Entropy **2013**, *15*, 3698-3713; doi:10.3390/e15093698

OPEN ACCESS

entropy

ISSN 1099-4300

www.mdpi.com/journal/entropy

Article

Correlation Distance and Bounds for Mutual Information

Michael J. W. Hall

Centre for Quantum Computation and Communication Technology (Australian Research Council), Centre for Quantum Dynamics, Griffith University, Brisbane, QLD 4111, Australia;
E-Mail: michael.hall@griffith.edu.au; Tel.: +61-7-3735-4333

Received: 21 June 2013; in revised form: 21 August 2013 / Accepted: 3 September 2013 / Published: 6 September 2013

Abstract: The correlation distance quantifies the statistical independence of two classical or quantum systems, via the distance from their joint state to the product of the marginal states. Tight lower bounds are given for the mutual information between pairs of two-valued classical variables and quantum qubits, in terms of the corresponding classical and quantum correlation distances. These bounds are stronger than the Pinsker inequality (and refinements thereof) for relative entropy. The classical lower bound may be used to quantify properties of statistical models that violate Bell inequalities. Partially entangled qubits can have lower mutual information than can any two-valued classical variables having the same correlation distance. The qubit correlation distance also provides a direct entanglement criterion, related to the spin covariance matrix. Connections of results with classically-correlated quantum states are briefly discussed.

Keywords: mutual information; variational distance; trace distance; Pinsker inequality; quantum entanglement

Classification: PACS 89.70.Cf;03.67.-a.

---

1. Introduction

The relative entropy between two probability distributions has many applications in classical and quantum information theory. A number of these applications, including the conditional limit theorem [1], quantum error correction [2], and secure random number generation and communication [3,4], make use

of lower bounds on the relative entropy in terms of a suitable distance between the two distributions. The best known such bound is the so-called Pinsker inequality [5]

$$H(P\|Q) := \sum_j P(j)[\log P(j) - \log Q(j)] \geq \frac{1}{2} D(P,Q)^2 \log e \tag{1}$$

where $D(P,Q) := \|P - Q\|_1 = \sum_j |P(j) - Q(j)|$ is the variational or L1 distance between distributions P and Q. Note that the choice of logarithm base is left open throughout this paper, corresponding to a choice of units. There are a number of such bounds [5], all of which easily generalise to the case of quantum probabilities [2,6,7].

However, in a number of applications of the Pinsker inequality and its quantum analog, a lower bound is in fact only needed for the special case that the relative entropy quantifies the mutual information between two systems. Such applications include, for example, secure random number generation and coding [3,4] (both classical and quantum), and quantum de Finnetti theorems [8]. Since mutual information is a special case of relative entropy, it follows that it may be possible to find strictly stronger lower bounds for mutual information.

Surprisingly little attention appears to have been paid to this possibility of better lower bounds (although upper bounds for mutual information have been investigated [9]). The results of preliminary investigations are given here, with explicit tight lower bounds being obtained for pairs of two-valued classical random variables, and for pairs of quantum qubits with maximally-mixed reduced states.

In the context of mutual information, the corresponding variational distance reduces to the distance between the joint state of the systems and the product of their marginal states, referred to here as the "correlation distance". It is shown that both the classical and quantum correlation distances are relevant for quantifying properties of quantum entanglement: the former with respect to the classical resources required to simulate entanglement, and the latter as providing a criterion for qubit entanglement. In the quantum case, it is also shown that the minimum value of the mutual information can only be achieved by entangled qubits if the correlation distance is more than ≈ 0.72654.

The main results are given in the following section. Lower bounds on classical and quantum mutual information for two-level systems are derived in Sections 3 and 5, and an entanglement criterion for qubits in terms of the quantum correlation distance is obtained in Section 4. Connections with classically-correlated quantum states are briefly discussed in Section 6, and conclusions are presented in Section 7.

2. Definitions and Main Results

For two classical random variables A and B, with joint probability distribution $P_{AB}(a,b)$ and marginal distributions $P_A(a)$ and $P_B(b)$, the Shannon mutual information and the classical correlation distance are defined respectively by

$$I(P_{AB}) := H(P_{AB}\|P_A P_B) = H(P_A) + H(P_B) - H(P_{AB}) \tag{2}$$

$$C(P_{AB}) := \|P_{AB} - P_A P_B\|_1 = \sum_{a,b} |P_{AB}(a,b) - P_A(a)P_B(b)| \tag{3}$$

where $H(P) := -\sum_j P(j) \log P(j)$ denotes the Shannon entropy of distribution P. The term "correlation distance" is used for $C(P_{AB})$, since it inherits all the properties of a distance from the

more general variational distance, and clearly vanishes for uncorrelated A and B. Both the mutual information and correlation distance have a minimum value of zero, and for n-valued random variables have maximum values

$$I(P_{AB}) \leq \log n, \qquad C(P_{AB}) \leq 2(n-1)/n \tag{4}$$

with saturation corresponding to the maximally correlated case $P_{AB}(a,b) = n^{-1}\delta_{ab}$. Note for $n = 2$ that $C(P_{AB}) \leq 1$.

For two quantum systems A and B described by density operator ρ_{AB} and reduced density operators ρ_A and ρ_B, the corresponding quantum mutual information and quantum correlation distance are analogously defined by

$$I(\rho_{AB}) := S(\rho_A) + S(\rho_B) - S(\rho_{AB}) \tag{5}$$

$$C(\rho_{AB}) := \|\rho_{AB} - \rho_A \otimes \rho_B\|_1 = \mathrm{tr}|\rho_{AB} - \rho_A \otimes \rho_B| \tag{6}$$

where $S(\rho) := -\mathrm{tr}[\rho \log \rho]$ denotes the von Neumann entropy of density operator ρ. Similarly to the classical case, these are direct measures of the correlation between A and B, vanishing only for uncorrelated A and B. It may be noted that trace distance has recently also been used to distinguish between and quantify quantum and classical contributions to this correlation [10,11]. For n-level quantum systems one has the maximum values

$$I(\rho_{AB}) \leq \log n^2, \qquad C(\rho_{AB}) \leq 2(n^2-1)/n^2 \tag{7}$$

with saturation corresponding to maximally entangled states. Thus, comparing with Equation (4), quantum correlations have a quadratic advantage with respect to both mutual information and correlation distance. For example, for $n = 2$ one has $C(\rho_{AB}) \leq 3/2$, allowing quantum correlation distances that are greater than the corresponding classical maximum value of unity.

In both the classical and quantum cases, one has the lower bound

$$I \geq \frac{1}{2} C^2 \log e \tag{8}$$

for mutual information, as a direct consequence of the Pinsker inequality (1) for classical relative entropies [2,5–7]. However, better bounds for mutual information can be obtained, which are stronger than any general inequality for relative entropy and variational distance.

For example, for two-valued classical random variables A and B one has the tight lower bound

$$I(P_{AB}) \geq \log 2 - H\left(\frac{1 + C(P_{AB})}{2}, \frac{1 - C(P_{AB})}{2}\right) \tag{9}$$

for classical mutual information. This inequality has been previously stated without proof in Reference [12], where it was used to bound the shared information required to classically simulate entangled quantum systems. It is proved in Section 3 below.

In contrast to Pinsker-type inequalities such as Equation (8), the quantum generalisation of Equation (9) is not straightforward. In particular, note for a two-qubit system that one cannot simply replace P_{AB} by ρ_{AB} in Equation (9), as the right hand side would be undefined for $C(\rho_{AB}) > 1$.

This can occur if the qubits are entangled. Indeed, as shown in Section 4, $C(\rho_{AB}) > 1$ is a sufficient condition for the entanglement of two qubits, as is the stronger condition

$$C(\rho_{AB}) > 2\sqrt{(1-\mathrm{tr}[\rho_A^2])\,(1-\mathrm{tr}[\rho_B^2])} \tag{10}$$

An explicit expression for the quantum correlation distance for two qubits, in terms of the spin covariance matrix, is also given in Section 4.

It is shown in Section 5 that the quantum equivalent of Equation (9), *i.e.*, a tight lower bound for the quantum mutual information shared by two qubits, is

$$I(\rho_{AB}) \geq \begin{cases} \log 2 - H\left(\frac{1+C(\rho_{AB})}{2}, \frac{1-C(\rho_{AB})}{2}\right), & C(\rho_{AB}) \leq C_0 \\ \log 4 - H\left(\frac{1}{4}+\frac{C(\rho_{AB})}{2}, \frac{1}{4}-\frac{C(\rho_{AB})}{6}, \frac{1}{4}-\frac{C(\rho_{AB})}{6}, \frac{1}{4}-\frac{C(\rho_{AB})}{6}\right), & C(\rho_{AB}) > C_0 \end{cases} \tag{11}$$

when the reduced density operators are maximally mixed, where $C_0 \approx 0.72654$ corresponds to the value of $C(\rho_{AB})$ for which the two expressions are equal. For $C(\rho_{AB}) > C_0$ this lower bound can only be achieved by entangled states, and cannot be achieved by any classical distribution P_{AB} having the same correlation distance. It is also shown that, for $C(\rho_{AB}) > C_0$, the bound is also tight if only one of the reduced states is maximally mixed. Support is given for the conjecture that the bound in Equation (11) in fact holds for all two-qubit states.

In Section 6 the natural role of "classically-correlated" quantum states, in comparing classical and quantum correlations, is briefly discussed. Such states have the general form $\rho_{AB} = \sum_{j,k} P(j,k)|j,k\rangle\langle j,k|$ [13], where $P(j,k)$ is a classical joint probability distribution and $\{|j\rangle\}$ and $\{|k\rangle\}$ are orthonormal basis sets for the two quantum systems. The lower bound in Equation (11) can be saturated by a classically-correlated state if and only if $C \leq C_0$.

3. Tight Lower Bound for Classical Mutual Information

3.1. Derivation of Bound

The tight lower bound in Equation (9) is derived here. The bound is plotted in Figure 1 below (top curve). Also plotted for comparison are the Pinsker lower bound in Equation (8) (bottom curve), and the lower bound following from the best possible generic inequality for relative entropy and variational distance, given in parametric form in Reference [5] (middle curve).

To derive the bound in Equation (9), it is convenient to label the two possible values of A and B by ± 1. Defining $R(a,b) := 4[P_{AB}(a,b) - P_A(a)P_B(b)]$, it follows by summing over each of a and b separately that $R(a,b) + R(a,-b) = 0 = R(a,b) + R(-a,b)$, implying that $R(a,b) = abr$ with $r := R(1,1)$. Hence, $C(P_{AB}) = |r|$. Further, writing $P_A(a) = (1+ax)/2$ and $P_B(b) = (1+by)/2$, for suitable $x, y \in [-1,1]$, the positivity condition $P_{AB}(a,b) \geq 0$ is equivalent to

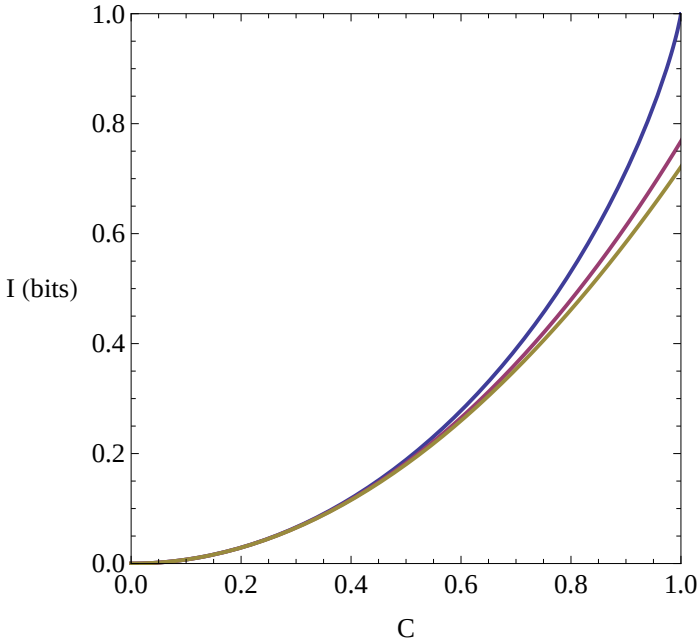
$$|x+y| - 1 \leq r + xy \leq 1 - |x-y| \tag{12}$$

Now, Equation (9) is equivalent to

$$f(r) := I(P_{AB}) - \log 2 + H\left(\frac{1+r}{2}, \frac{1-r}{2}\right) \geq 0 \tag{13}$$

It is easy to check that this inequality is always saturated for the case of maximally-random marginals, *i.e.*, when $x = y = 0$. In all other cases, the inequality may be proved by showing that $f(r)$ has a unique global minimum value of 0 at $r = 0$.

Figure 1. Lower bounds for the classical mutual information between two-valued variables.



In particular, note first that $f(0) = 0$ (one has $P_{AB} = P_A P_B$ in this case, so that the mutual information vanishes). Further, using $P_{AB}(a,b) = [(1+ax)(1+by)+abr]/4$, one easily calculates that, using logarithm base e for convenience,

$$f'(r) = \frac{1}{4}\sum_{a,b} ab \log P_{AB}(a,b) - \frac{1}{2}\sum_{a} a \log\frac{1+ar}{2} = \frac{1}{4}\log\frac{p_{AB}(+,+)\,p_{AB}(-,-)\,(1-r)^2}{p_{AB}(+,-)\,p_{AB}(-,+)\,(1+r)^2} \quad (14)$$

Hence, $f'(r) = 0$ if and only if the argument of the logarithm is unity, *i.e.*, if and only if

$$[(1+x)(1+y)+r]\,[(1-x)(1-y)+r]\,(1-r)^2 = [(1+x)(1-y)-r]\,[(1-x)(1+y)-r]\,(1+r)^2. \quad (15)$$

Expanding and simplifying yields two possible solutions: $r = 0$, or $r = (x^2+y^2-x^2y^2)/(2xy)$. However, in the latter case one has

$$|r+xy| = \frac{x^2+y^2+x^2y^2}{2|xy|} = \frac{\alpha}{\gamma} + \frac{\gamma}{2} \geq 1 + \frac{\gamma}{2} \geq 1 \quad (16)$$

where α and γ denote the arithmetic mean and geometric mean, respectively, of x^2 and y^2 (hence $\alpha \geq \gamma$). This is clearly inconsistent with the positivity condition (12) (unless $x = y = 0$, which trivially saturates Equation (13) for all r as noted above). The only remaining solution to $f'(r) = 0$ is then $r = 0$, implying $f(r)$ has a unique maximum or minimum value at $r = 0$. Finally, it is easily checked that it is a minimum, since

$$f''(0) = \frac{1}{16}\sum_{a,b}\frac{1}{p_A(a)p_B(b)} - 1 = \frac{1}{16P_A(+)P_A(-)P_B(+)P_B(-)} - 1 = \frac{1}{(1-x^2)(1-y^2)} - 1 \geq 0 \quad (17)$$

(with equality only for the trivially-saturating case $x = y = 0$). Thus, $f(r) \geq f(0) = 0$ as required.

3.2. Application: Resources for Simulating Bell Inequality Violation

The hallmark feature of quantum correlations is that they cannot be explained by any underlying statistical model that satisfies three physically very plausible properties: (i) no signaling faster than the speed of light; (ii) free choice of measurement settings; and (iii) independence of local outcomes. Various interpretations of quantum mechanics differ in regard to which of these properties should be given up. It is of interest to consider by *how much* they must be given up, in terms of the information-theoretic resources required to simulate a given quantum correlation [14]. For example, how many bits of communication, or bits of correlation between the source and the measurement settings, or bits of correlation between the outcomes, are required? The lower bound for classical mutual information in Equation (9) is relevant for the last of these questions.

In more detail, if $P_{AB}(a,b)$ denotes the joint probability of outcomes a and b, for measurements of variables A and B on respective spacelike-separated systems, and λ denotes any underlying variables relevant to the correlations, then Bayes theorem implies that

$$P_{AB}(a,b) = \sum_{\lambda} p_{AB}(\lambda)\, P_{AB}(a,b|\lambda) \tag{18}$$

where summation is replaced by integration over any continuous values of λ. The no-signaling property requires that the underlying marginal distribution of A, $p_A(a|\lambda)$, is independent of whether B or B' was measured on the second system (and vice versa), while the free-choice property requires that λ is independent of the choice of the measured variables A and B, *i.e.*, that $p_{AB}(\lambda) = p_{A'B'}(\lambda)$ for any A, A', B, B'. Finally, the outcome independence property requires that any observed correlation between A and B arises from ignorance of the underlying variable, *i.e.*, that $P_{AB}(a,b|\lambda) = P_A(a|\lambda)\, P_B(b|\lambda)$ for all A, B and λ. Thus the correlation distance of $P_{AB}(a,b|\lambda)$ vanishes identically:

$$C(P_{AB|\lambda}) \equiv 0 \tag{19}$$

As is well known, the assumption of all three properties implies that two-valued random variables with values ± 1 must satisfy the Bell inequality [15]

$$\langle AB\rangle + \langle AB'\rangle + \langle A'B\rangle - \langle A'B'\rangle \leq 2 \tag{20}$$

whereas quantum correlations can violate this inequality by as much as a factor of $\sqrt{2}$ [16]. It follows that quantum correlations can only be modeled by relaxing one or more of the above properties, as has recently been reviewed in detail in Reference [12].

For example, assuming that no-signaling and measurement independence hold (as they do in the standard Copenhagen interpretation of quantum mechanics), and defining $C_{\max}$ to be the maximum value of $C(P_{AB|\lambda})$ over all A, B and λ, it can be shown that Equation (20) generalises to the tight bound [12]

$$\langle AB\rangle + \langle AB'\rangle + \langle A'B\rangle - \langle A'B'\rangle \leq \frac{4}{2 - C_{\max}} \tag{21}$$

It follows that to simulate a Bell inequality violation $\langle AB\rangle + \langle AB'\rangle + \langle A'B\rangle - \langle A'B'\rangle = 2 + V$, for some $V > 0$, the observers must share random variables having a correlation distance of at least

$C_{\max} \geq 2V/(2+V)$. Hence, using the classical lower bound Equation (9) (stated without proof in Reference [12]), the observers must share a minimum mutual information of

$$I_{\min} = \log 2 - H\left(\frac{1+C_{\max}}{2}, \frac{1-C_{\max}}{2}\right) \geq \log 2 - H\left(\frac{2+3V}{4+2V}, \frac{2-V}{4+2V}\right) \tag{22}$$

Note this reduces to zero in the limit of no violation of Bell inequality (20), *i.e.*, when $V = 0$, and reaches a maximum of 1 bit of information in the limit of the maximum possible violation over arbitrary probability distributions, $V = 2$. To simulate the maximum quantum violation [16], *i.e.*, $V = 2\sqrt{2} - 2$, at least 0.264 bits are required.

4. Quantum Correlation Distance and Qubit Entanglement

The positivity condition (12) may be used to show that the classical correlation distance between any pair of two-valued random variables is never greater than unity, *i.e.*, that $C(P_{AB}) = |r| \leq 1$ [12]. In contrast, the quantum correlation distance between a pair of qubits can be greater than unity, with upper bound $C(\rho_{AB}) \leq 3/2$, as per Equation (7).

Non-classical values of the quantum correlation distance are closely related to the quintessential non-classical feature of quantum mechanics: entanglement. In particular, $C(\rho_{AB}) > 1$ is a direct signature of qubit entanglement. Indeed, even correlation distances smaller than unity can imply two qubits are entangled, as per the criterion given in Equation (10) and shown below. An explicit formula for qubit correlation distance in terms of the spin covariance matrix, needed for Section 5, is also obtained below.

4.1. Entanglement Criterion

Recall that the density operator ρ_{AB} of two qubits may always be written in the Fano form [17]

$$\begin{aligned} \rho_{AB} &= \frac{1}{4}\left[I \otimes I + u \cdot \sigma \otimes I + I \otimes v \cdot \sigma + \sum_{j,k} \langle \sigma_j \otimes \sigma_k \rangle \, \sigma_j \otimes \sigma_k \right] \\ &= \rho_A \otimes \rho_B + \frac{1}{4} \sum_{j,k} T_{jk} \, \sigma_j \otimes \sigma_k \end{aligned} \tag{23}$$

Here I is the unit operator; $\{\sigma_j\}$ denotes the set of Pauli spin observables on each qubit Hilbert space; the components of the 3-vectors u and v are the spin expectation values $u_j := \langle \sigma_j \otimes 1 \rangle$ and $v := \langle 1 \otimes \sigma_k \rangle$, for A and B respectively; and T denotes the 3×3 spin covariance matrix with coefficients

$$T_{jk} := \langle \sigma_j \otimes \sigma_k \rangle - \langle \sigma_j \otimes I \rangle \langle I \otimes \sigma_k \rangle \tag{24}$$

It immediately follows from Equation (23) that the quantum correlation distance may be expressed in terms of the spin covariance matrix as

$$C(\rho_{AB}) = \frac{1}{4} \mathrm{tr} \left| \sum_{j,k} T_{jk} \, \sigma_j \otimes \sigma_k \right| \tag{25}$$

This expression will be further simplified in Section 4.2.

Now consider the case where ρ_{AB} is a separable state, *i.e.*, of the unentangled form

$$\rho_{AB} = \sum_{\lambda} p(\lambda)\, \tau_A(\lambda) \otimes \omega_B(\lambda) \tag{26}$$

for some probability distribution $p(\lambda)$ and local density operators $\{\tau_A(\lambda)\}$, $\{\omega_B(\lambda)\}$. Defining $u_j(\lambda) := \mathrm{tr}[\tau_A(\lambda)\sigma_j]$, $v_k(\lambda) := \mathrm{tr}[\omega_B(\lambda)\sigma_k]$ implies $u = \sum_\lambda p(\lambda)u(\lambda)$ and $v = \sum_\lambda p(\lambda)v(\lambda)$, and substitution into Equation (25) then yields

$$\begin{aligned}
C(\rho_{AB}) &= \frac{1}{4}\left\| \sum_{\lambda} p(\lambda)\,[u(\lambda)-u]\cdot\sigma \otimes [v(\lambda)-v]\cdot\sigma \right\|_1 \\
&\leq \frac{1}{4}\sum_{\lambda} p(\lambda)\, \|[u(\lambda)-u]\cdot\sigma\|_1\, \|[v(\lambda)-v]\cdot\sigma\|_1 \\
&= \sum_{\lambda} p(\lambda)|u(\lambda)-u|\,|v(\lambda)-v| \leq \left[\sum_{\lambda} p(\lambda)|u(\lambda)-u|^2\right]^{1/2} \left[\sum_{\lambda} p(\lambda)|v(\lambda)-v|^2\right]^{1/2} \\
&= \left[\sum_{\lambda} p(\lambda)|u(\lambda)|^2 - |u|^2\right]^{1/2} \left[\sum_{\lambda} p(\lambda)|v(\lambda)|^2 - |v|^2\right]^{1/2} \\
&\leq \sqrt{(1-u\cdot u)(1-v\cdot v)}
\end{aligned} \tag{27}$$

Note that second line follows from the properties $\|X+Y\|_1 \leq \|X\|_1 + \|Y\|_1$ and $\|XY\|_1 \leq \|X\|_1\|Y\|_1$ of the trace norm [18]; the third line using $\|X\|_1 = \mathrm{tr}[\sqrt{X^\dagger X}]$ and the Schwarz inequality; and the last line via $|u(\lambda)|, |v(\lambda)| \leq 1$.

Equation (27) holds for all separable qubit states. Hence, a non-classical value of the correlation distance, $C(\rho_{AB}) > 1$, immediately implies that the qubits must be entangled. More generally, noting that $\rho_A = \frac{1}{2}(I + u\cdot\sigma)$ and $\rho_B = \frac{1}{2}(I + v\cdot\sigma)$, one has $\mathrm{tr}[\rho_A^2] = (1+u\cdot u)/2$, $\mathrm{tr}[\rho_B^2] = (1+v\cdot v)/2$, and the stronger entanglement criterion (10) immediately follows from Equation (27).

The fact that entanglement between two qubits is necessary (but not sufficient) for $C(\rho_{AB})$ to be greater than the maximum possible value of $C(P_{AB})$, for two-valued classical variables, is a nice distinction between quantum and classical correlation distances. It would be of interest to determine whether this result generalises to n-level systems. This would follow from the validity of Equation (10) for arbitrary quantum systems.

4.2. Explicit Expression for $C(\rho_{AB})$

To explicitly evaluate $C(\rho_{AB})$ in Equation (25), let $T = KDL^T$ denote a singular value decomposition of the spin covariance matrix. Thus, K and L are real orthogonal matrices and $D = \mathrm{diag}[t_1, t_2, t_3]$, with the singular values $t_1 \geq t_2 \geq t_3 \geq 0$ corresponding to the square roots of the eigenvalues of TT^T. Noting that any 3×3 orthogonal matrix is either a rotation matrix, or the product of a rotation matrix with the parity matrix $-I$, one therefore always has a decomposition of the form $T = \pm KDL^T$ where K and L are now restricted to be rotation matrices. Hence, defining unitary operators U and V corresponding to rotations K and L, via $U\sigma_j U^\dagger = \sum_{j,j'} K_{jj'}\sigma_{j'}$ and

$V\sigma_j V^\dagger = \sum_{j,j'} L_{jj'}\sigma_{j'}$, and using the invariance of the trace norm under unitary transformations, the quantum correlation distance in Equation (25) can be rewritten as

$$C(\rho_{AB}) = \frac{1}{4}\mathrm{tr}\left|\pm\sum_j t_j\, U\sigma_j U^\dagger \otimes V\sigma_j V^\dagger\right| = \frac{1}{4}\mathrm{tr}\left|\sum_j t_j\, \sigma_j \otimes \sigma_j\right| \tag{28}$$

Determining the eigenvalues of the Hermitian operator $\sum_j t_j\, \sigma_j \otimes \sigma_j$ is a straightforward 4×4 matrix calculation using the standard representation of the Pauli sigma matrices. Summing the absolute values of these eigenvalues then yields the explicit expression

$$\begin{aligned} C(\rho_{AB}) &= \frac{1}{4}\left[|t_1+t_2+t_3| + |t_1+t_2-t_3| + |t_1-t_2+t_3| + |-t_1+t_2+t_3|\right] \\ &= \frac{1}{2}\max\{t_1+t_2+t_3, 2t_1\} \end{aligned} \tag{29}$$

for the quantum correlation distance, in terms of the singular values of the spin covariance matrix.

For example, for the Werner state $\rho_{AB} = p|\psi\rangle\langle\psi| + (1-p)/4\, I \otimes I$, where $|\psi\rangle$ is the singlet state and $-1/3 \leq p \leq 1$ [19], one has $T = -pI$ and hence that $t_1 = t_2 = t_3 = |p|$. The corresponding correlation distance is therefore $3|p|/2$, which is greater than the classical maximum of unity for $p > 2/3$.

Equation (29) also allows the qubit entanglement criterion (10) to be directly compared with the strongest known criterion based on the spin covariance matrix [20]:

$$t_1 + t_2 + t_3 > 2\sqrt{(1-\mathrm{tr}[\rho_A^2])\,(1-\mathrm{tr}[\rho_B^2])} \tag{30}$$

For the above Werner state this criterion is tight, indicating entanglement for $p > 1/3$. Hence, the main interest in weaker entanglement criteria based on quantum correlation distance lies in their direct connection with non-classical values of the classical correlation distance.

5. Tight Lower Bound for Quantum Mutual Information

Here Equation (11) is derived for the case $\rho_A = \rho_B = \frac{1}{2}I$. Evidence is provided for the conjecture that Equation (11) in fact holds for all two-qubit states, including a partial generalisation of Equation (11) when only one of ρ_A and ρ_B is maximally-mixed.

5.1. Derivation for Maximally-Mixed ρ_A and ρ_B

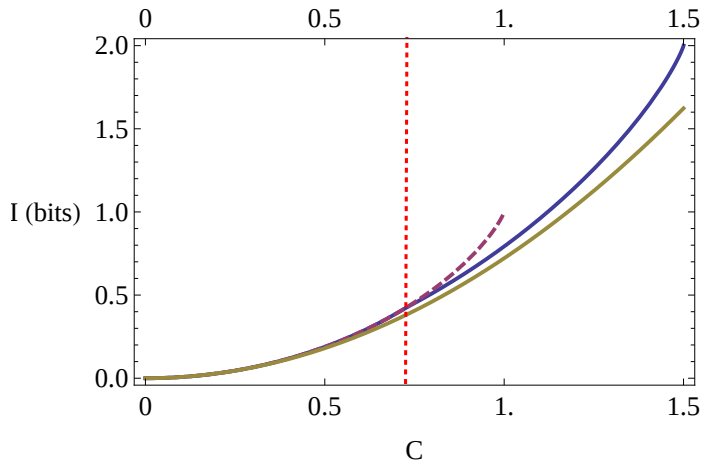
The tight lower bound for quantum mutual information in Equation (11), for maximally-mixed reduced states, is plotted in Figure 2 below (top solid curve). Also plotted for comparison are the Pinsker lower bound in Equation (8) (bottom solid curve), and classical lower bound in Equation (9) (dashed curve). The dotted vertical line indicates the value of $C_0 \approx 0.72654$ in Equation (11). It is seen that quantum correlations can violate the classical lower bound for correlation distances falling between C_0 and 1.

To derive Equation (11) for $\rho_A = \rho_B = \frac{1}{2}I$, note first that Equation (23) reduces to $\rho_{AB} = \frac{1}{4}[I \otimes I + \sum_{j,k} T_{jk}\, \sigma_j \otimes \sigma_k]$. By the same argument given in Section 4.2, this can be transformed via local unitary transformations to the state

$$\tilde{\rho}_{AB} = \frac{1}{4}\left[I \otimes I + \sum_j r_j\, \sigma_j \otimes \sigma_j\right] \tag{31}$$

where $r_j = \alpha t_j$, $\alpha = \pm 1$, and $t_1 \geq t_2 \geq t_3 \geq 0$ are the singular values of the spin covariance matrix T. Since the quantum mutual information and quantum correlation distance are invariant under local unitary transformations, one has $I(\rho_{AB}) = I(\tilde{\rho}_{AB})$ and $C(\rho_{AB}) = C(\tilde{\rho}_{AB})$. Hence Equation (11) only needs to be demonstrated for $\tilde{\rho}_{AB}$.

Figure 2. Lower bounds for the quantum mutual information between two qubits.



The mutual information of $\tilde{\rho}_{AB}$ is easily evaluated as

$$I(\tilde{\rho}_{AB}) = S(\tilde{\rho}_A) + S(\tilde{\rho}_B) - S(\tilde{\rho}_{AB}) = \log 4 - H(p_0, p_1, p_2, p_3) \tag{32}$$

where $p_0 = \frac{1}{4}(1-r_1-r_2-r_3)$, $p_1 = \frac{1}{4}(1-r_1+r_2+r_3)$, $p_2 = \frac{1}{4}(1+r_1-r_2+r_3)$, $p_3 = \frac{1}{4}(1+r_1+r_2-r_3)$ are the eigenvalues of $\tilde{\rho}_{AB}$. Inverting the relation between the r_j and p_j further yields

$$t_j = \alpha r_j = \alpha[1 - 2(p_0 + p_j)], \quad t_1 + t_2 + t_3 = \alpha(1 - 4p_0) \tag{33}$$

and hence the correlation distance follows from Equation (29) as

$$C(\tilde{\rho}_{AB}) = C := \frac{1}{2}\max\{\alpha(1 - 4p_0), \alpha(1 - 4p_0 + 1 - 4p_1)\} \tag{34}$$

Equation (32) implies that a tight lower bound for $I(\tilde{\rho}_{AB})$ corresponds to a tight upper bound for $H(p_0, p_1, p_2, p_3)$. To determine the maximum value of $H(p_0, p_1, p_2, p_3)$, for a fixed correlation distance C, consider first the case $\alpha = 1$. The ordering and positivity conditions on t_j then require

$$p_1 \leq p_2 \leq p_3, \quad p_0 + p_j \leq \frac{1}{2} \quad (j = 1, 2, 3) \tag{35}$$

(implying $p_0 \leq 1/4$). Further, from Equation (34), $C = \frac{1}{2}\max\{1 - 4p_0, 1 - 4p_0 + 1 - 4p_1\}$. Hence, if $p_1 \leq 1/4$, then $C = 1 - 2(p_0 + p_1) \leq 1$, implying the constraint $p_0 + p_1 = (1 - C)/2$. Noting the concavity of entropy, the maximum possible entropy under this constraint corresponds to equal values

$$p_0 = p_1 = (1 - C)/4, \quad p_2 = p_3 = (1 + C)/4 \tag{36}$$

(which are compatible with the above conditions on the p_j). Conversely, if $p_1 \geq 1/4$ then $C = (1 - 4p_0)/2 \leq 1/2$, and hence $p_0 = 1/4 - C/2$ is fixed, implying by concavity that the maximum possible entropy corresponds to

$$p_0 = 1/4 - C/2, \quad p_1 = p_2 = p_3 = 1/4 + C/6 \tag{37}$$

(which again satisfies the required conditions on the p_j). It follows that the maximum possible entropy is (i) the maximum of the entropies

$$H_1(C) := H((1-C)/4, (1-C)/4, (1+C)/4, (1+C)/4) \tag{38}$$

$$H_2(C) := H(1/4 - C/2, 1/4 + C/6, 1/4 - C/6, 1/4 - C/6) \tag{39}$$

for $C \leq 1/2$, and (ii) $H_1(C)$ for $1/2 < C \leq 1$. However, it is straightforward to show that $H_1(C) \geq H_2(C)$ over their overlapping range. Hence the maximum possible entropy is always $H_1(C)$ for the case $\alpha = 1$.

For the case $\alpha = -1$, the conditions on t_j require that

$$p_1 \geq p_2 \geq p_3, \quad p_0 + p_j \geq \frac{1}{2} \quad (j = 1, 2, 3) \tag{40}$$

(implying $p_0 \geq 1/4$), while from Equation (34) $C = \frac{1}{2}\max\{4p_0 - 1, 4p_0 - 1 + 4p_1 - 1)\}$. Carrying out a similar analysis to the above, one finds that the maximum possible entropy is (i) the maximum of the entropies $H_1(C)$ and

$$H_3(C) := H(1/4 + Q/2, 1/4 - Q/6, 1/4 - Q/6, 1/4 - Q/6) \tag{41}$$

for $C \leq 1$, and (ii) $H_3(C)$ for $1 < C \leq 3/2$.

Numerical comparison shows that $H_3(C) > H_1(C)$ for $C > C_0 \approx 0.72654$, and $H_3(C) \leq H_1(C)$ otherwise. Hence, from Equation (32) one has the tight lower bound

$$I(\tilde{\rho}_{AB}) \geq \begin{cases} \log 4 - H_1(C), & C \leq C_0 \\ \log 4 - H_3(C), & C > C_0 \end{cases} \tag{42}$$

Since $H_1(C) = \log 2 + H((1-C)/2, (1+C)/2)$, it follows that Equation (11) holds for $\tilde{\rho}_{AB}$ in Equation (31), and hence for all qubit states with maximally-mixed reduced density operators, as claimed.

The states saturating the lower bound in Equations (11) and (42) are easily constructed from the above derivation. In particular, they are given by

$$\rho(C) := \begin{cases} \frac{1}{4}\left[I \otimes I + C\,\sigma_1 \otimes \sigma_1\right], & C \leq C_0 \\ \frac{1}{4}\left[I \otimes I - (2C/3)\sum_j \sigma_j \otimes \sigma_j\right], & C > C_0 \end{cases} \tag{43}$$

and any local unitary transformations thereof, where the quantum correlation distance of $\rho(C)$ is C by construction.

Note that $\rho(C)$ is unentangled for $C \leq C_0$ (it can be written as a mixture of $(1/4)I \otimes I$, $|+\rangle\langle+| \otimes |+\rangle\langle+|$ and $|-\rangle\langle-| \otimes |-\rangle\langle-|$, where $\sigma_1|\pm\rangle = \pm|\pm\rangle$). Conversely, $\rho(C)$ is an entangled Werner state for $C \geq C_0$ (with singlet state weighting $p = 2C/3 > 1/3$). Hence, the lower bound in Equations (11) and (42) can only be achieved by entangled states for $C \geq C_0$, and cannot be achieved by any two-valued classical random variables.

5.2. Conjecture

It is conjectured that Equation (11) is in fact a tight lower bound for any two-qubit state. This conjecture would follow immediately if it could be shown that

$$I(\rho_{AB}) \geq I(\rho'_{AB}) \tag{44}$$

for arbitrary ρ_{AB}, where $\rho'_{AB} := \rho_{AB} - \rho_A \otimes \rho_B + (1/4)I \otimes I$, and where it must further be shown that ρ'_{AB} is a density operator. The conjecture would then follow since ρ'_{AB} is of the form of $\tilde{\rho}_{AB}$ in Equation (31), and hence $I(\rho'_{AB})$ satisfies Equation (42).

Partial support for Equation (44), and hence for the conjecture, is given by noting that any ρ_{AB} and corresponding ρ'_{AB} can be brought to the respective forms

$$\rho_{AB} = \rho_A \otimes \rho_B + \frac{1}{4}\sum_j r_j\, \sigma_j \otimes \sigma_j, \quad \rho'_{AB} = \frac{1}{4}\left[I \otimes I + \sum_j r_j\, \sigma_j \otimes \sigma_j\right] \tag{45}$$

via suitable local unitary transformations, similarly to the argument in Section 4.2. Defining the function

$$F(r_1, r_2, r_3) := I(\rho_{AB}) - I(\rho'_{AB}) \tag{46}$$

it is straightforward to show that $F = 0$ and $\partial F/\partial r_j = 0$ for $r_1 = r_2 = r_3 = 0$, consistent with $F \geq 0$. However, it remains to be shown that the gradient $\partial F/\partial r_j = 0$ does not vanish for other physically possible values of r_j (other than for the trivially saturating case $\rho_A = \rho_B = (1/2)I$).

The above conjecture is further supported by the generalisation of Equation (11) in the following section.

5.3. Generalisation to Maximally-Mixed ρ_A or ρ_B

It is straightforward to show that the lower bound on quantum mutual information is tight for $C \geq C_0$ when just *one* of the mixed density operators is mixed, *i.e.*, if ρ_A or ρ_B is equal to $(1/2)I$.

First, since $(1/2)I$ is invariant under unitary transformations, the same argument as in Section 4.2 implies the state can always be transformed by local unitary transformations to the generalised form

$$\tilde{\rho}_{AB} = \frac{1}{4}\left[\tilde{\rho}_A \otimes \tilde{\rho}_B + \alpha \sum_j t_j\, \sigma_j \otimes \sigma_j\right] \tag{47}$$

of Equation (31), where either $\tilde{\rho}_A$ or $\tilde{\rho}_B$ equals $(1/2)I$ and $\alpha = \pm 1$.

Second, let $\mathcal{T}$ denote the "twirling" operation, corresponding to applying a random unitary transformation of the form $U \otimes U$ [21]. It is easy to check that by definition $\mathcal{T}(I \otimes I) = I \otimes I$, $\mathcal{T}(I \otimes \sigma_j) = 0 = \mathcal{T}(\sigma_j \otimes 1)$ and $\mathcal{T}(\sigma_j \otimes \sigma_j) = \mathcal{T}(\sigma_k \otimes \sigma_k)$, for any j and k. Since Werner states are invariant under twirling [19,21], it follows that $\mathcal{T}(\sigma_j \otimes \sigma_j) = (1/3)\sum_k \sigma_k \otimes \sigma_k$. Using these properties, one finds that $\mathcal{T}(\tilde{\rho}_A \otimes \tilde{\rho}_B) = (1/4)I \otimes I$ if one of $\tilde{\rho}_A$ or $\tilde{\rho}_B$ is maximally mixed, and hence that

$$\mathcal{T}(\tilde{\rho}_{AB}) = \frac{1}{4}\left[I \otimes I + \alpha\bar{t} \sum_j \sigma_j \otimes \sigma_j\right] = \rho(-3\alpha\bar{t}/2) \tag{48}$$

where $\bar{t} := (t_1 + t_2 + t_3)/3$ and the second equality holds for $C \geq C_0$ (but not otherwise), with $\rho(C)$ defined as per Equation (43). Further, from Equation (29) one has

$$C(\mathcal{T}(\tilde{\rho}_{AB})) = C = \frac{1}{2}\max\{2\bar{t}, 3\bar{t}\} = 3\bar{t}/2 \tag{49}$$

Recalling that $\rho(C)$ saturates Equation (42), an analysis similar to Section 5.1 shows for $C \geq C_0$ that

$$I(\mathcal{T}(\tilde{\rho}_{AB})) = \log 4 - H_3(-\alpha C) \geq \log 4 - H_3(C) \tag{50}$$

with equality for $\alpha = -1$.

Third, again using $\mathcal{T}(\tilde{\rho}_A \otimes \tilde{\rho}_B) = (1/4)I \otimes I$, and the property that the relative entropy is non-increasing under the twirling operation, it follows that

$$I(\tilde{\rho}_{AB}) = S(\tilde{\rho}_{AB}\|\tilde{\rho}_A \otimes \tilde{\rho}_B) \geq S(\mathcal{T}(\tilde{\rho}_{AB})\|\mathcal{T}(\tilde{\rho}_A \otimes \tilde{\rho}_B)) = I(\mathcal{T}(\tilde{\rho}_{AB})) \geq \log 4 - H_3(C) \tag{51}$$

for $C \geq C_0$. Since Werner states are invariant under twirling, this inequality is tight for $\alpha = -1$, being saturated by the choice $\tilde{\rho}_{AB} = \rho(C)$. Recalling that mutual information and correlation distance are invariant under local unitary operations, the inequality is therefore tight for any ρ_{AB} for which one of ρ_A and ρ_B is maximally mixed, as claimed.

6. Classically-Correlated Quantum States

It is well known that a quantum system behaves classically if the state and the observables of interest all commute, *i.e.*, if they can be simultaneously diagonalised in some basis. Hence, a joint state will behave classically if the relevant observables of each system commute with each other and the state. It is therefore natural to define ρ_{AB} to be *classically correlated* if and only if it can be diagonalised in a joint basis [13], *i.e.*, if and only if

$$\rho_{AB} = \sum_{j,k} P(j,k)|j\rangle\langle j| \otimes |k\rangle\langle k| \tag{52}$$

for some distribution $P(j,k)$ and orthonormal basis set $\{|j\rangle \otimes |k\rangle\}$. Classical correlation is preserved by tensor products, and by mixtures of commuting states.

While, strictly speaking, a classically-correlated quantum state only behaves classically with respect to observables that are diagonal with respect to $|j\rangle \otimes |k\rangle$, they also have a number of classical correlation properties with respect to general observables [13,22], briefly noted here.

First, ρ_{AB} above is separable by construction, and hence is unentangled. Second, since it is diagonal in the basis $\{|j\rangle \otimes |k\rangle\}$, the mutual information and correlation distance are easily calculated as

$$I(\rho_{AB}) = I(P), \qquad C(\rho_{AB}) = C(P) \tag{53}$$

and hence can only take classical values.

Third, if M and N denote any observables for systems A and B respectively, then their joint statistics are given by

$$P_{MN}(m,n) = \sum_{j,k} p(m|j)\, p(n|k)\, P(j,k) = \sum_{j,k} \mathcal{S}_{m,n;j,k}\, P(j,k) \tag{54}$$

where $\mathcal{S}_{m,n;j,k} = p(m|j)\, p(n|k)$ is a stochastic matrix with respect to its first and second pairs of indices. Similarly, one finds

$$P_M(m)\, P_N(n) = \sum_{j,k} \mathcal{S}_{m,n;j,k}\, P(j)\, P(k) \tag{55}$$

for the product of the marginals. Since the classical relative entropy and variational distance can only decrease under the action of a stochastic matrix, it follows that one has the tight inequalities [13,22]

$$I(P_{MN}) \leq I(P) = I(\rho_{AB}), \qquad C(P_{MN}) \leq C(P) = C(\rho_{AB}) \tag{56}$$

with saturation for M and N diagonal in the bases $\{|j\rangle\}$ and $\{|k\rangle\}$ respectively. Maximising the first of these equalities over M or N immediately implies the well-known result that classically-correlated states have zero quantum discord [22].

Finally, for two-qubit systems, Equation (52) implies that ρ_{AB} is classically correlated if and only if it is equivalent under local unitary transformations to a state of the form

$$\rho'_{AB} = \frac{1}{4}\left[(1 + x\sigma_1) \otimes (1 + y\sigma_1) + r\, \sigma_1 \otimes \sigma_1\right] \tag{57}$$

where $x, y \in [-1, 1]$ and r satisfies Equation (12). Hence, the mutual information is bounded by the classical lower bound in Equation (9), and $\rho(C)$ in Equation (43) is classically correlated for $C \leq C_0$. It follows that the lower bound for quantum mutual information in Equation (11) can be attained by classically-correlated states if $C \leq C_0$. Conversely, the minimum possible bound cannot be reached by any classically-correlated two-qubit state if $C > C_0$.

7. Conclusions

Lower bounds for mutual information have been obtained that are stronger than those obtainable from general bounds for relative entropy and variational distance. Unlike the Pinsker inequality in Equation (8), the quantum form of these bounds is not a simple generalisation of the classical form.

Similarly to the case of upper bounds for (classical) mutual information [9], the tight lower bounds obtained here depend on the dimension of the systems. The results of this paper represent a preliminary investigation largely confined to two-valued classical variables and qubits. It would be of interest to generalise both the classical and quantum cases, and to further investigate connections between them.

Open questions include whether a quantum correlation distance greater than the corresponding maximum classical correlation distance is a signature of entanglement for higher-dimensional systems, and whether the related qubit entanglement criterion in Equation (10) holds more generally. The conjecture in Section 5.2, as to whether the quantum lower bound in Equation (11) is valid for all two-qubit states, also remains to be settled. Finally, it would be of interest to generalise and to better understand the role of the transition from classically-correlated states to entangled states in saturating information bounds, in the light of Equation (43) for qubits.

Acknowledgements

This research was supported by the ARC Centre of Excellence CE110001027.

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Cover, T.M.; Thomas, J.A. *Elements of Information Theory*, 2nd ed.; John Wiley & Sons: Hoboken, NJ, USA, 2006; Chapter 11.
2. Schumacher, B.; Westmoreland, M.D. Approximate quantum error correction. *Quantum Inf. Process.* **2002**, *1*, 5–12.
3. Hayashi, M. Large deviation analysis for classical and quantum security via approximate smoothing. *Quantum Phys.* **2013**, arXiv:1202.0322v5.
4. He, X.; Yener, A. Strong secrecy and reliable Byzantine detection in the presence of an untrusted relay. *IEEE Trans. Inf. Theory* **2013**, *59*, 177–192.
5. Fedotov, A.A.; Harremöes, P.; Tøpsoe, F. Refinements of Pinsker's inequality. *IEEE Trans. Inf. Theory* **2003**, *49*, 1491–1498.
6. Hiai, F.; Ohya, M.; Tsukada, M. Sufficiency, KMS conditions and relative entropy in von Neumann algebras. *Pac. J. Math.* **1981**, *96*, 99–109.
7. Rastegin, A.E. Fano type quantum inequalities in terms of q-entropies. *Quantum Inf. Process.* **2012**, *11*, 1895–1910.
8. Brandão, F.G.S.L.; Harrow, A.W. Quantum de Finetti theorems under local measurements with applications. *Quantum Phys.* **2013**, arXiv:1210.6367v3.
9. Zhang, Z. Estimating mutual information via Kolmogorov distance. *IEEE Trans. Inf. Theory* **2007**, *53*, 3280–3282.
10. Aaronson, B.; Franco, R.L.; Compagno, G.; Adesso, G. Hierarchy and dynamics of trace distance correlations. *Quantum Phys.* **2013**, arXiv:1307.3953v1.
11. Paula, F.M.; Montealegre, J.D.; Saguia, A.; de Oliveira, T.R.; Sarandy, M.S. Geometric classical and total correlations via trace distance. *Quantum Phys.* **2013**, arXiv:1307.3579v1.
12. Hall, M.J.W. Relaxed Bell inequalities and Kochen-Specker theorems. *Phys. Rev. A* **2012**, *84*, 022102.
13. Piani, M.; Horodecki, P.; Horodecki, R. No-local-broadcasting theorem for multipartite quantum correlations. *Phys. Rev. Lett.* **2008**, *100*, 090502.
14. Toner, B.F.; Bacon, D. Communication cost of simulating Bell correlations. *Phys. Rev. Lett.* **2003**, *91*, 187904.
15. Clauser, J.F.; Horne, M.A.; Shimony, A.; Holt, R.A. Proposed experiment to test local hidden-variable theories. *Phys. Rev. Lett.* **1969**, *23*, 880–884.
16. Csirel'son, B.S. Quantum generalizations of Bell's inequality. *Lett. Math. Phys.* **1980**, *4*, 93–100.
17. Fano, U. Pairs of two-level systems. *Rev. Mod. Phys.* **1983**, *55*, 855–874.
18. Bhatia, R. *Matrix Analysis*; Springer-Verlag: New York, NY, USA, 1997; Section IV.2.
19. Werner, R.F. Quantum states with Einstein-Podolsky-Rosen correlations admitting a hidden-variable model. *Phys. Rev. A* **1989**, *40*, 4277–4281.

20. Zhang, C.J; Zhang, Y.S.; Zhang, S.; Guo, G.C. Entanglement detection beyond the computable cross-norm or realignment criterion. *Phys. Rev. A* **2008**, *77*, 060301(R).
21. Bennett, C.H.; DiVincenzo, D.P. Mixed-state entanglement and quantum error correction. *Phys. Rev. A* **1996**, *54*, 3824–3851.
22. Wu, S.; Poulsen, U.V.; Mølmer, K. Correlations in local measurements on a quantum state, and complementarity as an explanation of nonclassicality. *Phys. Rev. A* **2009**, *80*, 032319.