Information-theoretic constraints on correlations with indefinite causal order

Issam Ibnouhsein
CEA-Saclay/IRFU/LARSIM, 91191 Gif-sur-Yvette, France and
Université Paris-Sud, 91405 Orsay, France
Alexei Grinbaum
CEA-Saclay/IRFU/LARSIM, 91191 Gif-sur-Yvette, France

(Dated: August 27, 2014)

Reconstructions of quantum theory usually implicitly assume that experimental events are ordered within a global causal structure. Using a generalization of the ‘process matrix’ framework, which accommodates quantum correlations that violate an inequality verified by all causally ordered correlations, we propose three principles constraining bipartite correlations to the quantum bound. We find a promising avenue for an information-theoretic reconstruction of causal structures in quantum theory by choosing a measure of dependence different from mutual information.

A physical theory is a coherent set of mathematical rules that correlate data recorded in experiments. Quantum theory is one such set of rules, however its different interpretations have produced no consensus on what these rules say about “reality”. A different approach to understanding quantum theory is to modify some of its rules and compare predictions of the modified theory with those of the original. Previous attempts include quaternionic models [1] or a model with non-linear terms in the Schrödinger equation [2]. More recently, quantum information has triggered a new development: instead of modifying the set of mathematical rules of quantum theory, one tries to derive (a subset of) these rules from clear informational principles. Reconstructing quantum theory then means that one should look for clearly motivated constraints on the correlations between experimental records, such that they (partially) reproduce the predictions of the quantum formalism [3].

For example, general non-signalling models have numerous properties in common with quantum theory, such as no-cloning [4, 5], no-broadcasting [6], monogamy of correlations [7], and information-disturbance trade-offs [7, 8]. Nonetheless, some supra-quantum models have powerful communication [9, 10] or non-local computation properties [11] unobserved in nature. The set of quantum correlations is then partially derived from various principles such as relaxed uncertainty relations [12, 13], non-locality swapping [14, 15] macroscopic locality [16], or information causality [17, 18].

Various reconstructions of quantum theory [19–24] assume, most often implicitly, that experimental events are ordered within a global causal structure. For example, Hardy proposed in [20] a reconstruction using as primitives the preparation, the transformation, and the measurement (PTM). Physical systems are defined in his reconstruction by two numbers: the number of degrees of freedom $K$, representing the minimum number of measurements to determine the state of the system, and the dimension $N$, corresponding to the maximum number of states perfectly distinguishable in one measurement of the system.

The assumption of a global causal structure is encoded in how systems compose. Indeed, consider a composite system with subsystems $A$ and $B$. Hardy’s fourth axiom expresses the operationally defined parameters $K_{AB}$ and $N_{AB}$ of the composite system in terms of the parameters of subsystems $A$ and $B$:

$$N_{AB} = N_A N_B, \quad K_{AB} = K_A K_B.$$  

This definition implies that only a super-observer can calculate $K_{AB}$ and $N_{AB}$, for it requires PTM on each subsystem by the same observer, even if $A$ and $B$ are not localized in the same laboratory. This in turn implies the existence of a global structure ordering PTM events that occur in the frame of the super-observer. To take another example, Rovelli argued informally that quantumness follows from a limit on the amount of “relevant” information that can be extracted from a system [25]. If the notion of relevance is to be connected to lattice orthomodularity in the quantum logical framework [26], the ensuing reconstruction of quantum theory will fundamentally depend on the order of binary questions asked to the system. For many systems, it requires the existence of a global causal structure ordering all incoming information.

The first step toward an information-theoretic formulation of the causal constraints on events is to build an operational framework that does not assume that events occur within a global causal structure. Efforts in this direction were initiated by Hardy [27, 28], followed by Chiribella et al. [29] and Oreshkov et al. [30]. We begin by presenting the latter framework in Section 3. A generalized notion of quantum state, called ‘process matrix’, describes all possible correlations between two physical systems under the assumption that quantum theory is valid in local laboratories, but without assuming that these laboratories are embedded in a global causal structure. Certain correlations allowed by this framework violate a ‘causal inequality’ verified by all correlations between causally ordered events. The value of the bound on such correlations, which we call ‘quantum bound’, was shown...
to be maximal for qubits and under a restricted set of lo-
cal operations involving traceless binary observables [31].
In Section [11] we introduce a class of causal games such
that any protocol defined within a global causal struc-
ture will only perform with a bounded efficiency. Taken
as an assumption in the general probabilistic framework,
this condition excludes supra-quantum correlations and
leads to a derivation of the quantum bound on correla-
tions with indefinite causal order. In Section [11] we pro-
tose two alternative informational principles based on
distinct measures of dependence, allowing to distinguish
between supra-quantum, causally ordered, and quantum
correlations with indefinite causal order. These results
further contribute to understanding the causal structure
of quantum theory via information-theoretic principles.

I. THE ‘PROCESS MATRIX’ FRAMEWORK

Consider a fixed number of laboratories equipped with
random bit generators and observers capable of free
choice. At each run of the experiment, each laboratory
receives exactly one physical system, performs transfor-
amations allowed by quantum theory and subsequently
sends the system out. Suppose each laboratory is iso-
lated from the rest of the world, except when it receives or
emits the system.

Framework. Denote the input and the output Hilbert
spaces of Alice by $H_A^1$ and $H_A^2$ and those of Bob by $H_B^1$
and $H_B^2$. We assume that these Hilbert spaces have the
same dimension.

The sets of all possible outcomes of a quantum in-
strument at Alice’s, respectively Bob’s, laboratory cor-
responds to the set of completely positive (CP) maps
$\{M_{1i,j}^{A_1A_2}\}^n_{i,j=1}$, respectively $\{M_{1i,j}^{B_1B_2}\}^n_{i,j=1}$. Using the Choi-
Jamiołkowski isomorphism, we can express a CP map
$M_{1i,j}^{A_1A_2} : \mathcal{L}(H_A^1) \rightarrow \mathcal{L}(H_A^2)$ at Alice’s laboratory
as a positive semi-definite operator $M_{1i,j}^{A_1A_2}$ acting on
$H_A^1 \otimes H_A^2$, and a CP map $M_{1i,j}^{B_1B_2} : \mathcal{L}(H_B^1) \rightarrow \mathcal{L}(H_B^2)$
at Bob’s laboratory as a positive semi-definite operator
$M_{1i,j}^{B_1B_2}$ acting on $H_B^1 \otimes H_B^2$. Using this correspondence,
the non-contextual probability for two measurement out-
comes can be expressed as a bilinear function of the cor-
responding Choi-Jamiołkowski operators:

$$P(M_{1i,j}^{A_1A_2}, M_{1i,j}^{B_1B_2}) = \text{Tr} \left[ W_{A_1A_2B_1B_2} \left( M_{1i,j}^{A_1A_2} \otimes M_{1i,j}^{B_1B_2} \right) \right],$$

where $W_{A_1A_2B_1B_2} \in \mathcal{L}(H_A^1 \otimes H_A^2 \otimes H_B^1 \otimes H_B^2)$ is fixed
for all runs of the experiment. Requiring that such prob-
abilities be non-negative for any choice of CP maps, and
equal to 1 for any choice of CPTP maps, yields a space
of valid $W$ operators referred to as process matrices.

Causal game. In this framework, two parties, Alice and
Bob, each receive a system in their laboratory. Each of
them tosses a coin, whose value is denoted by $a$ for Alice
and $b$ for Bob. They additionally share a random task
bit $b'$ with the following meaning: if $b' = 0$, Bob must
communicate $b$ to Alice; if $b' = 1$, Bob must guess the
value of $a$. Both parties always produce a guess, denoted
by $x$ for Alice and $y$ for Bob. It is crucial to assume that
the bits $a$, $b$, and $b'$ are random.

The goal of Alice and Bob is to maximize the proba-

$$P_{\text{success}} = \frac{1}{2} [p(x = b|b' = 0) + p(y = a|b' = 1)],$$

i.e. Alice should guess Bob’s toss, or vice versa, depend-
ing on the value of $b'$. If all events occur in a causal se-
nce, then

$$P_{\text{success}} \leq \frac{3}{4}. \quad (1)$$

Indeed, it is true that either Alice cannot signal to Bob
or Bob cannot signal to Alice. Consider the latter case.
If $b' = 1$, Alice and Bob could in principle achieve up to
$P(y = a|b' = 1) = 1$. However, if $b' = 0$, Alice can only
make a random guess, hence $P(x = b|b' = 0) = \frac{1}{2}$ and
the probability of success in this case satisfies (1).
The same argument shows that the probability of success will
not increase when Alice cannot signal to Bob or under
any mixing strategy.

Now consider the following process matrix using the
usual Pauli matrices $\sigma_x, \sigma_y$ and $\sigma_z$:

$$W_{A_1A_2B_1B_2} = \frac{1}{4} \left[ I_{A_1A_2B_1B_2} + \frac{1}{\sqrt{2}} (\sigma_x^{A_1} \sigma_z^{A_2} + \sigma_y^{A_1} \sigma_z^{A_2} + \sigma_z^{A_1} \sigma_z^{A_2}) \right],$$

where $A_1, A_2, B_1,$ and $B_2$ are two-level systems. Con-
ider the following CP maps at Alice’s and Bob’s labs
respectively:

$$\xi_{A_1A_2}(x, a, b') = \frac{1}{2} [ I + (-1)^a \sigma_x^{A_1} \otimes I + (-1)^b \sigma_x^{A_2}],$$

$$\eta_{B_1B_2}(y, b, b') = b' \cdot \eta_{B_1B_2}(y, b, b') + (b' \otimes 1) \cdot \eta_{B_1B_2}(y, b, b'),$$

where $\eta_{B_1B_2}(y, b, b') = \frac{1}{2} [I + (-1)^y \sigma_x^{B_1} \otimes I_{B_2} + \sigma_y^{B_1} \otimes (-1)^{b} \sigma_y^{B_2}].$ Computa-
tions show that the success probability associated to (2)
and (3) violates causal inequality (1):

$$P_{\text{success}} = \frac{2 + \sqrt{2}}{4} > \frac{3}{4}. \quad (4)$$

Hence it is impossible to interpret these events as occur-
ing within a global causal structure. This is an example
of a causally non-separable process, viz. a process that
cannot be written as (a mixture of) causal processes:

$$W \neq \lambda W_{A_1A_2} + (1 - \lambda) W_{B_1B_2},$$

where $0 \leq \lambda \leq 1$, $W_{A_1A_2}$ is a process in which Alice
cannot signal to Bob and $W_{B_1B_2}$ a process in which Bob
cannot signal to Alice. “Cannot signal” here means ei-
ther that the channels go in the other direction or that
parties share a bipartite state. If a process matrix $W$
can be written in the form (5), it will be called causally
separable.
II. A NEW CLASS OF CAUSAL GAMES

Generalized probabilistic framework. Suppose the task bit $b'$, which might not be random in the most general situation, is shared between Alice and Bob, and consider a box verifying a condition we call ‘causal realism’: $P(x, y|a, b)$ results from a convex mixture of causal orders. Formally:

$$P(x, y|a, b) = p(A \leq B)p(b'|A \leq B)\sum_{b''} p(x|a, b'', A \leq B)p(y|a, b, b', A \leq B) + p(A \nleq B)p(b'|A \nleq B)\sum_{b''} p(y|b, b', A \nleq B)p(x|a, b, b', A \nleq B).$$

This expresses the fact that causal order is predefined, even if unknown with certainty. Causal inequality (1) then plays a role similar to the violation of Bell inequalities invalidating the ‘local realism’ hypothesis [32, 33]. One can easily check that the probability distribution:

$$P(x, y|a, b, b') = \frac{1}{2} [b' \cdot \delta_{y=a} + (b' \oplus 1) \cdot \delta_{x=b} ] \quad (6)$$

is causally realistic (CR) and verifies $P_{\text{success}} = 1$.

In the process matrix framework, it has been proved that events within each laboratory are necessarily causally ordered [34]. This statement becomes an assumption in the generalized probabilistic framework (Fig.1).

In the process matrix framework, causal realism is verified by the bipartite probability distributions defined through process matrices and local CP maps. Indeed, assuming that $b'$ is given after $b$ but before $y$, one can show that, for each value of $b'$, all possible correlations are equivalent to the correlations obtained by ‘classical’ local operations, hence are causally separable. First, note that for each value of $b'$, the most general strategy for Bob is to apply a fixed quantum instrument on the input system, whose outcome yields $y$, and to subject the output system of that instrument to a subsequent CPTP map dependent on the value of $b$. The first quantum instrument can be implemented by a fixed unitary on the input system plus an ancilla, followed by a projective measurement on part of the resulting joint system. Second, note that a CPTP map dependent on the value of $b$ can be implemented by a fixed unitary applied on the output of the first quantum instrument, an ancilla, and a qubit prepared in the state $|b\rangle$ (we feed $b$ in the form of a quantum state $|b\rangle$, where different vectors $|b\rangle$ are orthogonal). These assumptions can be used to fix the probability $p(|b')$ for each value of $b'$.

Third, consider an equivalent process describing the same scenario, with Bob performing in his laboratory respectively. The arrows at Alice’s and Bob’s laboratories show the necessary causal ordering of events within each laboratory.

**Our causal game.** Consider two runs of the experiment described in the causal game in section II with bits $\{x_1, a_1, y_1, b_1\}$ and $\{x_2, a_2, y_2, b_2\}$ respectively. The random task bit $b'$ now corresponds to a pair of bits $b'_1, b'_2$ denoting the four possible combinations of tasks for two runs of the experiment: $b' = 0_10_2$ means that in both runs Alice must guess Bob’s bit, $b' = 0_11_2$ means that Alice must guess Bob’s bit in the first run and Bob must guess Alice’s bit in the second run, and so forth. It is straightforward to generalize this notation for $n$ runs.

Assume that different runs of the experiment use the same box as a resource:

$$p(b_i \oplus x_i | b'_i = 0) = p(b_j \oplus x_j | b'_j = 0), \forall i, j.$$  

One can also assume without loss of generality that:

$$p(b_i \oplus x_i | b'_i = 0) = p(a_i \oplus y_i | b'_i = 1), \forall i. \quad (7)$$

Indeed, suppose that a specific process $W^{A_1, A_2, B_1, B_2}$ and local operations $\xi^{A_1, A_2}(x, a)$ and $\eta^{B_1, B_2}(y, b, b') = b'\eta_1^{B_1}(y, b, b') + (b' \oplus 1)\eta_2^{B_2}(y, b, b')$ allow to reach an
The condition that the sum of guesses over two runs equal \( P \) to terms in \( \xi B \) constructed from \( P \) from one-run probabilities provided the number of wrong guesses is even, so that the sum over \( n \) guesses be zero. These terms are equal to:

\[
\frac{1}{2^n}(1+E)^n + \frac{1}{2^n}\sum_{k=1}^{\lfloor n/2 \rfloor} \binom{n}{k} (1-E)^{2k}(1+E)^{n-2k} = \frac{1+E^n}{2},
\]

hence \( P_{Q,n} = \frac{1+E^n}{2} \).

We now treat the two bits in \( b' \) as binary notation of a decimal number and identify \( b' \) with this number. For example, \( b' = 01 \) corresponds to 1 and \( b' = 10 \) to 2. For a given decimal \( b' = i \), we group the two rounds by specifying an expression to be set to 0, which we denote by \( g_i \oplus t_i = 0 \), where \( g_i \) is the sum of output bits (‘guesses’) and \( t_i \) the sum of input bits (‘tosses’). To continue the examples, for \( b' = 1 \) we set \( x_1 \oplus b_1 \oplus y_2 \oplus a_2 = 0 \) with the bit of guesses \( y_1 = x_1 \oplus y_2 \) and the bit of tosses \( t_1 = b_1 \oplus a_2 \). For \( b' = 2 \) the corresponding expression is \( y_1 \oplus a_1 \oplus x_2 \oplus b_2 = 0 \) with the bit of guesses \( g_2 = y_1 \oplus x_2 \) and the bit of tosses \( t_2 = a_1 \oplus b_2 \).

Lemma. The following inequality holds:

\[
\sum_{i=0}^{2^n-1} h(P(g_i \oplus t_i = 0|b' = i)) \geq 2^n - I(n), \tag{8}
\]

where \( I(n) = \sum_{i=0}^{2^n-1} I(g_i : t_i|b' = i) \) is a measure of efficiency of the \( n \) runs protocol, \( I(X : Y) \) denotes mutual information between random variables \( X \) and \( Y \), and \( h \) is the binary entropy.

Proof. We have:

\[
I(n) = \sum_{i=0}^{2^n-1} I(g_i : t_i|b' = i) = \sum_{i=0}^{2^n-1} H(g_i|b' = i) + H(t_i|b' = i) - H(g_i, t_i|b' = i),
\]

where \( H \) is Shannon entropy. Moreover:

\[
H(g_i|b' = i) - H(g_i, t_i|b' = i) = -H(t_i|g_i, b' = i),
\]

and

\[
H(t_i|g_i, b' = i) = H(t_i \oplus g_i|g_i, b' = i) \leq H(t_i \oplus t_i|b' = i) = h(P(g_i \oplus t_i = 0|b' = i)).
\]

It follows that \( I(g_i : t_i|b' = i) \geq H(t_i|b' = i) - h(P(g_i \oplus t_i = 0|b' = i)) \), hence \( I(n) \).

Proposition 1. The following inequality holds:

\[
\frac{(2E^2)^n}{2ln(2)} \leq I(n) \leq (2E^2)^n, \tag{9}
\]

where \( E = 2P_{success} - 1 \).
Indeed, in a fixed causal structure for a given value \(b = i\) the only possible finite bound on sequence \(I(n)\) is 1. Any causally separable process verifies:

\[
I(n) \leq 1, \forall n. \tag{11}
\]

This result is somewhat analogous to the principle of information causality. The latter proceeds as follows. Given a set of ‘classical’ resources (shared no-signalling correlations and one-way signalling) and a class of games \([17, 18]\), the quantum bound on correlations can be derived only by keeping an entropic figure of merit, which quantifies the performance of the parties in winning such games using these resources. Note that this similarity is only intuitive and not at all rigorous, because, in the context of no-signalling games, one can show that the principle of information causality is distinct from the ‘no-supersignalling’ principle which encodes the idea that protocol efficiency must not increase \([36]\).

### III. Constraints on Mutual Information and Beyond

We now explore the relation between the quantum bound on correlations with indefinite causal order and natural constraints on two measures of dependence (in the sense of Rényi \([37, 38]\)) between the variables of a CR-box.

**Constraints on mutual information.** The first measure of dependence we explore is mutual information. Using equation (11), one can show that the condition on mutual information:

\[
I(x : b \mid b = 0) + I(y : a \mid b = 1) \leq 1 \tag{12}
\]

is violated by certain supra-quantum CR-boxes. The meaning of this condition is that, in the context of the causal game of section I correlations should be compatible with at most one bit of shared information between parties. However, this condition is not sufficient for limiting correlations to the ones allowed by the process matrix framework. Indeed, using again equation (11), one can show that there exist supra-quantum CR-boxes obeying (12). One way to obtain only quantum correlations is to introduce multiple boxes and condition (11) on mutual information. An alternative approach is to impose a slightly stronger constraint on the behavior of mutual information for boxes.

**Proposition 2.** Consider two boxes \((E_1, x_1, y_1, a_1, b_1)\) and \((E_2, x_2, y_2, a_2, b_2)\) with parameters \(E_1 = 2p_1 - 1, E_2 = 2p_2 - 1\), where \(p_i\) are the probabilities of success in simulating box \([6]\), and associated task bits \(b_1', b_2'\). The following two conditions are equivalent:

(i):

\[
I(x_1 : b_1 \mid b_1' = 0) \geq I(x_1 \oplus x_2 : b_1 \oplus b_2 \mid b_1' = 0, b_2' = 0) + I(x_1 \oplus y_2 : b_1 \oplus a_2 \mid b_1' = 0, b_2' = 1), \tag{13}
\]

\[
I(y_1 : a_1 \mid b_1' = 1) \geq I(y_1 \oplus x_2 : a_1 \oplus b_2 \mid b_1' = 1, b_2' = 0) + I(y_1 \oplus y_2 : a_1 \oplus a_2 \mid b_1' = 1, b_2' = 1), \tag{14}
\]

(ii):

\[
P_{Q,1} = \frac{1}{2} [p(x_2 \oplus b_2 | b_2' = 0) + p(y_2 \oplus a_2 | b_2' = 1)] \leq \frac{2 + \sqrt{2}}{4}. \tag{15}
\]

**Proof.** Suppose that \(P_{Q,1} = \frac{1 + E_2}{2} \leq \frac{2 + \sqrt{2}}{4}\). Assumption \([7]\) implies that the two terms on the right-hand side in equations (13) and (14) are equal. Focusing on (13), one only needs to show that:

\[
I(x_1 \oplus x_2 : b_1 \oplus b_2 | b_1' = 0, b_2' = 0) \leq \frac{1}{2} I(x_1 : b_1 | b_1' = 0). \tag{16}
\]

Define the variables \(X = b_1 | b_1' = 0, Y = x_1 | b_1' = 0\) and \(Z = x_1 \oplus x_2 \oplus b_2 | b_1' = 0, b_2' = 0\), where the entire expression on the left-hand side of the bar is conditioned by the
Since \(Y, Z\) maximal correlation of variables \(Y\) and \(Z\), we want to explain the quantum bound on correlations with indefinite causal orders through a condition on mutual information. If we wish to explain the quantum bound on correlations with indefinite causal orders, we can introduce a class of causal games in which causally separable processes perform with bounded efficiency as measured by mutual information. Using bounded efficiency as a condition, we derived the quantum bound on correlations with indefinite causal order. The quantum bound can also be derived from a DPI-like constraint on the behavior of CR-boxes. Using an alternative measure of dependence, we established a relation between the bound on causally separable processes and a constraint on the total amount of communication between parties. Another relation was found between the quantum bound and the initial efficiency of communication. Central to these derivations were standard properties of mutual information. The most interesting finding is that our approach highlights both qualitatively and quantitatively the fact that mutual information is not the most convenient measure of dependence in causal games. Whether “natural” properties of alternative measures, e.g., HGR maximal correlation, lead to the quantum bound is curiously an informational consistency condition for classical systems, so that mutual information between independent systems equal 0, along with one of conditions (11) or (17).

IV. CONCLUSION

We introduced a class of causal games in which causally separable processes perform with bounded efficiency as measured by mutual information. Using bounded efficiency as a condition, we derived the quantum bound on correlations with indefinite causal order. The quantum bound can also be derived from a DPI-like constraint on the behavior of CR-boxes. Using an alternative measure of dependence, we established a relation between the bound on causally separable processes and a constraint on the total amount of communication between parties. Another relation was found between the quantum bound and the initial efficiency of communication. Central to these derivations were standard properties of mutual information. The most interesting finding is that our approach highlights both qualitatively and quantitatively the fact that mutual information is not the most convenient measure of dependence in causal games. Whether “natural” properties of alternative measures, e.g., HGR maximal correlation, lead to the quantum bound is currently under investigation.

ACKNOWLEDGMENTS

We thank O. Oreshkov, F. Costa, M. Pawłowski and J. Baumesler for helpful discussions. This work was supported by the European Commission Project Q-ESSENCE (No.248095).
