I. INTRODUCTION

Entanglement is a form of correlations between two or more quantum systems whose amount exceeds anything that can be obtained by the laws of classical physics. Bell was first to clarify the difference between these correlations by measuring statistical regularities between local measurements on two separated systems \([1]\). However, even though it is now well established that quantum systems can be more correlated than classical ones, the central question is whether we can use these excess correlations to do something useful. There are many indications from the field of quantum information that entanglement can result in a computational speed up, nonetheless, there is no precise and proven link between the two at present. In this paper, we show that there is indeed a more basic relationship between entanglement and its usefulness in thermodynamics. We derive an inequality showing that we can extract more work out of a heat bath via entangled systems than via classically correlated ones. We also analyze the work balance of the process as a heat engine, in connection with the Second Law of thermodynamics.

II. WORK-EXTRACTION SCHEME

Suppose that we have a two dimensional classical system, such as a “one-molecule gas” which can only be in either the right or the left side of a chamber. If we have full information about this molecule, i.e., we know its position with certainty, we can extract \(k_B T \ln 2\) of work out of a heat bath of temperature \(T\) by letting the gas expand isothermally. If we have only partial information about the system, the extractable work becomes \(k_B T \ln 2 [1 - H(X)]\), where \(H(X)\) is the Shannon entropy and \(X\) is a binary random variable representing the position of the molecule \([2, 3, 5, 6]\). For simplicity, we set \(k_B T \ln 2 = 1\) hereafter so that we can identify the amount of work with the amount of information in bits.

The same argument can be applied to quantum cases provided that we know the nature of projection operators \(\{P_0, P_1, (= P_0^\perp)\}\) employed to obtain the information. The corresponding Shannon entropy is \(H(p) = -p \log_2 p - (1 - p) \log_2 (1 - p)\), where \(p = \text{Tr}(P_0 \rho)\) and \(\rho\) is the density matrix for the state. In order to extract work, we can store the measurement results in classical bits so that the same process as above can be applied. Or, equivalently, we can copy the information about the quantum state to an ancilla by using controlled-NOT (CNOT) operation with respect to the basis defined by projectors \(\{P_0, P_1\}\), and use this ancilla to extract work after letting it dephase. A CNOT is a two-bit unitary operation, which flips one of the two bits, the target bit, if the other, the control bit, is in 1 and does nothing otherwise.

Let us now consider a bipartite correlated system, retained by Alice and Bob. Suppose a set of identically prepared copies of the system for which Alice chooses \(A_\theta = \{P_0, P_0^\perp\}\) and Bob chooses \(B_\theta' = \{P_0^\perp, P_0\}\) as bases of their measurements with \(\theta(\theta')\) representing the direction of the basis. Alice performs her measurement
with $A_\theta$ and sends all ancillae containing results to Bob. Then, by analogy with the single molecule case, Bob can extract $1 - H(B_{\theta'}|A_\theta)$ bits of work per pair on his side after compressing the information of his measurement outcomes, where $H(X|Y)$ is the Shannon entropy of $X$, conditional on the knowledge of $Y$.

When their system is in a maximally entangled state, such as $|\Phi^+\rangle = (|00\rangle + |11\rangle)/\sqrt{2}$, $H(A_\theta|B_\theta)$ vanishes for all $\theta$, unlike any other forms of correlation for which $H(A_\theta|B_\theta)$ can take any value between 0 and 1. This means that we can extract more work from entangled pairs than from classically correlated pairs.

Figuratively speaking, demons, Alice and Bob, who share entangled states with each other, can outdo those who have only classically correlated ones in terms of the amount of extractable work. By demons we mean here any fictitious entities that manipulate microscopic objects using information available to them, in analogy with Maxwell’s demon \cite{maxwell}. Their everlasting wish is to violate the Second Law of thermodynamics by extracting extra work from a heat bath thanks to entanglement, although this attempt turns out to be thwarted as we will see below.

The overall protocol to extract work from correlated pairs is depicted in Fig. 1. They first divide their shared ensemble into groups of two pairs to make the process symmetric with respect to Alice and Bob. For each group, they both choose a projection operator randomly and independently out of a set, $\{A_1, A_2, \cdots, A_n\}$ for Alice and $\{B_1, B_2, \cdots, B_n\}$ for Bob, just before their measurement. Then, Alice measures one of the two qubits in a group with the projector she chose and informs Bob of the outcome as well as her basis choice. Bob performs the same on his qubit of the other pair in the group. As a result of collective manipulations on the set of those groups for which they chose $A_i$ and $B_j$, they can extract $2 - H(A_i|B_j) - H(B_j|A_i)$ bits of work per two pairs at maximum.

III. THERMODYNAMICAL SEPARABILITY CRITERION

Let us find a general description of correlations in terms of work to clarify the difference between classical and quantum ones. It turns out that it suffices to sum up all the work that can be obtained by varying the basis continuously over a great circle on theBloch sphere, i.e., the circle of maximum possible size on a sphere (see Fig. 2). This is similar in approach to the chained Bell’s inequalities discussed in Ref. \cite{bell}. The circle should be chosen to maximize the sum.

Let $\xi_\rho(A_i, B_j)$ denote the extractable work from two copies of bipartite system $\rho$ in the asymptotic limit when Alice chooses $A_i$ and $A_i^\perp$ as her measurement and work-extracting basis while Bob chooses $B_j$ and $B_j^\perp$. For a two-dimensional bipartite system, $\xi$ is given by

$$\xi_\rho(A_i, B_j) := 2 - H(A_i|B_j) - H(B_j|A_i)$$

$$= 2 - 2H(A_i, B_j) + H(A_i) + H(B_j), \quad (1)$$

which is symmetric with respect to $A_i$ and $B_j$.

The quantity we consider is $\lim_{n \to \infty} 1/(2n - 1)[\sum_{k=1}^{n} \xi(A(\theta_k), B(\theta_k)) + \sum_{k=1}^{n-1} \xi(B(\theta_k), A(\theta_{k+1}))]$ for a state $\rho$ and we let $\Xi(\rho)$ denote it as

$$\Xi(\rho) := \frac{1}{2\pi} \int_0^{2\pi} \xi_\rho(A(\theta), B(\theta))d\theta, \quad (2)$$

where $\theta$ is the angle representing the direction of measurement on the great circle. The great circle is the one that maximizes the integral, as mentioned above. Note that $\Xi(\rho)$ represents the extractable work under local operations and classical communication, which is a standard framework to deal with entanglement in quantum information theory.

We now present an inequality that shows a connection between the thermodynamically extractable work and the separability of bipartite quantum systems.
An inequality

\[ \Xi(\rho) \leq \Xi(00) \]  

(3)
is a necessary condition for a two-dimensional bipartite state \( \rho \) to be separable, that is, \( \rho = \sum_i p_i \rho_i^A \otimes \rho_i^B \). The state |00⟩ in the right-hand side (RHS) can be any pure product state |ψ⟩|ψ⟩. We obtained the value of \( \Xi(00) \) numerically as 0.8854 bits. We will refer to this inequality as “thermodynamical separability criterion”.

**Proof.** Without loss of generality, all \( \rho_i^A \) and \( \rho_i^B \) can be assumed as pure states. The key point of the proof is that even if the information from the other side could always be used to specify the pure state component \( \rho_i^B \) (or \( \rho_i^A \)) on his/her side, the average extractable work \( \Xi(\rho) \) is always not larger than \( \Xi(00) \). This is because for any pure product state |ψ⟩|ψ⟩, \( \Xi(|ψ⟩|ψ⟩) \leq \Xi(00) \) with equality when both |ψ⟩ and |ψ⟩′ are on the integral path for \( \Xi \). Thus, \( \Xi^{\text{pcs}}(\rho) = \sum_i p_i \Xi(\rho_i^A \otimes \rho_i^B) \leq \sum_i p_i \Xi(00) = \Xi(00) \), where \( \Xi^{\text{pcs}}(\rho) \) is the extractable work from \( \rho \) when “pure component specification (pcs)” is possible. We show below that \( \Xi(\rho) \leq \Xi^{\text{pcs}}(\rho) \).

The conditional entropy in the definition of \( \xi_\rho(A,B) \) can be written as

\[ H(B|A) = \sum_j p_j \sum_i p_i \text{Tr}(A^i \rho_j^A) \text{Tr}(B^j \rho_i^B)H(p_i^B | p_j^A). \]

(4)

On the other hand, if the pcs is possible, the corresponding entropy can be written as

\[ H^{\text{pcs}}(B|A) = \sum_i p_i H(B^0 \rho_i^B). \]

(5)

By letting \( p_j^A \) denote \( \text{Tr}(A^j \rho_i^A) \) and noting \( p_j^A = \sum_i p_i p_j^A \), we can have an inequality as

\[ H(B|A) = \sum_j p_j H(\sum_i p_i \rho_j^A) \text{Tr}(B^j \rho_i^B) \leq \sum_i \sum_j p_j^A \rho_j^A H(p_i^B) \]

\[ = \sum_i p_i H(p_i^B) = H^{\text{pcs}}(B|A), \]

(6)
due to the concavity of Shannon entropy. As all \( p_j^A \) and \( p_i^B \) can be regarded as distinct, the equality in Eq. (6) holds when there exists only one pure component, i.e., \( p_k = 1 \) for a certain \( k \) and \( p_i = 0 \) for \( i \neq k \). Similarly, \( H(A|B) \geq H^{\text{pcs}}(A|B) \). Hence, for all separable, or classically correlated, states \( \rho \), \( \Xi(\rho) \leq \Xi^{\text{pcs}}(\rho) \) and thus \( \Xi(\rho) \leq \Xi(00) \). □

The dashed lines in Fig. 3 show numerically plotted \( \Xi(\sigma_{cl}) \) as a function of \( c_0 \), where \( \sigma_{cl} = c_0|00⟩⟨00| + c_1|\varphi⟩⟨\varphi| \) is a classically correlated state with respect to two vectors, \( |0⟩ \) and \( |\varphi⟩ := \cos(\varphi/2)|0⟩ + \sin(\varphi/2)|1⟩ \). There are five dashed lines, each of which corresponds to \( \sigma_{cl} \) with one \( \varphi \) out of the set \( \{0.2\pi, 0.4\pi, \cdots, \pi\} \). The near-horizontal one corresponds to \( \varphi = 0.2\pi \), which gives a state “close” to |00⟩. We can see that none of \( \Xi(\sigma_{cl}) \) exceeds \( \Xi(00) \), whose value is 0.8854 bits, as the above proposition claims. Also, above the threshold \( \Xi(00) \), the extractable work from a (pure) entangled state is a monotonic function of the amount of entanglement \( \alpha^2 \), taking its maximum when maximally entangled.

As the RHS of Eq. (3) represents the maximum amount of work obtainable from classically correlated states with our protocol, any excess of extractable work should be the manifestation of entanglement. Thus, violating the inequality \( \Xi(\rho) \leq \Xi(00) \), demons can exploit the extra work from entanglement, which is unavailable from classically correlated states.

Even if we choose a part of the great circle as the integral path of Eq. (2) without closing it, the inequality corresponding to Eq. (3) should be violated for states that are entangled strongly enough. This is because for a strongly entangled state \( |\phi⟩, \xi_\phi(A_\theta, B_\theta) \) is close to 2 for all \( \theta \), while \( \xi_{00}(A_\theta, B_\theta) \) is always less than 2 unless \( \theta \) indicates |0⟩. This also means that the violation criterion, i.e., the RHS of Eq. (3), depends on the range of the path in such a case.
We can also perform the integral in Eq. (2) over the whole Bloch sphere, instead of the great circle, to have another separability criterion. Then, Eq. (8) becomes

$$\Xi_{BS}(\rho) := \frac{1}{4\pi} \int_{BS} \xi_p(A, B) d\Omega \leq \Xi_{BS}(00),$$

(7)

where BS stands for the Bloch sphere and we obtained $$\Xi_{BS}(00) = 0.5573$$ numerically. The proposition also holds for $$\Xi_{BS}(\rho)$$ as there is no need to change the proof except that $$\Xi_{BS}^{\rho}$$ is always equal to $$\Xi_{BS}(00)$$ in this case. Let us now compute the value of $$\Xi_{BS}(\rho_W)$$, where $$\rho_W = p|\Psi+\rangle\langle\Psi+| + (1 - p)/4 \cdot I$$ is the Werner state [3], to see the extent to which the inequality can be satisfied when we vary $$p$$. It has been known that Bell-CHSH inequalities [13] are violated for $$p > 1/\sqrt{2} = 0.7071$$, while $$\rho_W$$ is inseparable iff $$p > 1/3$$. A bit of algebraic calculations lead to $$\Xi_{BS}(\rho_W) = (1 - p) \log_2(1 - p) + (1 + p) \log_2(1 + p)$$ and this is greater than $$\Xi_{BS}(00)$$ when $$p > 0.6006$$. Therefore, the inequality (7) is significantly stronger than Bell-CHSH inequalities.

Since we have obtained the inequalities (3) and (7) without discussing the non-locality of quantum mechanics, they are different from Bell’s inequalities, but similar to them in the sense that they discriminate non-classical correlations from classical ones. With reference to Bell’s inequalities, the formal of extractable work, Eq. (1), reminds us of the “information theoretic Bell’s inequalities” derived by Schumacher by defining the information distance using conditional entropies [14].

Schumacher obtained his inequalities from the triangle inequality for a metric. However, it is possible to derive the same inequalities directly from our definition of the extractable work, Eq. (1). The result becomes

$$\xi(A, B) + \xi(B, C) \leq 2 + \xi(A, C),$$

(8)

with equality when $$H(A) = 0$$ or $$H(C) = 0$$. This is a direct consequence of Eq. (1) and the strong subadditivity of the Shannon entropy. From Eq. (8), we can obtain “chained” inequalities

$$\xi(A_1, B_1) + \xi(B_1, A_2) + \cdots + \xi(A_n, B_n) \leq 2(2n - 2) + \xi(A_1, B_n).$$

(9)

Note that the left-hand side of Eq. (9) becomes the same as that of Eq. (3) if we take bases $$A_1, \ldots, A_n$$ as those in Fig. 2 divide the sum by $$2n - 1$$, and let $$n$$ tend to infinity. However, the same procedure on the RHS gives 2, thus Eq. (9) becomes a trivial inequality $$\Xi(\rho) \leq 2$$. Hence, our inequality (8) is essentially independent of Bell-Schumacher inequalities and gives a stronger bound on $$\Xi(\rho)$$ in the continuous limit.

IV. ANALYSIS OF DEMONS’ ATTEMPT

Let us analyze how demons’ attempt to violate the Second Law of thermodynamics will end in failure. In order to discuss the Second Law, which states “there does not exist any cycle of a heat engine that converts heat to work without leaving any change in its environment,” we need to restore the initial state after extracting work and check the balance of work. The impossibility of breaking the Second Law by the protocol described above follows from the fact that the net work investment is always nonnegative, regardless of the direction of measurement/work extraction, thus, it is necessarily nonnegative after averaging over the great circle or the Bloch sphere.

Suppose that Bob performs measurement on his part and Alice extracts work on her side. The state after work extraction is $$\sigma = I/2 \otimes |00\rangle$$, if Bob’s outcome was 1, he can flip it with $$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$ without energy consumption. Generically, unitary operations require no energy consumption as entropy stays constant with them. Let us consider first the case in which there was only classical consumption as entropy stays constant with them. Let

**FIG. 4:** Restoration of the initial state $$\rho^{cl}$$. It is simply a scheme to prepare $$nS(\rho^{cl})$$ of $$I/2$$ from $$n$$ copies of the state $$I/2 \otimes |00\rangle$$. The work, $$w^A_{\text{ext}}$$ and $$w^B_{\text{ext}}$$, are explained in the main text.

| n copies of $$I/2 \otimes |00\rangle$$ | Work consumed by Alice
|---|---|
| n copies of $$\rho^{cl}$$ | Quantum decompression |

work $$w^A_{\text{ext}}$$ extracted by Bob

- n copies of $$I/2 \otimes I/2$$
- n copies of $$I/2 \otimes |00\rangle$$
calculate the work investment $W_{\text{inv}}$ to close the cycle: 

$$W_{\text{cl, inv}} = w_{\text{con}}^A - w_{\text{ext}}^B + H(B(\theta)) - [1 - H(A(\theta)|B(\theta))] = H(A(\theta), B(\theta)) - S(\rho^{cl}) \text{ bits of work for an initial state of } \rho^{cl}$$

and, similarly, $W_{\text{ent, inv}} = H(A(\theta), B(\theta)) \text{ bits for an entangled state, } |\psi^{AB}\rangle$. These $W_{\text{inv}}$ must be nonnegative in order for the Second Law not to be violated. It turns out that both $W_{\text{cl, inv}}$ and $W_{\text{ent, inv}}$ are indeed nonnegative due to the properties of the Shannon and von Neumann entropies and the effect of projective measurements.

V. SUMMARY

We have devised a scenario in which two demons, Alice and Bob, can distill more work from entanglement than from classical correlation. We have cast this discrimination of correlations in the form of the thermodynamical separability criteria. Although we can re-derive Schumacher’s inequalities by considering a set of discrete basis of measurement, our inequalities are essentially different. Interestingly, our inequality (7) is more effective than Bell-CHSH inequalities, as well as Schumacher’s, in detecting inseparability of the Werner state. Lastly, our analysis of the energy balance after closing the thermodynamical cycle illustrates, quite expectedly, that the demons cannot violate the Second Law even with the excess work extracted from entangled states.

One important step we should make next is to devise a method to test the inequality experimentally, for example, by using NMR as proposed in Ref. [17]. Also, it would be very interesting if we can find a connection between our inequalities and the non-locality of quantum mechanics.

Acknowledgments

KM acknowledges financial support by Fuji Xerox. FM appreciates the warmth and hospitality of the quantum information group at Imperial College. VV is supported by the European Community, the Engineering and Physical Sciences Research Council, Hewlett-Packard and Elisag Spa.

[1] J. S. Bell, Physics (Long Island City, N.Y.) 1, 195 (1964).
[2] J. Oppenheim, M. Horodecki, P. Horodecki, and R. Horodecki, Phys. Rev. Lett. 89, 180402 (2002).
[3] W. H. Zurek, Phys. Rev. A 67, 012320 (2003); W. H. Zurek, Rev. Mod. Phys. 75, 715 (2003).
[4] B. M. Terhal, Phys. Lett. A 271, 319 (2000).
[5] R. F. Werner, Phys. Rev. A. 40, 4277 (1989).
[6] L. Szilard, Z. Phys. 53, 840–856 (1929).
[7] C. H. Bennett, Int. J. Theor. Phys. 21, 905 (1982).
[8] J. von Neumann, Mathematical Foundations of Quantum Mechanics, (English translation by R. T. Beyer, Princeton University Press, Princeton, 1955).
[9] H. S. Leff and A. F. Rex (eds.), Maxwell’s Demon 2, (Institute of Physics Publishing, Bristol and Philadelphia, 2003).
[10] S. L. Braunstein and C. M. Caves, Ann. Phys. 202, 22 (1990).
[11] Since we are focusing on qubits, the entropy $H(B|A)$ can be determined with a probability for only one outcome on Bob’s side, say 0, for instance. This is the same as taking $p = \text{Tr}(P_0 \rho)$ for the (binary) Shannon entropy $H(p)$ in Section [11].
[12] V. Vedral, M. B. Plenio, M. A. Rippin, and P. L. Knight, Phys. Rev. Lett. 78, 2275 (1997).
[13] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt, Phys. Rev. Lett. 23, 880 (1969).
[14] B. W. Schumacher, Phys. Rev. A. 44, 7047 (1991).
[15] B. Schumacher, Phys. Rev. A. 51, 2738 (1995).
[16] R. Landauer, IBM J. Res. Dev. 5, 183 (1961).
[17] S. Lloyd, Phys. Rev. A. 56, 3374 (1997).