LETTERS

Information causality as a physical principle

Marcin Pawłowski[1], Tomasz Paterek[2], Dagomir Kaszlikowski[2], Valerio Scarani[2], Andreas Winter[2,3] & Marek Żukowski[1]

Quantum physics has remarkable distinguishing characteristics. For example, it gives only probabilistic predictions (non-determinism) and does not allow copying of unknown states (no-cloning[1]). Quantum correlations may be stronger than any classical ones[2], but information cannot be transmitted faster than light (no-signalling). However, these features do not uniquely define quantum physics. A broad class of theories exist that share such traits and allow even stronger (than quantum) correlations[3]. Here we introduce the principle of 'information causality' and show that it is respected by classical and quantum physics but violated by all no-signalling theories with stronger than (the strongest) quantum correlations. The principle relates to the amount of information that an observer (Bob) can gain about a data set belonging to another observer (Alice), the contents of which are completely unknown to him. Using all his local resources (which may be correlated with her resources) and allowing classical communication from her, the amount of information that Bob can recover is bounded by the information volume (m) of the communication. Namely, if Alice communicates m bits to Bob, the total information obtainable by Bob cannot be greater than m. For $m = 0$, information causality reduces to the standard no-signalling principle. However, no-signalling theories with maximally strong correlations would allow Bob access to all the data in any m-bit subset of the whole data set held by Alice. If only one bit is sent by Alice ($m = 1$), this is tantamount to Bob's being able to access the value of any single bit of Alice's data (but not all of them). Information causality may therefore help to distinguish physical theories from non-physical ones. We suggest that information causality—a generalization of the no-signalling condition—might be one of the foundational properties of nature.

Classical (as opposed to quantum) physics rests on the assumption that all physical quantities have well-defined values simultaneously. Relativity is based on clear-cut physical statements: the speed of light and the electric charge are the same for all observers. In contradistinction, the definition of quantum physics is still a description of its formalism: the theory in which systems are described by Hilbert spaces and dynamics is reversible. This situation is all the more unexpected because quantum physics is the most successful physical theory and quite a lot is known about it. Some of its counterintuitive features are almost popular knowledge: all scientists, and many laymen as well, know that quantum physics predicts only probabilities, that some physical quantities (such as position and momentum) cannot be simultaneously well defined and that the act of measurement generically modifies the state of the system. Entanglement and no-cloning are rapidly claiming their place in the list of well-known quantum features; in next place are the feats of quantum information such as the possibility of secure cryptography[4,5] or the teleportation of unknown states[6].

These features are so striking that one could hope that some of them provide the physical ground behind the formalism. Is quantum physics, for instance, the most general theory that allows violations of Bell inequalities, while satisfying no-signalling? When this question was investigated[3] the answer was found to be negative: impossibility of being represented in terms of local variables is a property shared by a broad class of no-signalling theories. Such theories predict intrinsic randomness, no-cloning[7,8] and an information-disturbance trade-off[9] and permit secure cryptography[10–12]. As regards teleportation and entanglement swapping[13], after a first negative attempt[14], it seems that they can also be defined within the general no-signalling framework[15,16]. In summary, most of the features that have been highlighted as 'typically quantum' are shared by all possible no-signalling theories. Only a few discrepancies have been noticed: some no-signalling theories would lead to an implausible simplification of distributed computational tasks[17–20] and would have very limited dynamics[21]. This highlights the importance of the no-signalling principle but leaves us still uncertain about the specificity of quantum theory.

Here we define and study a previously unnoticed feature, which we call 'information causality'. Information causality generalizes no-signalling and is respected by both classical and quantum physics. However, as we shall show, it is violated by all no-signalling theories that are endowed with correlations that are stronger than the strongest quantum correlations. It can therefore be used as a principle to distinguish physical theories from non-physical ones and is a good candidate for one of the foundational assumptions that are at the very root of quantum theory.

Formulated as a principle, information causality states: ''the information gain that Bob can reach about a previously unknown to him data set of Alice, by using all his local resources and m classical bits communicated by Alice, is at most m bits''. The standard no-signalling condition is just information causality for $m = 0$. The principle assumes classical communication: if quantum bits were allowed to be transmitted, the information gain could be higher, as demonstrated in the quantum super-dense coding protocol[22]. The efficiency of this protocol is based on the use of quantum entanglement, and information causality holds true even if the quantum bits are transmitted provided that they are disentangled from the systems of the receiver. This follows from the Holevo bound, which limits information gain after transmission of m such qubits to m classical bits.

We show that in a world in which certain tasks are 'too simple' (compare with refs 17, 18) and there exists implausible accessibility of remote data, information causality is violated. Consider a generic situation in which Alice has a database of N bits described by a string $\vec{a}$. She would like to grant Bob access to as big a portion of the database as possible within a fixed amount of classical communication. If there were no pre-established correlations between them, communication of m bits would open access to at most m bits of the database. With previously shared correlations they could expect to do better (however, as we show here, in the real world they would be mistaken). For concreteness, consider a generic task illustrated in Fig. 1. It is a

[1]Institute of Theoretical Physics and Astrophysics, University of Gdańsk, 80-952 Gdańsk, Poland. [2]Centre for Quantum Technologies and Department of Physics, National University of Singapore, 3 Science Drive 2, 117543 Singapore, Singapore. [3]Department of Mathematics, University of Bristol, Bristol BS8 1TW, UK.

©2009 Macmillan Publishers Limited. All rights reserved

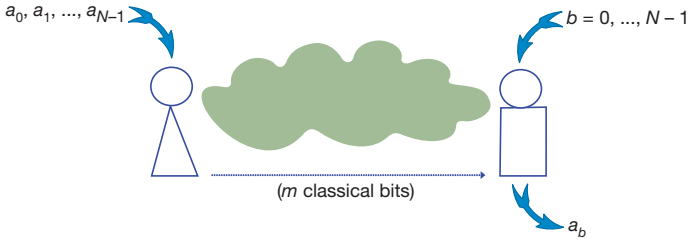

Figure 1 | The task. Alice receives N random and independent bits $\vec{a}=(a_0, \ldots, a_{N-1})$. In a separate location, Bob receives a random variable $b\in\{0, 1, \ldots, N-1\}$. Alice sends m classical bits to Bob, with the help of which Bob is asked to guess the value of the bth bit in Alice's list, a_b. Alice and Bob can share any no-signalling resources. Information causality limits the efficiency of solutions to this task. It bounds the mutual information between Alice's data and all that Bob has at hand after receiving the message.

distributed version of random access coding[23,24], oblivious transfer[14,25] and related communication complexity problems[26]. Alice receives a string of N random and independent bits, $\vec{a}=(a_0, a_1, \ldots, a_{N-1})$. Bob receives a random value of $b=0, \ldots, N-1$ and is asked to give a value of the bth bit of Alice after receiving from her a message of m classical bits. The restrictions are only on the communication that can take place after the inputs have been provided. The resources that Alice and Bob may have shared in advance are assumed to be no-signalling because allowing signalling resources would open other communication channels. In a classical world, these additional resources would be correlated lists of bits; in a quantum world, Alice and Bob may share an arbitrary quantum state. However, the task itself is open to accommodate any hypothetical resource producing no-signalling correlations, even those that go beyond the possibilities of quantum physics. We shall call these imaginary resources no-signalling boxes, or NS-boxes for short. The impact of stronger-than-quantum correlations on the efficiency of random access coding has been studied recently from a different angle[24].

There exists a protocol that allows Bob to give the correct value of at least m bits. If Alice sends him an m-bit message $\vec{x}=(a_0, \ldots, a_{m-1})$ Bob will guess a_b perfectly whenever $b\in\{0, \ldots, m-1\}$. The price to pay is that he is bound to make a completely random guess for $b\in\{m, \ldots, N-1\}$. Because the previously shared correlations contain no information about $\vec{a}$, for every strategy there will be a trade-off between the probabilities for guessing different bits of $\vec{a}$. Let us denote Bob's output by β. The efficiency of Alice's and Bob's strategy can be quantified by

$$I \equiv \sum_{K=0}^{N-1} I(a_K : \beta | b=K) \tag{1}$$

where $I(a_K:\beta|b=K)$ is the Shannon mutual information between a_K and β, computed under the condition that Bob has received $b=K$. One can also show that

$$I \geq N - \sum_{K=0}^{N-1} h(P_K) \tag{2}$$

where $h(x) = -x\log_2 x - (1-x)\log_2(1-x)$ is the binary entropy of x, and P_K is the probability that $a_K=\beta$, again for the case of $b=K$. To obtain the inequality, the a_K have been assumed to be unbiased and independently distributed (details are given in Supplementary Information).

Ideally, we wish to define that information causality holds if, after transfer of the m-bit message, the mutual information between Alice's data $\vec{a}$ and everything that Bob has—that is, the message $\vec{x}$ and his part B of the previously shared correlation—is bounded by m. Intuitively appealing though such a definition is, it has the severe issue that it is not theory-independent. Specifically, a mutual information expression '$I(\vec{a}:\vec{x}, B)$' has to be defined for a state involving objects from the underlying theory (the possibilities include classical correlation, a shared quantum state and NS-boxes). It is far from clear whether mutual information can be defined consistently for all non-local correlations, nor whether such a definition would be unique.

Instead, we shall show that if a mutual information can be defined that obeys certain elementary properties, then (a) information causality holds and (b) $I(\vec{a}:\vec{x}, B)\geq I$. Thus we obtain the following necessary condition for information causality:

$$I \leq m \tag{3}$$

We stress that the parameter I is independent of any underlying physical theory: I does not involve any details of a particular physical model but is fully determined by Alice's and Bob's input bits and Bob's output. In this sense it resembles Bell's parameter[2], which also involves only random variables and can be used to test different physical theories.

For a system composed of parts A, B and C, prepared in a state allowed by the theory, we need to assign symmetric and non-negative mutual informations $I(A:B)$, etc. The elementary properties mentioned above are the following. First, consistency: if the subsystems A and B are both classical, then $I(A:B)$ should coincide with Shannon's mutual information. Second, data-processing inequality: acting on one of the parts locally by any state transformation allowed in the theory cannot increase the mutual information. That is, if $B\rightarrow B'$ is a permissible map between systems, then $I(A:B)\geq I(A:B')$. This says that any local manipulation of data can only decay information. Third, a chain rule: there exists a conditional mutual information $I(A:B|C)$ such that the following identity is satisfied for all states and triples of parts: $I(A:B,C) = I(A:C) + I(A:B|C)$. This implies an identity between ordinary mutual informations:

$$I(A:B, C) - I(A:B) = I(A:B|C) = I(A, C:B) - I(B:C)$$

Information causality holds in both classical and quantum physics; we may focus on the latter because the former is a special case of it. This is because one can define quantum mutual information in a formal extension of Shannon's quantity, using von Neumann entropy[27], and all three of the above properties are fulfilled[28]. Details are given in Supplementary Information, but in brief one argues as follows.

To show (a), denote by B Bob's quantum system holding the shared quantum state ρ_{AB}, Alice's data $\vec{a}=(a_0, \ldots, a_{N-1})$, and the m-bit message $\vec{x}$; our objective is to prove that $I(\vec{a}:\vec{x}, B)\leq m$. First, the chain rule for mutual information yields $I(\vec{a}:\vec{x}, B) = I(\vec{a}:B) + I(\vec{a}:\vec{x}|B)$. Second, $I(\vec{a}:B)=0$ because without the message Alice's data and Bob's quantum state are independent (expressing the no-signalling condition). Third, we use the chain rule again to express the conditional mutual information as $I(\vec{a}:\vec{x}|B) = I(\vec{x}:\vec{a}, B) - I(\vec{x}:B) \leq I(\vec{x}:\vec{a}, B)$. Finally, the latter can be upper-bounded by $I(\vec{x}:\vec{x})\leq m$, invoking data processing. Similarly, (b) is obtained by repeated application of the chain rule, data-processing inequality and non-negativity of mutual information (details are given in Supplementary Information).

To study how other no-signalling theories can violate information causality, we focus on the necessary condition in equation (3). First consider the simplest example of two-bit input by Alice, (a_0, a_1); it is described in Fig. 2. The probability that Bob correctly gives the value of the bit a_0 is

$$P_{\mathrm{I}} = \frac{1}{2}[P(A\oplus B=0|0, 0) + P(A\oplus B=0|1, 0)] \tag{4}$$

and the analogous probability for the bit a_1 reads

$$P_{\mathrm{II}} = \frac{1}{2}[P(A\oplus B=0|0, 1) + P(A\oplus B=1|1, 1)] \tag{5}$$

where the symbol $\oplus$ denotes summation modulo 2.

One can recognize that these probabilities are intimately linked with the Clauser–Horne–Shimony–Holt parameter[29] S, which can be used to quantify the strength of correlations. Indeed,

©2009 Macmillan Publishers Limited. All rights reserved

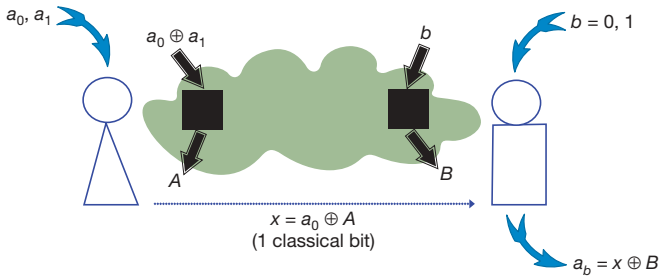

Figure 2 | Van Dam's protocol[17]. This is the simplest case in which information causality can be violated (see also ref. 25). Alice receives two bits (a_0, a_1) and is allowed to send only one bit to Bob. A convenient way of thinking about no-signalling resources is to consider paired black boxes shared between Alice and Bob (NS-boxes). The correlations between inputs $a, b = 0, 1$ and outputs $A, B = 0, 1$ of the boxes are described by probabilities $P(A \oplus B = ab|a, b)$. The no-signalling is satisfied because of uniformly random local outputs. With suitable NS-boxes Alice and Bob violate information causality. She uses $a = a_0 \oplus a_1$ as an input to the shared NS-box and obtains the outcome A, which is used to compute her message bit $x = a_0 \oplus A$ for Bob. Bob, on his side, inputs $b = 0$ if he wants to learn a_0, and $b = 1$ if he wants to learn a_1; he gets the outcome B. On receiving x from Alice, Bob computes his guess $\beta = x \oplus B = a_0 \oplus A \oplus B$. The probability that Bob correctly gives the value of the bit a_0 is $P_{\mathrm{I}} = \frac{1}{2}[P(A \oplus B = 0|0, 0) + P(A \oplus B = 0|1, 0)]$, and the analogous probability for the bit a_1 reads $P_{\mathrm{II}} = \frac{1}{2}[P(A \oplus B = 0|0, 1) + P(A \oplus B = 1|1, 1)]$, which follow by inspection of the different cases.

$$S = \sum_{a=0}^{1}\sum_{b=0}^{1} P(A \oplus B = ab|a, b) = 2(P_{\mathrm{I}} + P_{\mathrm{II}}) \qquad (6)$$

The classical correlations are bounded by $S \leq S_{\mathrm{C}} = 3$ (the equivalent form of Bell's inequality[2,29]). Quantum correlations exceed this limit up to $S \leq S_Q = 2 + \sqrt{2}$ (the so-called Tsirelson bound[30]). The maximal algebraic value of $S_{\mathrm{NS}} = 4$ is reached by the Popescu–Rohrlich (PR) box[3], which is an extremal no-signalling resource. PR-boxes maximally violate information causality because they predict $P_{\mathrm{I}} = P_{\mathrm{II}} = 1$; that is, $I = 2$ for $m = 1$, so here occurs an extreme violation of information causality. Bob can learn either bit perfectly. $I = 2$ measures the sum total of the information accessible to Bob. However, he cannot learn both of Alice's bits—the latter would imply signalling.

The protocol works just as well for any Boolean function of the inputs, $f(\vec{a}, b)$. It is sufficient that Alice inserts to her PR-box the sum $f(\vec{a}, 0) \oplus f(\vec{a}, 1)$. If information causality is maximally violated, Bob can learn the value of $f(\vec{a}, b)$ for any one of his inputs, irrespective of Alice's input data. Even more surprisingly, this is also true if he does not know the function to be computed.

We shall now demonstrate that information causality is violated as soon as the quantum Tsirelson limit for the CHSH inequality is exceeded. This result of ours can be also seen as an information-theoretic proof of the Tsirelson bound, independently of the formalism of Hilbert spaces, relying instead only on the existence of a consistent information calculus for certain correlations.

First we note that, using a suitable local randomization procedure that does not change the value of the parameter S, any NS-box can be brought to a simple form[7]: the local outcomes are uniformly random and the correlations are given by

$$P(A \oplus B = ab|a, b) = \frac{1}{2}(1 + E) \qquad (7)$$

with $0 \leq E \leq 1$. The case $E = 1$ corresponds to the PR-box; $E = 0$ describes uncorrelated random bits. The classical bound $S \leq S_{\mathrm{C}}$ is violated as soon as $E > \frac{1}{2}$; the Tsirelson bound of quantum physics becomes $E \leq E_{\mathrm{Q}} = \frac{1}{\sqrt{2}}$, attained by performing suitable measurements on the singlet state of two two-level systems[2,30].

The bound that information causality imposes on correlations can be identified by using a pyramid of NS-boxes and nesting the simple protocol described above. Now Alice receives $N = 2^n$ bits, and correspondingly Bob receives n input bits b_n, which describe the index of the bit he is interested in, $b = \sum_{k=0}^{n-1} b_k 2^k$. Alice is allowed to send a single bit, $m = 1$. An example of this protocol for $n = 2$ is presented in Fig. 3. Generally, Alice and Bob use a pyramid of $N - 1$ pairs of boxes placed on n levels. Looking at the binary decomposition of b, Bob aims $(n - r)$ times at the left bit and r times at the right, where $r = b_0 + \ldots + b_{n-1}$. His final guess is the sum of $\beta = x \oplus B_0 \oplus \ldots \oplus B_{n-1}$. Bob's final guess is therefore correct whenever he has made an even number of errors in the intermediate steps. This leads to the equation

$$P_K = \frac{1}{2}[1 + E^n] \qquad (8)$$

for the probability of his correct final guess (the details of this calculation are given in Supplementary Information).

Inserting this expression into equation (1), one finds that the information causality condition $I \leq 1$ is violated as soon as $2E^2 > 1$ and n is large enough; that is, $E > E_{\mathrm{Q}}$. Because all NS-boxes can be brought to the form in equation (7) without changing the value of S, we conclude indeed that every NS-box with stronger than quantum correlations violates the information causality condition. In Supplementary Information the more general result is proved, that for any $\frac{1}{2}(E_{\mathrm{I}}^2 + E_{\mathrm{II}}^2) > E_{\mathrm{Q}}^2$ where $E_j = 2P_j - 1$ (see equations (4) and (5)), information causality is violated, and conversely that if it is fulfilled there exists a quantum correlation with these probabilities.

Here we have identified the principle of information causality, which precisely distinguishes physically realized correlations from non-physical ones (in the sense that quantum mechanics cannot reach them). It is phrased in operational terms and in a theory-independent way; we therefore suggest that it is at the same foundational level as the no-signalling condition itself, of which it is a generalization.

The new principle is respected by all correlations accessible with quantum physics and excludes all no-signalling correlations, which violate the quantum Tsirelson bound. Among the correlations that do not violate that bound it is not known whether information

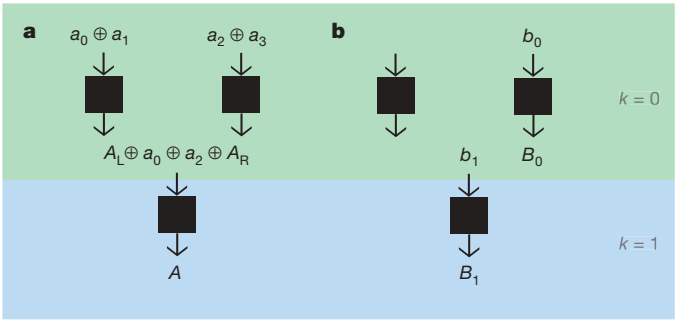

Figure 3 | Information causality identifies the strongest quantum correlations. Alice receives $N = 2^n$ input bits and correspondingly Bob receives n input bits b_n, which describe the index of the bit he is interested in, $b = \sum_{k=0}^{n-1} b_k 2^k$. Alice is allowed to send a single bit, $m = 1$. **a**, For $n = 2$, to encode information about her data Alice uses a pyramid of NS-boxes. Note that Fig. 2 shows how Bob can correctly guess the first or second bit Alice has using a single pair of the boxes (the case of $n = 1$). If Alice has more bits, then they recursively use this protocol in the following way. For example, for four input bits of Alice, two pairs of NS-boxes on the level $k = 0$ allow Bob to make the guess of a value of any one of Alice's bits as soon as he knows either $a_0 \oplus A_{\mathrm{L}}$ or $a_2 \oplus A_{\mathrm{R}}$, where A_{L} and A_{R} are the output of her left and right boxes, respectively, on the level $k = 0$, which are the one-bit messages of the protocol in Fig. 2. These can be encoded using the third box, on the level $k = 1$, by inserting their sum to Alice's box and sending $x = a_0 \oplus A_{\mathrm{L}} \oplus A$ to Bob (A is the output of her box on the level $k = 1$). Depending on the bit he is interested in, he now reads a suitable message using the box on the level $k = 1$ and uses one of the boxes on the level $k = 0$. **b**, An example of a situation in which Bob aims at the value of a_2 or a_3. Bob's final answer is $x \oplus B_0 \oplus B_1$, where B_k is the output of his box on the kth level.

©2009 Macmillan Publishers Limited. All rights reserved

causality singles out exactly those allowed by quantum physics. If it does, the new principle would acquire even stronger status.

Received 8 May; accepted 13 August 2009.

1. Wootters, W. K. & Zurek, W. H. A single quantum cannot be cloned. *Nature* **299,** 802–803 (1982).
2. Bell, J. S. On the Einstein Podolsky Rosen paradox. *Physics* **1,** 195–200 (1964).
3. Popescu, S. & Rohrlich, D. Quantum nonlocality as an axiom. *Found. Phys.* **24,** 379–385 (1994).
4. Bennett, C. H. & Brassard, G. in *Proceedings IEEE Int. Conf. on Computers, Systems and Signal Processing, Bangalore, India* 175–179 (IEEE, 1984).
5. Ekert, A. K. Quantum cryptography based on Bell's theorem. *Phys. Rev. Lett.* **67,** 661–663 (1991).
6. Bennett, C. H. *et al.* Teleporting an unknown quantum state via dual classical and Einstein–Podolsky–Rosen channels. *Phys. Rev. Lett.* **70,** 1895–1899 (1993).
7. Masanes, L., Acín, A. & Gisin, N. General properties of nonsignaling theories. *Phys. Rev. A* **73,** 012112 (2006).
8. Barnum, H., Barrett, J., Leifer, M. & Wilce, A. Generalized no-broadcasting theorem. *Phys. Rev. Lett.* **99,** 240501 (2007).
9. Scarani, V. *et al.* Secrecy extraction from no-signaling correlations. *Phys. Rev. A* **74,** 042339 (2006).
10. Barrett, J., Hardy, L. & Kent, A. No signaling and quantum key distribution. *Phys. Rev. Lett.* **95,** 010503 (2005).
11. Acín, A., Gisin, N. & Masanes, L. From Bell's theorem to secure quantum key distribution. *Phys. Rev. Lett.* **97,** 120405 (2006).
12. Masanes, L. Universally-composable privacy amplification from causality constraints. *Phys. Rev. Lett.* **102,** 140501 (2009).
13. Żukowski, M., Zeilinger, A., Horne, M. A. & Ekert, A. K. 'Event-ready-detectors' Bell experiment via entanglement swapping. *Phys. Rev. Lett.* **71,** 4287–4290 (1993).
14. Short, A. J., Popescu, S. & Gisin, N. Entanglement swapping for generalized nonlocal correlations. *Phys. Rev. A* **73,** 012101 (2006).
15. Barnum, H., Barrett, J., Leifer, M. & Wilce, A. Teleportation in general probabilistic theories. Preprint at ⟨http://arxiv.org/abs/0805.3553v1⟩ (2008).
16. Skrzypczyk, P., Brunner, N. & Popescu, S. Emergence of quantum correlations from nonlocality swapping. *Phys. Rev. Lett.* **102,** 110402 (2009).
17. Van Dam, W. *Nonlocality and Communication Complexity*. PhD thesis, Univ. Oxford (2000); Implausible consequences of superstrong nonlocality. Preprint at ⟨http://arxiv.org/abs/quant-ph/0501159v1⟩ (2005).
18. Brassard, G. *et al.* Limit on nonlocality in any world in which communication complexity is not trivial. *Phys. Rev. Lett.* **96,** 250401 (2006).
19. Linden, N., Popescu, S., Short, A. J. & Winter, A. Quantum nonlocality and beyond: limits from nonlocal computation. *Phys. Rev. Lett.* **99,** 180502 (2007).
20. Brunner, N. & Skrzypczyk, P. Non-locality distillation and post-quantum theories with trivial communication complexity. *Phys. Rev. Lett.* **102,** 160403 (2009).
21. Barrett, J. Information processing in generalized probabilistic theories. *Phys. Rev. A* **75,** 032304 (2007).
22. Bennett, C. H. & Wiesner, S. J. Communication via one- and two-particle operators on Einstein–Podolsky–Rosen states. *Phys. Rev. Lett.* **69,** 2881–2884 (1992).
23. Ambainis, A., Nayak, A., Ta-Shma, A. & Vazirani, U. Quantum dense coding and quantum finite automata. *J. ACM* **49,** 496–511 (2002).
24. Ver Steeg, G. & Wehner, S. Relaxed uncertainty relations and information processing. Preprint at ⟨http://arxiv.org/abs/0811.3771v2⟩ (2009).
25. Wolf, S. & Wullschleger, J. Oblivious transfer and quantum non-locality. Preprint at ⟨http://arxiv.org/abs/quant-ph/0502030v1⟩ (2005).
26. Brassard, G. Quantum communication complexity. *Found. Phys.* **33,** 1593–1616 (2003).
27. Cerf, N. J. & Adami, C. Negative entropy and information in quantum mechanics. *Phys. Rev. Lett.* **79,** 5194–5197 (1997).
28. Nielsen, M. A. & Chuang, I. L. *Quantum Computation and Quantum Information* (Cambridge Univ. Press, 2000).
29. Clauser, J. F., Horne, M. A., Shimony, A. & Holt, R. A. Proposed experiment to test local hidden-variable theories. *Phys. Rev. Lett.* **23,** 880–884 (1969).
30. Cirel'son, B. S. Quantum generalizations of Bell's inequality. *Lett. Math. Phys.* **4,** 93–100 (1980).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank M. Christandl, V. Vedral and S. Wehner for stimulating discussions. This work was supported by the National Research Foundation and the Ministry of Education in Singapore, and by the European Commission through the Integrated Project Qubit Applications. A.W. acknowledges support by the UK Engineering and Physical Sciences Research Council through the Quantum Information Processing Interdisciplinary Research Collaboration and an Advanced Fellowship, by a Royal Society Wolfson Merit Award, and by a Philip Leverhulme Prize.

Author Contributions All authors contributed to the initial conception of the ideas, to the working out of details, and to the writing and editing of the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to M.P. (dokmpa@univ.gda.pl).

©2009 **Macmillan Publishers Limited. All rights reserved**