C}_1$ , with  $\mathcal{C}_0$  consisting of two von Neumann measurements on the same random basis, and  $\mathcal{C}_1$  consisting of two rotations of  $\pi$  around a random pair of orthogonal axes in the Bloch sphere. More generally, the superposition of causal structures presented here allows for a single-query, zero-error discrimination between an arbitrary pair of qubit channels with commuting Kraus operators and an arbitrary pair of qubit channels with anticommuting Kraus operators. These results contribute to the exploration of a new research avenue that aims at demonstrating new physical phenomena and power-ups to information processing arising from the application of quantum theory in the lack of a definite causal structure [18–20].

Before presenting the result, let us make precise what we mean by quantum superposition of causal structures. We can start from the simplest case, where the channels  $\mathcal{C}_0$  and  $\mathcal{C}_1$  are of the product form  $\mathcal{C}_i = \mathcal{A}_i \otimes \mathcal{B}_i$ ,  $i = 0, 1$ . First, let us write down the Kraus forms  $\mathcal{A}_i(\rho) = \sum_k A_{ik} \rho A_{ik}^\dagger$  and  $\mathcal{B}_i(\rho) = \sum_l B_{il} \rho B_{il}^\dagger$ . For each value of  $k$  and  $l$ , a sequential circuit with causal structure  $A \leq B$  [like the circuit in Eq. (2)] yields the (unnormalized) pure state

$$|\Psi_{ikl}\rangle = (B_{il} \otimes I_{R'}) W(A_{ik} \otimes I_R) |\Psi\rangle. \quad (4)$$

Whereas a sequential circuit with causal structure  $B \leq A$  yields

$$|\tilde{\Psi}_{ikl}\rangle = (A_{ik} \otimes I_{\tilde{R}'}) \tilde{W}(B_{il} \otimes I_{\tilde{R}}) |\tilde{\Psi}\rangle, \quad (5)$$

where  $\tilde{R}$  and  $\tilde{R}'$  are suitable ancillary systems,  $\tilde{\Psi} \in B \otimes \tilde{R}$  is a pure state, and  $\tilde{W}$  is an isometry from  $B' \otimes \tilde{R}$  to  $A \otimes \tilde{R}'$ . Note that by suitably choosing the ancillary systems  $R, R', \tilde{R}, \tilde{R}'$  we can assume without loss of generality that the input and output systems of the two circuits are the same, i.e.,  $A \otimes R \simeq B \otimes \tilde{R}$  and  $B \otimes R' \simeq A \otimes \tilde{R}'$ . Suppose now that we have at our disposal a coherent mechanism that chooses the first circuit when the state of a control qubit is  $|0\rangle$ , and the second when the state is  $|1\rangle$ . Such a mechanism could be implemented in a quantum network where the connections among devices are not predetermined, but instead can be controlled by the state of some quantum system, as in the *quantum switch* of Ref. [18]. If the control qubit is prepared in the state  $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$ , then the output of the network is  $\alpha|\Phi_{ikl}\rangle|0\rangle + \beta|\Phi'_{ikl}\rangle|1\rangle$ . Taking the corresponding density matrix and summing over

all possible Kraus elements we then get the output state

$$\begin{aligned} \rho_i^{\text{sup}} := & |\alpha|^2 (\mathcal{B}_i \otimes \mathcal{I}_{R'}) \mathcal{W}(\mathcal{A}_i \otimes \mathcal{I}_R) (|\Psi\rangle\langle\Psi|) \otimes |0\rangle\langle 0| \\ & + |\beta|^2 (\mathcal{A}_i \otimes \mathcal{I}_{R'}) \tilde{\mathcal{W}}(\mathcal{B}_i \otimes \mathcal{I}_R) (|\tilde{\Psi}\rangle\langle\tilde{\Psi}|) \otimes |1\rangle\langle 1| \\ & + \alpha\beta^* \sum_{k,l} |\Psi_{ikl}\rangle\langle\tilde{\Psi}_{ikl}| \otimes |0\rangle\langle 1| \\ & + \alpha^*\beta \sum_{k,l} |\tilde{\Psi}_{ikl}\rangle\langle\Psi_{ikl}| \otimes |1\rangle\langle 0|, \end{aligned} \quad (6)$$

with  $\tilde{\mathcal{W}}(\rho) = \tilde{W}\rho\tilde{W}^\dagger$ . The first two terms in Eq. (6) are the classical ones, corresponding to the random choice of two possible circuits with causal structures  $A \leq B$  and  $B \leq A$ . The off-diagonal terms have no classical interpretation: they represent the quantum interference between the two different causal structures. Note that the state in Eq. (6) does not depend on the particular Kraus representation chosen for the channels  $\mathcal{A}_i$  and  $\mathcal{B}_i$ : had we chosen another Kraus representation, after summation we would have obtained the same result. Equation (6) can be extended by linearity to the case of generic no-signalling channels by writing the Kraus form  $\mathcal{C}_i(\rho) = \sum_m C_{im} \rho C_{im}^\dagger$  and expanding the Kraus operators as  $C_{im} = \sum_k A_{imk} \otimes B_{imk}$ .

The availability of a network implementing the quantum superposition of causal structures can be interpreted as a new information-theoretic primitive that takes as input a query to a generic no-signalling channel  $\mathcal{C}_i$  and produces as output one query to the channel defined by  $\mathcal{C}_i^{\text{sup}}(\rho) := \rho_i^{\text{sup}}$ , where  $\rho_i^{\text{sup}}$  is the state defined in Eq. (6) and  $\rho$  is the projector on the state  $\alpha|\Psi\rangle|0\rangle + \beta|\tilde{\Psi}\rangle|1\rangle$ . We will now show that having access to this primitive can reduce by a factor of 2 the number of queries needed for the discrimination of a pair of no-signalling channels. In our example, all systems are qubits:  $A \simeq A' \simeq B \simeq B' \simeq \mathbb{C}^2$ . Channel  $\mathcal{C}_0$  consists of two von Neumann measurements on the same random basis

$$\mathcal{C}_0 := \int dU \mathcal{M}_U^{(A)} \otimes \mathcal{M}_U^{(B)}, \quad (7)$$

where  $dU$  is the normalized Haar measure on  $\text{SU}(2)$  and  $\mathcal{M}_U$  is the single-qubit channel given by  $\mathcal{M}_U(\rho) := \langle 0|U^\dagger \rho U|0\rangle U|0\rangle\langle 0|U^\dagger + \langle 1|U^\dagger \rho U|1\rangle U|1\rangle\langle 1|U^\dagger$ ,  $\{|0\rangle, |1\rangle\}$  being the computational basis. Channel  $\mathcal{C}_1$  consists of two rotations of  $\pi$  around a pair of random orthogonal axes in the Bloch sphere:

$$\mathcal{C}_1 := \int dV \mathcal{X}_V^{(A)} \otimes \mathcal{Y}_V^{(B)}, \quad (8)$$

where  $\mathcal{X}_V(\rho) := (VXV^\dagger)\rho(VXV^\dagger)$  and  $\mathcal{Y}_V(\rho) := (VYV^\dagger)\rho(VYV^\dagger)$ ,  $X$  and  $Y$  being the Pauli matrices representing rotations of  $\pi$  around the  $x$  and  $y$  axes, respectively.

Suppose that an experimenter has access to the bipartite black box and is asked to discriminate between  $\mathcal{C}_0$  or  $\mathcal{C}_1$  using a single query. The discrimination between  $\mathcal{C}_0$  and  $\mathcal{C}_1$  is equivalent to the discrimination between two product channels  $\mathcal{C}_{0,U} := \mathcal{A}_{0,U} \otimes \mathcal{B}_{0,U}$  and  $\mathcal{C}_{1,V} := \mathcal{A}_{1,V} \otimes \mathcal{B}_{1,V}$ , with  $\mathcal{A}_{0,U} \equiv \mathcal{B}_{0,U} := \mathcal{M}_U$  and  $\mathcal{A}_{1,V} := \mathcal{X}_V$  and  $\mathcal{B}_{1,V} := \mathcal{Y}_V$ , where the unitaries  $U$  and  $V$  are completely unknown. To achieve perfect discrimination between  $\mathcal{C}_{0,U}$  and  $\mathcal{C}_{1,V}$  one can take a quantum

superposition of the following two circuits:

$$\begin{aligned}
 (a) \quad & \langle \varphi | \text{---} \boxed{\mathcal{A}_{i,U_i}} \text{---} \boxed{\mathcal{B}_{i,U_i}} \text{---} \\
 (b) \quad & \langle \varphi | \text{---} \boxed{\mathcal{B}_{i,U_i}} \text{---} \boxed{\mathcal{A}_{i,U_i}} \text{---}
 \end{aligned} \tag{9}$$

where  $\varphi$  is a fixed pure state and  $U_0 := U$  and  $U_1 := V$ . The key idea is that the Kraus operators of  $\mathcal{C}_0$  and  $\mathcal{C}_1$  behave very differently when we switch the ordering from  $A \leq B$  to  $B \leq A$ : the Kraus operators of  $\mathcal{A}_{0,U}$  and  $\mathcal{B}_{0,U}$  commute for every  $U$  (they are projectors on the same basis vectors), whereas the Kraus operators of  $\mathcal{A}_{1,V}$  and  $\mathcal{B}_{1,V}$  anticommute for every  $V$ . The difference between commutation and anticommutation cannot be detected by any ordinary circuit using a single query to the black boxes, but becomes visible in the interference terms when we superpose the two circuits (a) and (b) with amplitudes  $\alpha = \beta = \frac{1}{\sqrt{2}}$ . Using Eqs. (4)–(6) with  $R \simeq R' \simeq \tilde{R} \simeq \tilde{R}' \simeq \mathbb{C}$ ,  $\Psi = \tilde{\Psi} = \varphi$ , and  $W = \tilde{W} = I$ , we obtain the output states

$$\begin{aligned}
 \rho_0^{\text{sup}} &= \mathcal{M}_U(|\varphi\rangle\langle\varphi|) \otimes |+\rangle\langle+|, \\
 \rho_1^{\text{sup}} &= \mathcal{Z}_V(|\varphi\rangle\langle\varphi|) \otimes |-\rangle\langle-|,
 \end{aligned}$$

where  $\mathcal{Z}_V$  is the unitary channel  $\mathcal{Z}_V(\rho) := (VZV^\dagger)\rho(VZV^\dagger)$ ,  $Z$  being the Pauli matrix  $Z := -iXY$ , and  $|\pm\rangle := (|0\rangle \pm |1\rangle)/\sqrt{2}$ . By measuring the control qubit on the basis  $|+\rangle, |-\rangle$  the experimenter can perfectly distinguish between  $\mathcal{C}_0$  and  $\mathcal{C}_1$ , *no matter what the unknown unitaries  $U$  and  $V$  are*. More generally, the above scheme allows one to distinguish an arbitrary pair  $(\mathcal{A}_0, \mathcal{B}_0)$  of channels with commuting Kraus operators  $\mathcal{A}_{0i} \mathcal{B}_{0j} = \mathcal{B}_{0j} \mathcal{A}_{0i} \forall i, j$  from an arbitrary pair  $(\mathcal{A}_1, \mathcal{B}_1)$  of channels with anticommuting Kraus operators  $\mathcal{A}_{1i} \mathcal{B}_{1j} = -\mathcal{B}_{1j} \mathcal{A}_{1i} \forall i, j$ .

We now show that no quantum circuit with a fixed causal structure can perfectly distinguish between  $\mathcal{C}_0$  and  $\mathcal{C}_1$  with a single query. The proof requires the formalism of *quantum combs* [22,23] (see also the framework of *quantum strategies* by Gutoski and Watrous [24]), which describes the most general sequential strategies. This formalism makes extensive use of the Choi isomorphism [25] between a channel  $\mathcal{C}$  transforming states on  $\mathcal{H}$  and the positive operator  $C$  on  $\mathcal{H} \otimes \mathcal{H}$  defined by  $C := (\mathcal{C} \otimes \mathcal{I})(|I\rangle\langle I|)$ , where  $\mathcal{I}$  is the identity map and  $|I\rangle$  is the maximally entangled vector  $|I\rangle := \sum_n |n\rangle|n\rangle \in \mathcal{H} \otimes \mathcal{H}$ ,  $\{|n\rangle\}$  being a fixed orthonormal basis for  $\mathcal{H}$ . In general, we will use the ‘‘double ket’’ notation  $|\Psi\rangle\rangle := (\Psi \otimes I)|I\rangle\rangle$ , where  $\Psi$  is any operator on  $\mathcal{H}$ . Defining  $\mathcal{H}_1 := A$ ,  $\mathcal{H}_2 := A'$ ,  $\mathcal{H}_3 := B$ ,  $\mathcal{H}_4 = B'$ , the Choi operator of the channel  $\mathcal{C}_i$ ,  $i = 0, 1$  is the operator on  $\mathcal{H}_4 \otimes \mathcal{H}_3 \otimes \mathcal{H}_2 \otimes \mathcal{H}_1$  given by

$$C_i := \int dU (\mathcal{U} \otimes \mathcal{U}^* \otimes \mathcal{U} \otimes \mathcal{U}^*)(\Lambda_i), \tag{10}$$

where  $\Lambda_0 = \sum_{m,n=0,1} |m\rangle\langle m| \otimes |m\rangle\langle m| \otimes |n\rangle\langle n| \otimes |n\rangle\langle n|$ ,  $\Lambda_1 = |Y\rangle\rangle\langle\langle Y| \otimes |X\rangle\rangle\langle\langle X|$ , and  $\mathcal{U}$  ( $\mathcal{U}^*$ ) is the unitary channel defined by  $\mathcal{U}(\rho) := U\rho U^\dagger$  ( $\mathcal{U}^*(\rho) := U^*\rho U^T$ ),  $U^*$  ( $U^T$ ) denoting the complex conjugate (the transpose) of the matrix  $U$ .

Let us consider discrimination strategies with causal structure  $A \leq B$ . The discrimination is represented by a binary *quantum tester*  $\{T_0, T_1\}$ , consisting of two positive operators

$T_0$  and  $T_1$  on  $\mathcal{H}_4 \otimes \mathcal{H}_3 \otimes \mathcal{H}_2 \otimes \mathcal{H}_1$  that give the probabilities of the measurement outcomes according to the generalized Born rule  $p(i|C_j) = \text{Tr}[T_i C_j]$  [7]. The normalization of the tester is given by the condition  $T_0 + T_1 = I_4 \otimes \Xi$ , where  $\Xi$  is a positive operator on  $\mathcal{H}_3 \otimes \mathcal{H}_2 \otimes \mathcal{H}_1$  satisfying the relation

$$\text{Tr}_3[\Xi] = I_2 \otimes \rho, \quad \text{Tr}[\rho] = 1, \tag{11}$$

$\text{Tr}_3$  denoting the partial trace over  $\mathcal{H}_3$  and  $\rho$  being a quantum state on  $\mathcal{H}_1$  (see Ref. [7] for more details). Now, it is clear from Eq. (10) that the outcome probabilities are not affected if we replace each  $T_i$ ,  $i = 0, 1$  with its average  $T'_i := \int dU (\mathcal{U} \otimes \mathcal{U}^* \otimes \mathcal{U} \otimes \mathcal{U}^*)(T_i)$ . Since the average commutes with all the unitaries  $U \otimes U^* \otimes U \otimes U^*$ , we can assume without loss of generality the commutation relation

$$[\Xi, U^* \otimes U \otimes U^*] = 0 \quad \forall U \in \text{SU}(2). \tag{12}$$

From Ref. [7], we know that distinguishing between the two channels  $\mathcal{C}_0$  and  $\mathcal{C}_1$  with the tester  $\{T_0, T_1\}$  is equivalent to distinguishing between the two states  $\Gamma_0$  and  $\Gamma_1$  given by  $\Gamma_i := (I_4 \otimes \Xi^{\frac{1}{2}})C_i(I_4 \otimes \Xi^{\frac{1}{2}})$ ,  $i = 0, 1$ . Hence,  $\mathcal{C}_0$  and  $\mathcal{C}_1$  are perfectly distinguishable if and only if  $\Gamma_0$  and  $\Gamma_1$  have orthogonal support, that is,  $\text{Tr}[\Gamma_0 \Gamma_1] = 0$ . Note that we have  $\text{Tr}[\Gamma_0 \Gamma_1] = \text{Tr}[\tilde{C}_0(I_4 \otimes \Xi)\tilde{C}_1(I_4 \otimes \Xi)]$ , having defined  $\tilde{C}_i := (I \otimes Y \otimes I \otimes Y)C_i(I \otimes Y \otimes I \otimes Y)$  for  $i = 1, 2$ , and  $\tilde{\Xi} := (Y \otimes I \otimes Y)\Xi(Y \otimes I \otimes Y)$ .

We now prove that the condition  $\text{Tr}[\Gamma_0 \Gamma_1] = 0$  cannot be satisfied. First note that by definition  $\tilde{\Xi}$  must satisfy Eq. (11) for some density matrix  $\rho$ . Moreover, from Eq. (12) and from the relation  $U^* = YUY$ ,  $\forall U \in \text{SU}(2)$  it follows that  $\tilde{\Xi}$  must satisfy the commutation relation  $[\tilde{\Xi}, U \otimes U \otimes U] = 0$ ,  $\forall U \in \text{SU}(2)$ . Hence, by the Schur lemmas  $\tilde{\Xi}$  must be a combination of projectors onto the irreducible subspaces of  $U \otimes U \otimes U$ . The latter are easily obtained by coupling the three angular momenta: we have the subspace  $\mathcal{L}_{\frac{3}{2}}$  corresponding to  $j = \frac{3}{2}$  and two subspaces  $\mathcal{L}_{\frac{1}{2}}^{(1)}$  and  $\mathcal{L}_{\frac{1}{2}}^{(0)}$  corresponding to  $j = \frac{1}{2}$ , which are spanned by the vectors  $\{\Phi_0^{(1)}, \Phi_1^{(1)}\}$  and  $\{\Phi_0^{(0)}, \Phi_1^{(0)}\}$ , respectively,

$$\begin{aligned}
 |\Phi_0^{(1)}\rangle &:= \frac{1}{\sqrt{6}}(|0\rangle_1|Y\rangle_{2,3} + |Y\rangle_{1,3}|0\rangle_2), \\
 |\Phi_1^{(1)}\rangle &:= \frac{1}{\sqrt{6}}(|1\rangle_1|Y\rangle_{2,3} + |Y\rangle_{1,3}|1\rangle_2), \\
 |\Phi_0^{(0)}\rangle &:= \frac{|Y\rangle_{1,2}|0\rangle_3}{\sqrt{2}}, \quad |\Phi_1^{(0)}\rangle := \frac{|Y\rangle_{1,2}|1\rangle_3}{\sqrt{2}}.
 \end{aligned}$$

The most general expression for a positive operator  $\tilde{\Xi}$  commuting with  $U \otimes U \otimes U$  is then

$$\tilde{\Xi} = aP_{\frac{3}{2}} + \sum_{m,n=0,1} b_{mn}T_{\frac{1}{2}}^{mn}, \tag{13}$$

where  $a \geq 0$ ,  $P_{\frac{3}{2}}$  is the projector onto  $\mathcal{L}_{\frac{3}{2}}$ ,  $(b_{mn})$  is a positive two-by-two matrix, and  $T_{\frac{1}{2}}^{mn} := \sum_{k=0,1} |\Phi_k^{(m)}\rangle\langle\Phi_k^{(n)}|$ . Let us analyze now the normalization (11). First, note that the state  $\rho$  in Eq. (11) must be the invariant state  $\rho = I/2$  due to the symmetry  $[\tilde{\Xi}, U \otimes U \otimes U] = 0$ , which implies  $[\rho, U] = 0$ . On the other hand, taking the partial trace we

obtain  $\text{Tr}_3[P_{\frac{3}{2}}] = \frac{4}{3}P_1$ ,  $\text{Tr}_3[T_{\frac{1}{2}}^{11}] = \frac{2}{3}P_1$ ,  $\text{Tr}_3[T_{\frac{1}{2}}^{00}] = 2P_0$ , and  $\text{Tr}_3[T_{\frac{1}{2}}^{01}] = \text{Tr}[T_{\frac{1}{2}}^{10}] = 0$ .

Hence, Eq. (11) with  $\rho = I/2$  implies  $(\frac{4a+2b_{11}}{3})P_1 + 2b_{00}P_0 = \frac{I \otimes I}{2}$ , which is equivalent to the relations  $2a + b_{11} = \frac{3}{4}$  and  $b_{00} = \frac{1}{4}$ . On the other hand, direct calculation shows that the overlap  $\text{Tr}[\Gamma_0\Gamma_1]$  is zero if and only if  $a = b_{11} = b_{01} = 0$  [26]. Since this condition is incompatible with the normalization condition  $2a + b_{11} = \frac{3}{4}$ , we proved that no circuit with causal structure  $A \preceq B$  can perfectly discriminate between  $C_0$  and  $C_1$  with a single query. Moreover, since the Choi operators  $C_0$  and  $C_1$  are invariant under the exchange  $(A, A') \leftrightarrow (B, B')$ , the same derivation can be used to prove that perfect discrimination cannot be achieved with a single query by any circuit with causal structure  $B \preceq A$ . In conclusion, perfect discrimination in a circuit with fixed causal structure requires at least two queries. This number is actually sufficient, because with two queries the quantum superposition of causal structures can be simulated in an ordinary circuit using controlled swap operations [18].

Iterating the result for  $N$  different pairs of no-signalling channels  $\{C_{i_n}^{(n)} | i_n \in \{0, 1\}, n = 1, \dots, N\}$  it is easy to see that superposing two causal structures for each pair allows us to distinguish with probability 1 among  $2^N$  channels using one query to the black box  $\bigotimes_{n=1}^N C_{i_n}^{(n)}$ . Without the superposition of causal structures, the probability of successfully distinguishing all pairs of channels with a single query would go to zero exponentially fast in  $N$ . This fact can be proven using the product rule of Ref. [27], which shows that the maximum probability  $p_{\text{succ}}^{(N)}$  of distinguishing correctly all the  $N$  channels is equal to the product of the probabilities of distinguishing each channel separately, that is,  $p_{\text{succ}}^{(N)} = [p_{\text{succ}}^{(1)}]^N \rightarrow 0$ .

Before concluding, it is worth highlighting a remarkable feature of our result: perfect discrimination is achieved by superposing two strategies that, considered separately, are very inefficient. It is easy to show that the probability of success

of the strategies (a) and (b) in Eq. (9) is  $p_{\text{succ}}^{(a)} = p_{\text{succ}}^{(b)} = \frac{2}{3}$ , a value that can be easily beaten even by parallel strategies. For example, applying the unknown channel  $C_i$ ,  $i = 0, 1$  on one side of a maximally entangled state, thus obtaining the Choi state  $\rho_i^{\text{Choi}} = C_i/4$ ,  $i = 0, 1$ , yields the much higher success probability  $p_{\text{succ}}^{\text{Choi}} = \frac{11}{12}$ . Quite paradoxically, it is exactly by superposing two suboptimal strategies that one can achieve perfect discrimination. This feature suggests an analogy with Parrondo's paradox in classical game theory [21], where the alternate choice of two losing games yields a winning game (i.e., a game where the optimal strategy yields a winning probability larger than  $1/2$ ). In the quantum example the counterintuitive feature is even more striking: the probability of winning the discrimination game jumps to  $p_{\text{succ}} = 1$  thanks to the quantum superposition of the losing strategies (a) and (b).

In conclusion, we demonstrated that the quantum superposition of causal structures is a new primitive that offers an advantage over causally ordered quantum circuits in the problem of quantum channel discrimination. Such a result is similar in spirit to that of Oreshkov, Costa, and Brukner [20], who showed the advantage of noncausal strategies in a nonlocal (Bell-inequality-type) game. These results, along with the quantum switch of Ref. [18], are starting to unveil some deep relationship between quantum theory, causal order, and space-time, and more developments in this direction are expected to come in the near future.

I acknowledge G. M. D'Ariano, P. Perinotti, J. Watrous, and M. Piani for stimulating discussions and two anonymous referees for useful comments. This work has been supported by the National Basic Research Program of China (973) 2011CBA00300 (2011CBA00301), and, in its initial stage, from Perimeter Institute for Theoretical Physics. Research at Perimeter Institute for Theoretical Physics is supported in part by the Government of Canada through NSERC and by the Province of Ontario through MRI.

- 
- [1] A. Acín, *Phys. Rev. Lett.* **87**, 177901 (2001).  
 [2] G. M. D'Ariano, P. Lo Presti, and M. G. A. Paris, *Phys. Rev. Lett.* **87**, 270404 (2001).  
 [3] M. F. Sacchi, *Phys. Rev. A* **71**, 062340 (2005).  
 [4] M. F. Sacchi, *Phys. Rev. A* **72**, 014305 (2005).  
 [5] R. Duan, Y. Feng, and M. Ying, *Phys. Rev. Lett.* **98**, 100503 (2007).  
 [6] R. Duan, Y. Feng, and M. Ying, *Phys. Rev. Lett.* **100**, 020503 (2008).  
 [7] G. Chiribella, G. M. D'Ariano, and P. Perinotti, *Phys. Rev. Lett.* **101**, 180501 (2008).  
 [8] S. Lloyd, *Science* **321**, 1463 (2008).  
 [9] S. H. Tan, B. I. Erkmen, V. Giovannetti, S. Guha, S. Lloyd, L. Maccone, S. Pirandola, and J. H. Shapiro, *Phys. Rev. Lett.* **101**, 253601 (2008).  
 [10] J. H. Shapiro and S. Lloyd, *New J. Phys.* **11**, 063045 (2009).  
 [11] R. Duan, Y. Feng, and M. Ying, *Phys. Rev. Lett.* **103**, 210501 (2009).  
 [12] A. W. Harrow, A. Hassidim, D. W. Leung, and J. Watrous, *Phys. Rev. A* **81**, 032339 (2010).  
 [13] T. Eggeling, D. Schlingemann, and R. F. Werner, *Europhys. Lett.* **57**, 782 (2002).  
 [14] The restriction to pure states and isometries can be done without loss of generality thanks to the purification property of quantum theory.  
 [15] M. Piani, M. Horodecki, P. Horodecki, and R. Horodecki, *Phys. Rev. A* **74**, 012305 (2006).  
 [16] G. M. D'Ariano, S. Facchini, and P. Perinotti, *Phys. Rev. Lett.* **106**, 010501 (2011).  
 [17] Note that no-signalling is defined as an information-theoretic property of the black box. Here we do not assume that Alice and Bob are spacelike separated. What we assume instead is that the no-signalling box is able to provide Alice with an output as soon as she supplies an input, independently of Bob's actions, and vice-versa. This assumption is very natural in the case of product channels, of the form  $\mathcal{A} \otimes \mathcal{B}$ , and of randomizations of product channels, of the form  $\sum_k p_k \mathcal{A}_k \otimes \mathcal{B}_k$  for some probabilities  $\{p_k\}$ .  
 [18] G. Chiribella, G. M. D'Ariano, and P. Perinotti, [arXiv:0912.0195](https://arxiv.org/abs/0912.0195).



- [19] L. Hardy, in *Quantum Reality, Relativistic Causality, and Closing the Epistemic Circle: Essays in Honour of Abner Shimony*, edited by W. C. Myrvold and J. Christian (Springer, New York, 2009).
- [20] O. Oreshkov, F. Costa, and C. Brukner, *Nat. Commun.* **3**, 1092 (2012).
- [21] G. P. Harmer and D. Abbott, *Nature* **402**, 864 (1999).
- [22] G. Chiribella, G. M. D'Ariano, and P. Perinotti, *Phys. Rev. Lett.* **101**, 060401 (2008).
- [23] G. Chiribella, G. M. D'Ariano, and P. Perinotti, *Phys. Rev. A* **80**, 022339 (2009).
- [24] G. Gutoski and J. Watrous, *Proceedings of the 39th ACM Symposium on Theory of Computing (STOC)* (ACM Press, New York, NY, 2007), pp. 565–574.
- [25] M.-D. Choi, *Lin. Alg. Appl.* **10**, 285 (1975).
- [26] See Supplemental Material at <http://link.aps.org/supplemental/10.1103/PhysRevA.86.040301> for the detailed proof; see also G. Chiribella, [arXiv:1109.5154](https://arxiv.org/abs/1109.5154).
- [27] G. Chiribella, [arXiv:1207.6172](https://arxiv.org/abs/1207.6172).