

ARTICLE

Received 22 May 2014 | Accepted 27 May 2015 | Published 7 Jul 2015

DOI: 10.1038/ncomms8669

OPEN

The minimal work cost of information processing

Philippe Faist¹, Frédéric Dupuis^{1,2,3}, Jonathan Oppenheim⁴ & Renato Renner¹

Irreversible information processing cannot be carried out without some inevitable thermodynamical work cost. This fundamental restriction, known as Landauer's principle, is increasingly relevant today, as the energy dissipation of computing devices impedes the development of their performance. Here we determine the minimal work required to carry out any logical process, for instance a computation. It is given by the entropy of the discarded information conditional to the output of the computation. Our formula takes precisely into account the statistically fluctuating work requirement of the logical process. It enables the explicit calculation of practical scenarios, such as computational circuits or quantum measurements. On the conceptual level, our result gives a precise and operational connection between thermodynamic and information entropy, and explains the emergence of the entropy state function in macroscopic thermodynamics.

¹Institute for Theoretical Physics, ETH Zurich, Wolfgang-Pauli-Str. 27, Zurich 8093, Switzerland. ²Department of Computer Science, Aarhus University, IT-Parken, Aabogade 34, Denmark 8200. ³Faculty of Informatics, Masaryk University, Botanická 68a, Brno 612 00, Czech Republic. ⁴Department for Physics and Astronomy, University College of London, Gower Street, WC1E 6BT UK. Correspondence and requests for materials should be addressed to P.F (email: pfaist@phys.ethz.ch).

Thermodynamics in essence is an information theory—its purpose is to make statements about systems for which we only have certain partial information, such as a gas of many particles for which only macroscopic quantities like temperature, volume and pressure are accessible. Following this point of view, Jaynes showed that the entropy function derived in statistical mechanics corresponds to the information-theoretic entropy of the gas associated with a macroscopic observer who is maximally ignorant of the microscopic degrees of freedom¹, resorting to Shannon’s mathematical theory of information² developed in the context of telecommunications.

When the observers have access to knowledge about microscopic quantities, such as positions and velocities of particles in a gas, the second law of thermodynamics seems to break down, as was illustrated by Maxwell’s demon. To address this problem, Szilard³ studied a one-particle gas that can be located on either side of a box, left (‘L’) or right (‘R’), and noted that by isothermally compressing the gas or letting the gas expand, one can trade this one bit of information for $kT \ln 2$ work, as depicted in Fig. 1a (in the presence of a heat bath at temperature T , and where k is Boltzmann’s constant). Landauer and Bennett later realized that the information content of data stored in a memory register, independently of the nature of its physical representation, counts as thermodynamic entropy when considering thermodynamical operations on that register^{4–13}. For example, given a bit in an unknown state, any operation that resets it to zero must dissipate at least $kT \ln 2$ heat, and thus the corresponding amount of work must be supplied (this is known as Landauer’s principle). This fact salvages the second law of thermodynamics and resolves the paradox of Maxwell’s demon.

More recently with the advent of quantum information, efforts were made to understand the laws of quantum thermodynamics

from an information-theoretic viewpoint^{14–18}, while the increasing technological ability to control and manipulate nanoscale systems^{19,20} has prompted the study of particular operational models and frameworks, leading to characterization of the work cost of various information-theoretic tasks such as erasure and work extraction^{21–32}. For a more specific review of existing results, we refer to (Supplementary Note 1).

The aim of this work is to study thermodynamics in such generalized scenarios, where one may have knowledge about microscopic degrees of freedom, by resorting to modern tools of information theory^{33,34}. We provide a fundamental lower bound to the work cost of a physical implementation of a logical process, discuss several examples and illustrate how traditional thermodynamics emerges from our microscopic result in the limit of macroscopic systems.

Results

The Framework. We determine a general expression for the minimal amount of work needed to carry out any given logical process \mathcal{E} . This can be for example an AND gate or any quantum or classical computation; most generally \mathcal{E} is defined as any completely positive, trace-preserving map from quantum states on an input Hilbert space H_X to quantum states on an output Hilbert space $H_{X'}$. We assume these spaces to be of finite dimension for simplicity; note that such a space can be a subspace of an infinite-dimensional Hilbert space in which the relevant computation or logical process takes place. The terminology ‘logical process’ is meant to emphasize that the mathematical object \mathcal{E} only specifies for each input state the corresponding output state and does not prescribe its physical realization, which would consist of a full description of a physical system including the parts of its environment that are relevant to determine its time evolution. Note that in performing a logical process one does not merely transform one quantum state into another; rather, the output must be related to the input in a precisely specified way. In the case where the input is a classical value, this means that the output depends on the particular input value received, and not only on the distribution of inputs. This might be checked in practice, for example, if one keeps a copy of the input as a reference system and observes the correlations between the output and the reference system.

There are, generally, many ways of actually realizing a logical process with an actual physical device. The device and its interactions with the environment (for example, a heat bath) may for example be described by a Hamiltonian or a Liouvillian. For our purposes, it is sufficient to specify the set of operations which the device is allowed to perform as well as the associated work cost. We then optimize the work expenditure over precisely those strategies, which realize the given logical process \mathcal{E} . Observe that the more permissive our framework is, the more robust our bound will be. In our model, we shall be allowed to implement at no work cost any trace-preserving completely positive map that is unital, that is, which preserves the identity operator. Note that if we were to allow any logical process that is not unital to be performed for free, one could flagrantly violate the second law of thermodynamics on a macroscopic scale: in this sense, unital maps are the most permissive logical operation that we can allow for free. The model must also include a description of a ‘battery’ that provides the energy required to drive the process. For this we resort to Bennett’s idea of an ‘information fuel tape’^{5,11}: such a battery consists of a large number of qubits with a degenerate Hamiltonian. Initially, a certain number λ_1 of these qubits are in the maximally mixed state and the rest are pure. We may freely implement any joint unital map on the system and battery. At the end of the operation, the state of the battery consists of a possibly different number λ_2 of qubits in a maximally mixed state, while

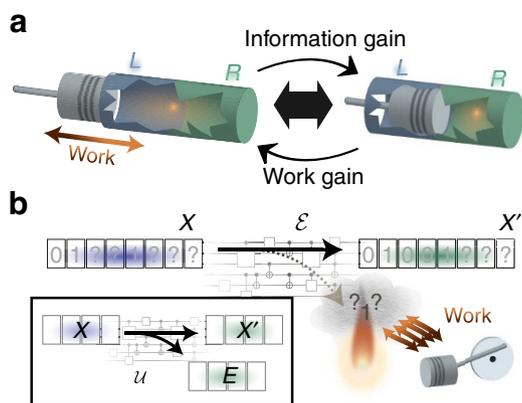


Figure 1 | Work and information are related by physical processes.

(a) A gas formed by a single particle can either be on the left (‘L’) or the right (‘R’) side of the cylinder (known as a Szilard box³). This one bit of information can be reversibly traded for $kT \ln 2$ work by isothermally compressing the gas with a piston or letting the gas isothermally expand. This illustrates that discarding 1 bit of entropy (or uncertainty) requires $kT \ln 2$ work. (b) An implementation of the logical process \mathcal{E} mapping a system X to an output X' interacts with the thermal bath may discard information and in general costs work. The logical process \mathcal{E} may be written as part of a global unitary U on an additional hypothetical system E , which represents the discarded information (inset). Our main result states that the minimum work required in a physical implementation of \mathcal{E} is the amount of discarded information, which the implementation has to dump into the environment.

the rest should be pure (The requirement that these λ_2 qubits be maximally mixed is not a restriction, see Methods section.). We then count the amount of work consumed as $W = kT \ln 2 \cdot (\lambda_2 - \lambda_1)$, which is the amount of work required to restore the battery system into its initial state. Indeed, a vast amount of literature has well underscored the correspondence between possessing a pure degenerate qubit, or storing $kT \ln 2$ work, and vice versa^{3,5,11,12}. The quantity W may be negative, indicating that work can be extracted from the battery when restoring it to its initial state. In addition, we assume that the input to the logical process \mathcal{E} is encoded in a system whose initial Hamiltonian is degenerate. The same is assumed about the output system at the end of the computation. Note that this does not exclude making use of systems with nontrivial Hamiltonians during the implementation of the process. Also, this requirement is in practice not a limitation, as many other frameworks may be mapped to this setting^{26,28,29}; indeed the assumption should rather be regarded as a technicality to ensure a clean way of accounting for work.

To obtain physically relevant results, we also have to exclude overwhelmingly unlikely events from our considerations. This is actually quite common in thermodynamics and is usually done implicitly. For example, consider a stone lying on the ground. There is a very small chance that by thermal fluctuation the stone spontaneously jumps in the air. However, this event is so disproportionately unlikely that in a physical theory we may safely choose to ignore this possibility. Within our framework, we do this more explicitly. That is, we consider a parameter that specifies the total probability of all events we want to exclude. In the quantum regime, where events are generally not well-defined, this idea is captured by ϵ -approximations: the stone has a very small amplitude of being found in the air, but its state is ϵ -close to a state completely located on the ground. Analogously, we study the work requirement of logical processes that are ϵ -approximations of the desired logical process. This is a standard procedure in information theory^{33,34}, and is justified by the fact that an ϵ -approximation cannot be distinguished from the original logical process with probability greater than ϵ .

The main result. To formulate our main claim, we represent the logical process \mathcal{E} by its Stinespring dilation³⁵. This is an isometry \mathcal{U} (which can be seen as part of a unitary) that maps X onto X' as well as an extra system E such that the original map \mathcal{E} is retrieved by ignoring E (see Fig. 1b). Our main result asserts that W^ϵ , the work one needs to supply to execute the operation up to an ϵ -approximation, is lower bounded by

$$W^\epsilon \geq kT \ln(2) H_{\max}^\epsilon(E|X'). \tag{1}$$

The right hand side is the smooth max-entropy of E conditioned on X' and may be interpreted as a measure for the irreversibility of the logical process. More precisely, the smooth max-entropy is an information-theoretic measure defined in the Methods section, and quantifies the uncertainty one has about E when given access to X' . The parameters ϵ and $\bar{\epsilon}$ are related by $\bar{\epsilon} = \sqrt{2\epsilon}$; ϵ may be chosen arbitrarily. We stress that the system E is an abstract mathematical concept used to represent the logical map \mathcal{E} , and can be interpreted as the information discarded by the mapping. In particular, our bound is independent of the choice of this representation.

The form of the bound (1) naturally expresses our intuition that the amount of work that needs to be provided corresponds to the amount of information that is logically discarded, and which therefore has to be dumped into the environment. This consideration is done from the viewpoint of the observer who has completed the computation, and thus has access to X' ,

explaining the occurrence of the conditional entropy. Also, if E is classical, the max-entropy has the operational interpretation of being the amount of memory space needed to compress the information contained in E when possessing knowledge of X' (ref. 36) (In the fully quantum case, it corresponds to quantum state merging³⁷).

The proof of our main result proceeds by first considering the special case in which $\epsilon = 0$. The bound one then obtains is

$$W^{\epsilon=0} \geq kT \ln 2 \cdot \log_2 \|\mathcal{E}(\Pi_X)\|_\infty, \tag{2}$$

where Π_X is the projector onto the support of the input state. This expression proves particularly useful for calculating some simple practical examples.

The proof of this special case, and its generalization to the regime where $\epsilon > 0$, is presented in the Methods section. An alternative proof, using techniques from majorization, is given in (Supplementary Note 5).

Classical mappings and dependence on the logical process. Our result, which is applicable to arbitrary quantum processes, applies to all classical computations as a special case. Classically, logical processes correspond to stochastic maps, of which deterministic functions are a special case. As a simple example, consider the AND gate. This is one of the elementary operations computing devices can perform, from which more complex circuits can be designed. The gate takes two bits as input, and outputs a single bit that is set to 1 exactly when both input bits are 1, as illustrated in Fig. 2a.

The logical process is manifestly irreversible, as the output alone does not allow to infer the input uniquely. If one of the inputs is zero, then the logical process effectively has to reset a three-level system to zero, forgetting which of the three possible inputs 00, 01 or 10 was given; this information can be viewed as being discarded, and hence dumped into the environment. We can confirm this intuition with our main result, using the fact

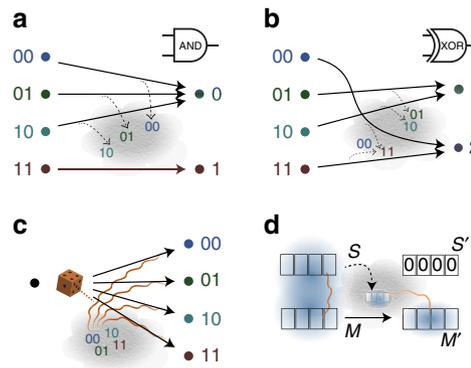


Figure 2 | Examples of logical processes. (a) The AND gate is one of the building blocks of computers. Our result implies that any successful implementation of this logically irreversible gate requires at least work $\log_2 3 \cdot kT \ln 2 \approx 1.6 kT \ln 2$ due to the entropy of the discarded information (dotted arrows). (b) The XOR gate only requires $kT \ln 2$ work, as it discards less entropy per output event than the AND gate. (c) Work can be extracted if randomness is being produced: the discarded information is entangled with the output (orange wavy lines), and the conditional entropy on the right hand side of (1) is negative. (d) The erasure of a quantum system S with access to a quantum memory M must transfer the content of S into the system E containing the discarded information, while preparing S' in a pure state and mapping M to M' identically. The corresponding minimal work cost is $kT \ln(2) \cdot H_{\max}^\epsilon(S|M)$; this can be achieved using the procedure of del Rio *et al.*²⁵ If the system is entangled with the memory, this quantity is negative and work may be extracted.

that a general classical mapping is given by the specification of the conditional probability $p(x'|x)$ of observing x' at the output if the input was x . Embedding the classical probability distributions into the diagonals of quantum states, the infinity norm in expression (2) becomes simply

$$W^{\epsilon=0} \geq kT \ln 2 \cdot \log_2 \max_{x'} \sum_x p(x'|x), \quad (3)$$

where the sum ranges only over those x that have a non-zero probability of occurring. In the case of deterministic mappings $p(x'|x) \in \{0,1\}$, this corresponds to the maximum number of input states that map to a same output state. For the AND gate, provided all four states 00, 01, 10 and 11 have non-negligible probability of occurring, there are three input states mapping to the same output state, so (3) gives us simply $W_{\text{AND}}^{\epsilon=0} \geq \log_2 3 \cdot kT \ln 2 \approx 1.6kT \ln 2$. Also, in simple examples as considered here, the expression (3) is stable to considering an ϵ -approximation (Supplementary Note 4); this quantity is thus physically justified.

Crucially, our result reveals that the minimal work requirement in general depends on the specific logical process, and not only on the input and output states. This contrasts with traditional thermodynamics for large systems, where the minimal work requirement of a state transformation can always be written as a difference of a thermodynamical potential, such as the free energy. For example, the minimal work cost of performing specifically an AND gate may differ from that of another logical process mapping an input distribution $(p_{00}, p_{01}, p_{10}, p_{11})$ (with $\sum_i p_i = 1$) to the distribution $(p'_0, p'_1) = (p_{00} + p_{01} + p_{10}, p_{11})$ (Recall that the classical counterpart of a quantum state is a probability distribution.). To see this, consider the XOR gate, which outputs a 1 exactly when both inputs are different (see Fig. 2b). The minimal work cost requirement of this gate, as given by (3), is now only $kT \ln 2$, as in the worst case, only a single bit of information is erased (again supposing that all four input states have non-negligible probability of occurring). Now, suppose that, for some reason, the input distribution is such that $p_{01} + p_{10} = p_{11}$, that is, the input 11 occurs with the same probability as of either 01 or 10 appearing. Then, the XOR gate reproduces the exact same output distribution as the AND gate: in both cases, we have $p'_0 = p_{00} + p_{10} + p_{01} = p_{00} + p_{11}$ and $p'_1 = p_{11} = p_{01} + p_{10}$. In other words, both logical processes have the same input and output state, yet the XOR gate only requires work $kT \ln 2$ compared with the AND gate, which requires $1.6kT \ln 2$. Furthermore, we point out that this difference, which appears small in this case, may be arbitrarily large in certain scenarios (Supplementary Note 4).

On the one hand, we are by definition interested in the work cost of a given logical process, so one might have expected that this work cost should not only depend on the input and output states. On the other hand, it might seem contradictory that the full logical process matters even though we have fixed an input state σ_X . However, this makes sense if we consider preparing the input state as part of a pure state on the input system and a reference system. In this case, the logical process that is implemented influences the (in principle detectable) correlations between the output and the reference system, even if the reduced state on the input is the fixed state σ_X .

We emphasize that the phenomenon observed here is fundamentally different from the notion of thermodynamic irreversibility. Here we always consider the optimal procedure for implementing the logical process, whereas a thermodynamically irreversible process is in fact an 'inefficient' physical process that could be replaced by a more efficient, reversible one. In our framework, the thermodynamically irreversible processes are those physical implementations that do not achieve the

bound (1). A longer discussion with examples is provided in (Supplementary Note 2).

Work extraction. While erasure requires work, it is well known that in a wide range of frameworks one can in general extract work with the reverse logical process, which corresponds to taking a register of bits that are all in the zero state and making them maximally mixed^{3,5}. Our result intrinsically reproduces this fact: the Stinespring dilation $\mathcal{U}_{X \rightarrow X'E}$ of a logical process that generates randomness in fact creates entanglement between the output X' and E (see Fig. 2c). The conditional entropy $H_{\text{max}}^{\epsilon}(E|X')$ then becomes negative, such that the bound (1) allows work to be extracted. We remark that, even if the logical process $\mathcal{E}_{X \rightarrow X'}$ is classical, the relevant state for the entropic term in (1) is entangled, and thus all but classical; this is due to the construction of E as a purifying system for the logical process.

Erasure with a quantum memory and tightness of our bound. Recently, del Rio *et al.*²⁵ have constructed an explicit procedure capable of resetting a quantum system S to a pure state using an erasure mechanism assisted by a quantum memory M , and doing so at a work cost of approximately

$$W_{\text{erasure}}^{\epsilon} \approx kT \ln 2 \cdot H_{\text{max}}^{\epsilon}(S|M). \quad (4)$$

The approximation holds up to terms of the order of the logarithm of ϵ and are negligible in typical scenarios (Supplementary Note 4).

Our main result implies that their procedure is nearly optimal (Fig. 2d). Indeed, consider the total system $X = S \otimes M$, in the initial state σ_{SM} , with the logical process $\mathcal{E}(\cdot) = \text{tr}_S(\cdot) \otimes |0\rangle\langle 0|_S$, denoting symbolically with a prime the output system S' (The state on M remains unchanged.). One then straightforwardly sees that the resulting joint state on E and the output $X' = S' \otimes M$ is obtained from the initial state on S and M by isometrically 'transferring' the S part to E and replacing it by a fixed pure state. The entropy term in our bound (1) then becomes $H_{\text{max}}^{\epsilon}(E|S'M) = H_{\text{max}}^{\epsilon}(E|M) = H_{\text{max}}^{\epsilon}(S|M)$, the latter entropy being evaluated on the input state. This matches the term in (4).

Conversely, this optimal erasure procedure can be used to show that for any arbitrary logical process, the minimal amount of work our result associates to it can be in principle achieved to good approximation. Given a logical process \mathcal{E} and an input state σ_X , calculate its Stinespring dilation $\mathcal{U}_{X \rightarrow X'E}$ as explained above, and consider an ancillary system A_E of the same dimension as E . This ancilla system is initialized in a pure state $|0\rangle_{A_E}$. One can then carry out a unitary $\mathcal{U}'_{XA_E \rightarrow X'A'_E}$ on X and A_E , chosen such that

$$\mathcal{U}'_{XA_E \rightarrow X'A'_E} \left(\sigma_X \otimes |0\rangle\langle 0|_{A_E} \right) = \mathcal{U}_{X \rightarrow X'A'_E}(\sigma_X). \quad (5)$$

In effect, A'_E impersonates the abstract system E while we perform a unitary corresponding to the Stinespring dilation of \mathcal{E} (see inset of Fig. 1b). This unitary operation can be implemented at no work cost because it is reversible. The aforementioned optimal erasure procedure can then be used to restore the ancilla A'_E to its original pure state, using the output system X' as the quantum memory, at a work cost of approximately $kT \ln(2) \cdot H_{\text{max}}^{\epsilon}(A'_E|X')$. As A'_E corresponds to E , this matches our bound (1) and therefore proves its tightness.

The work requirement of a quantum measurement. The problem of determining the amount of work needed to carry out a quantum measurement has been the subject of much literature^{38–40}, especially in the context of Maxwell's demon^{5,6,12,41}. A quantum measurement is a logical process (depicted in Fig. 3a) acting on a system X to be measured and a classical register C initially set to a pure state, and outputting systems C' and X' , with

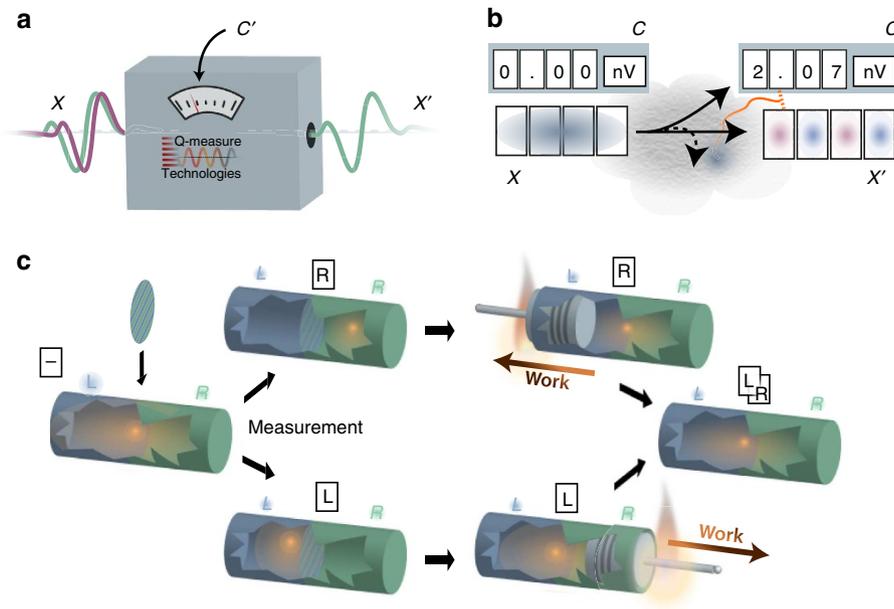


Figure 3 | The work cost of quantum measurements. (a) A quantum measurement may be thought of as a device which produces a post-measurement state X' and a classical reading C' from an input state X . (b) The corresponding logical process maps the input system X and classical register C to a classical outcome on the output register C' and a post-measurement state on X' . The initial register C is prepared in a pure state. Our main result implies that the measurement costs no work in principle. (c) Maxwell's demon with a Szilard box, as proposed by Bennett⁵. A measurement detects on which side of the inserted separator the particle is, and extracts work with a piston in either case. The cylinder is left in its original state, apparently creating a perpetuum mobile with net work gain. However, the measurement outcome (represented by 'L' or 'R') had to be stored in a memory register, which was initially in some pure state (represented by '-') and the work cost of resetting it to a pure state again compensates the work gain. The register could have been reset using the post-measurement state at no work cost, but the latter was consumed during work extraction.

C' containing the measurement result and X' the quantum post-measurement state. We will consider a projective measurement for simplicity, treating the more general case in (Supplementary Note 4). The logical process corresponding to the measurement described by a complete set of projectors $\{P_i\}_i$ takes the form

$$\mathcal{E}_{CX \rightarrow C'X'}(\sigma_{CX}) = \sum_i |i\rangle\langle i|_{C'} \otimes (P_i \sigma_X P_i). \quad (6)$$

Our bound (2) for this map is at most zero (since $\|\mathcal{E}(|0\rangle\langle 0|_C \otimes \Pi_X)\|_\infty \leq 1$), implying that the measurement can be carried out in principle at no work cost, as was already stated by Bennett⁵. Note that a work cost is required if the classical register C was not initially pure⁴⁰.

A related question is the work cost of erasing the information contained in the register C' after the measurement. Doing so would allow us to construct a cycle. The cost of this erasure can be reduced using the post-measurement state as a quantum memory, by employing the procedure presented above, to $kT \ln 2 \cdot H_{\max}^c(C'|X')$. But because C' and X' may only be classically correlated, no work may be extracted in this way²⁵. In some cases this work cost may be zero, for example for projective measurements on a maximally mixed state (Supplementary Note 4). This might seem to save Maxwell's demon from Bennett's information-theoretic exorcism, which argues that the demon must pay work to reset its memory⁵ (see Fig. 3c). However, the key point is to notice that the demon cannot use the post-measurement state to both extract work and to reset its internal memory register.

Discussion

Our main result exposes various features of thermodynamics in the microscopic regime that are not present in the standard

setting of large systems. In particular, as argued above, the minimum work cost of a logical process cannot be given in terms of a state function, such as the entropy or the free energy in thermodynamics.

Traditional thermodynamics is concerned with macroscopic systems, and we may retrieve this limit by considering logical processes that consist of many individual operations. Under appropriate independence assumptions and using typicality arguments⁴², one can show that the average minimal work cost per process as determined by (1) simply takes the form $kT \ln(2) \cdot [H(X) - H(X')]$, where $H(X) = -\text{tr}(\rho_X \log_2 \rho_X)$ is the usual von Neumann entropy (see Methods section): the minimal work requirement is now given by a function of state $H(X)$, and no longer depends on the logical process that maps X to X' (see Methods).

Our result thus provides the following fresh view on the macroscopic regime. Thermodynamics can be seen as a general framework, in which the second law postulates the existence of a state function, the thermodynamic entropy, which relates to the heat flow in processes. Many standard results of thermodynamics follow from that starting point. It is now the role of a microscopic theory to construct a state function with this property, based on the microscopic dynamics of the particular system. In textbook statistical mechanics, this construction is given for several physical setups, such as gases or lattices; one usually considers, for example, the configuration entropy, or an appropriately normalized Shannon or von Neumann entropy of the density of the statistical ensemble. Our result generalizes this construction and clarifies when it is justified: the state function, in general, appears whenever the inherent fluctuations due to the microscopic stochastic nature of the process vanish by typicality. The existence of an entropy state function is therefore not a

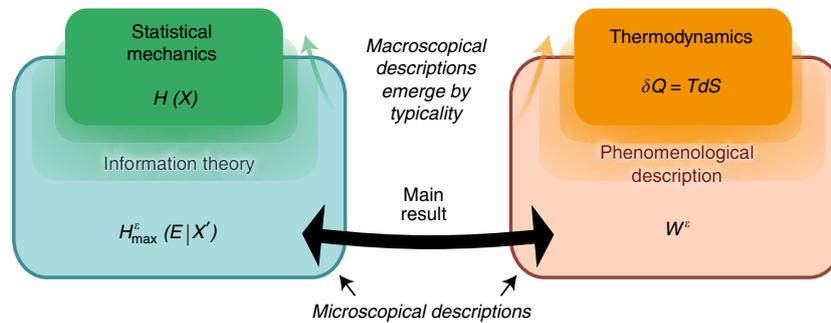


Figure 4 | Relation between information-theoretic and thermodynamic quantities. Our result relates two quantities that depend on the microscopic details of the system: the information-theoretic entropy $H_{\max}^e(E|X')$ that quantifies the amount of information discarded by the logical process, and the amount of work W^e needed to carry out a logical process on the microscopic level. Standard thermodynamics is obtained in the limit of macroscopic systems. In this limit, it follows from typicality arguments that the entropic measure H_{\max}^e converges to the von Neumann entropy $H(X)$, which may thus be seen as an emergent quantity. Furthermore, in this regime, the minimum amount of work W^e used by a process corresponds to the heat Q that is reversibly transferred to the environment, which in turn is related to the thermodynamic entropy, S , as defined by Clausius. Our result thus permits the identification of the information-theoretic entropy $H(X)$ for a macroscopic observer, that is, the entropy considered in statistical mechanics, with the thermodynamic entropy S .

property of the microscopic system; it is rather an emergent quantity that appears whenever the full system is typical, such as in the limit of macroscopic processes (Fig. 4).

Finally, one should note that the system in consideration need not be large for the typicality arguments to apply. For example, if one considers the work requirement of performing many independent repetitions of a single given logical process (seen as one big joint process), then the work requirement W^e per repetition converges to the average work requirement as calculated via statistical mechanics, even if the individual system is small: in this case, the entropy function emerges. This further justifies the usage of the von Neumann entropy in statistical mechanics even for small systems. Conversely, a large system does not necessarily display typicality; such is the case for systems out of thermodynamic equilibrium. An explicit example is provided in (Supplementary Note 4).

In summary, our main result quantifies the minimal required work to perform a logical process on the microscopic level. On the conceptual level, our result shows how, for macroscopic systems, the information-theoretic von Neumann entropy emerges as a state function and can thus be strictly identified with the thermodynamic entropy.

Methods

Mathematical formulation and proof of the main result. The task is to implement the logical process $\mathcal{E}_{X \rightarrow X'}$. Recall the framework allows for the implementation of any unital map, that is, $\mathcal{E}(1) = 1$, to be performed on the systems at hand. We first adapt a well-known classical result about doubly stochastic and doubly sub-stochastic matrices⁴³ to relate unital quantum maps to so-called subunital maps, that is, maps $\tilde{\mathcal{E}}$ that satisfy $\tilde{\mathcal{E}}(1) \leq 1$. Note also that the composition of two unital maps is unital, and similarly the composition of two subunital maps is subunital. We will need the following proposition, which we prove in (Supplementary Note 6) as Prop. 17.

Proposition 1 (dilation of a subunital map). Let H_K and H_L be finite dimensional Hilbert spaces, and let $\tilde{\mathcal{E}}_{K \rightarrow L}$ be a completely positive, trace-nonincreasing, subunital map. Then there exists finite dimensional Hilbert spaces H_Q and $H_{Q'}$, and a completely positive, trace-preserving, unital map $\tilde{\mathcal{E}}_{KQ \rightarrow LQ'}$ such that

$$\tilde{\mathcal{E}}_{K \rightarrow L}(\cdot) = \left(\mathbb{1}_L \otimes (f|_{Q'}) \right) \tilde{\mathcal{E}}_{KQ \rightarrow LQ'} \left[\left(\mathbb{1}_K \otimes |i\rangle_Q \right) (\cdot) \left(\mathbb{1}_K \otimes \langle i|_Q \right) \right] \left(\mathbb{1}_L \otimes |f\rangle_{Q'} \right), \quad (7)$$

for some pure states $|i\rangle_Q, |f\rangle_{Q'}$. In addition, $\dim(H_K \otimes H_Q) = \dim(H_L \otimes H_{Q'})$.

Let's now denote by A the 'information battery' system, which is the physical system that tracks how much work we have used or extracted. The system A may be as large as we might wish (but finite) and starts in a state $2^{-\lambda_1} \mathbb{1}_{2^{\lambda_1}} \otimes |0\dots 0\rangle\langle 0\dots 0|$ with some given number of mixed qubits λ_1 . The system X starts in a given state σ_X , and we assume that the Hamiltonians of X and A vanish at the beginning and at the end of the physical process.

Our framework specifies that we are allowed to perform any sequence of joint unital operations on any subsystems of X and A . The final state on $X' \otimes A'$ should be a product state, with the state on A' of the form $2^{-\lambda_2} \mathbb{1}_{2^{\lambda_2}} \otimes |0\dots 0\rangle\langle 0\dots 0|$. Note

that the structure imposed on this state is not a restriction: if the final state on A' is not of this form, an additional unital map can be applied on the support of the final state on A' to replace the latter by a maximally mixed state on its support. However, this condition does assume that there is no way to extract work while transforming a state ρ to a maximally mixed state of the same rank, or, equivalently, that the worst-case erasure cost of a state ρ is $kT \ln 2 \log_2 \text{rank } \rho$. This can usually be seen as a consequence of the choice of framework, and is in line with the findings of refs 28, 29. Alternatively, given a state ρ , let m be its rank, p_{\min} its smallest non-zero eigenvalue and Π the projector on its support. The state ρ may be written as a statistical mixture of $\frac{1}{m} \mathbb{1}_m$ with probability $m \cdot p_{\min}$ and some state $(\rho - p_{\min} \Pi) / (1 - m \cdot p_{\min})$ with probability $1 - m \cdot p_{\min}$. In the event where the system is prepared in the maximally mixed state of rank m , the work requirement for erasure is deterministic because the state is uniform, and equals $kT \ln 2 \log_2 m$ (refs 3,5,11,12); it follows that the work required for erasing ρ with certainty is at least $kT \ln 2 \log_2 \text{rank } \rho$.

Observe that our framework is equivalent to allowing the agent to perform a single unital operation on the whole of X and A , leaving both systems in the state $\rho_{X'} \otimes (2^{-\lambda_2} \mathbb{1}_{2^{\lambda_2}} \otimes |0\dots 0\rangle\langle 0\dots 0|)_{A'}$: indeed the composition of unital maps is unital, and extending a unital map by an identity map still yields a unital map.

Even though we have presented our results while hinting that X and X' represent the same system, and are thus of the same dimension, this need not be the case: our results are valid for arbitrary finite dimensions of X and X' . However, we will assume that one can bring in ancillas of arbitrary finite dimension in pure states and dispose of ancillas restored to a pure state for free. Henceforth, we will assume that such ancillas are counted as part of the pure systems composing the work storage systems A and A' (The systems A and A' hence need not be of same dimension.).

We must in addition require that the physical process implement the logical process \mathcal{E} . Let $|\sigma\rangle_{XR}$ be a purification of σ_X on a system R . If one applies the physical process to X while leaving R untouched, then the state on $X' \otimes R$ that results from the physical process must be equal to the state $\rho_{X'R}$ that would result by applying the mapping $\mathcal{E} \otimes \text{id}_R$ on σ_{XR} , that is, $\rho_{X'R} = \mathcal{E} \otimes \text{id}_R(\sigma_{XR})$. Observe that this constraint is equivalent to requiring the logical mapping corresponding to the physical process to be exactly \mathcal{E} on the support of σ_X , due to the Choi-Jamiołkowski isomorphism. So, even with a fixed given input state σ_X , the full information about the mapping can be observed in the resulting state on $X' \otimes R$, by keeping a purification of σ_X : in other words, the full information about the mapping and the input state is one-to-one encoded in the bipartite state $\rho_{X'R}$.

Let's now state a formal version of our problem, in the case where we do not yet consider an ϵ -approximation. The task is to find the minimal $kT \ln 2 \cdot (\lambda_2 - \lambda_1)$, such that there exists a unital, trace-preserving, map $\mathcal{E}_{XA \rightarrow X'A'}$ satisfying

$$\tilde{\mathcal{E}}_{XA \rightarrow X'A'}(\sigma_{XR} \otimes (2^{-\lambda_1} \mathbb{1}_{2^{\lambda_1}})) = \rho_{X'R} \otimes (2^{-\lambda_2} \mathbb{1}_{2^{\lambda_2}}), \quad (8)$$

where $\rho_{X'R} = \mathcal{E}(\sigma_{X'R})$ and where an identity mapping on R is implicitly understood (We henceforth omit the pure states on system A , that is, the factors $|\otimes|0\dots 0\rangle\langle 0\dots 0|$ above, for readability.).

At this point, note that whenever for given λ_1, λ_2 , there is such a unital map, then there is also a subunital map achieving the same logical process and vice versa. Let's write this as a proposition:

Proposition II. Let $\lambda_1, \lambda_2 \geq 0$ and let \mathcal{E} be given. Then are equivalent

(1) For a large enough A , and corresponding A' , there exists a trace-preserving unital map $\tilde{\mathcal{E}}_{AX \rightarrow A'X'}$ such that

$$\tilde{\mathcal{E}}_{AX \rightarrow A'X'}(\sigma_{XR} \otimes (2^{-\lambda_1} \mathbb{1}_{2^{\lambda_1}})_A) = \rho_{X'R} \otimes (2^{-\lambda_2} \mathbb{1}_{2^{\lambda_2}})_{A'}; \quad (9)$$

(2) For a large enough B , and large enough B' , there exists a trace-nonincreasing subunital map $\tilde{\mathcal{E}}_{XB \rightarrow X'B'}$ such that

$$\tilde{\mathcal{E}}_{XB \rightarrow X'B'}(\sigma_{XR} \otimes (2^{-\lambda_1} \mathbb{1}_{2^{\lambda_1}})_B) = \rho_{X'R} \otimes (2^{-\lambda_2} \mathbb{1}_{2^{\lambda_2}})_{B'}. \quad (10)$$

Proof. The forward direction is straightforward, as a unital map is in particular subunital. For the converse, we will dilate the given subunital map $\tilde{\mathcal{E}}_{XB \rightarrow X'B'}$ to a unital map using Prop. 1, with $H_K = H_X \otimes H_B$ and $H_L = H_{X'} \otimes H_{B'}$: let $H_Q, H_{Q'}$ and $\tilde{\mathcal{E}}_{XQ \rightarrow X'Q'} = \tilde{\mathcal{E}}_{XBQ \rightarrow X'B'Q'}$ be given by the Proposition. Now define $H_A = H_B \otimes H_Q$ and $H_{A'} = H_{B'} \otimes H_{Q'}$. We would like to show that $\tilde{\mathcal{E}}_{XA \rightarrow X'A'}(\sigma_{XR} \otimes (2^{-\lambda_1} \mathbb{1}_{2^{\lambda_1}})_A) = \rho_{X'R} \otimes (2^{-\lambda_2} \mathbb{1}_{2^{\lambda_2}})_{A'}$, where we have defined $(2^{-\lambda_1} \mathbb{1}_{2^{\lambda_1}})_A = (2^{-\lambda_1} \mathbb{1}_{2^{\lambda_1}})_B \otimes |i\rangle\langle i|_Q$ and $(2^{-\lambda_2} \mathbb{1}_{2^{\lambda_2}})_{A'} = (2^{-\lambda_2} \mathbb{1}_{2^{\lambda_2}})_{B'} \otimes |f\rangle\langle f|_{Q'}$ (as pure states, $|i\rangle_Q$ and $|f\rangle_{Q'}$ do not alter the amount of work stored in the work storage systems A and A'). Define also the shorthand $\tilde{E}_{X'B'Q'R} := \tilde{\mathcal{E}}_{XBQ \rightarrow X'B'Q'}(\sigma_{XR} \otimes (2^{-\lambda_1} \mathbb{1}_{2^{\lambda_1}})_B \otimes |i\rangle\langle i|_Q)$. By construction, and using (7), we have

$$\left(\mathbb{1}_{X'B'R} \otimes \langle f|_{Q'} \right) \tilde{E}_{X'B'Q'R} \left(\mathbb{1}_{X'B'R} \otimes |f\rangle_{Q'} \right) = \rho_{X'R} \otimes (2^{-\lambda_2} \mathbb{1}_{2^{\lambda_2}})_{A'}. \quad (11)$$

Since $\tilde{\mathcal{E}}$ is trace-preserving, we have $\text{tr}(\tilde{E}_{X'B'Q'R}) = 1$ and

$$\text{tr} \left[\left(\mathbb{1}_{X'B'R} \otimes (1 - |f\rangle\langle f|)_{Q'} \right) \tilde{E}_{X'B'Q'R} \right] = 1 - \text{tr} \left[\left(\mathbb{1}_{X'B'R} \otimes |f\rangle\langle f|_{Q'} \right) \tilde{E}_{X'B'Q'R} \right] = 0,$$

as the expression in (11) has unit trace. It follows that $\tilde{E}_{X'B'Q'R}$ lies in the support of $\mathbb{1}_{X'B'R} \otimes |f\rangle\langle f|_{Q'}$, and from (11) we conclude as requested that

$$\tilde{E}_{X'A'R} = \rho_{X'R} \otimes (2^{-\lambda_2} \mathbb{1}_{2^{\lambda_2}})_{A'}.$$

We can now characterize the allowed operations in our framework and their work costs with the following proposition.

Proposition III. Let $\sigma_X, \mathcal{E}_{X \rightarrow X'}$ be given. Choose system B big enough and let be given integers $\lambda_1, \lambda_2 \geq 0$. Then are equivalent:

(1) There exists a trace-nonincreasing subunital map $\tilde{\mathcal{E}}_{XB \rightarrow X'B'}$ such that

$$\tilde{\mathcal{E}}_{XB \rightarrow X'B'}(\sigma_{XR} \otimes (2^{-\lambda_1} \mathbb{1}_{2^{\lambda_1}})_B) = \rho_{X'R} \otimes (2^{-\lambda_2} \mathbb{1}_{2^{\lambda_2}})_{B'};$$

(2) There exists a trace-nonincreasing map \mathcal{T} , mapping linear operators on H_X to linear operators on $H_{X'}$, such that $\mathcal{T}_{X \rightarrow X'}(\mathbb{1}) \leq 2^{-(\lambda_1 - \lambda_2)} \mathbb{1}$, and $\mathcal{T}_{X \rightarrow X'}(\sigma_{XR}) = \rho_{X'R}$;

(3) The map $\tilde{\mathcal{E}}_{XB \rightarrow X'B'}$ satisfies $\|\mathcal{E}(\Pi_X)\|_\infty \leq 2^{-(\lambda_1 - \lambda_2)}$, where Π_X is the projector onto the support of σ_X .

Proof. (i) \Rightarrow (ii): Define $\mathcal{T}_{X \rightarrow X'}(\cdot) = \text{tr}_B[\mathbb{1}_{2^{\lambda_2}} \tilde{\mathcal{E}}_{XB \rightarrow X'B'}(\cdot \otimes (2^{-\lambda_1} \mathbb{1}_{2^{\lambda_1}}))] \mathbb{1}_{2^{\lambda_2}}$. Then, $\mathcal{T}(\sigma_{XR}) = \text{tr}_{B'}[\mathbb{1}_{2^{\lambda_2}} (\mathcal{E}(\sigma_{XR}) \otimes (2^{-\lambda_2} \mathbb{1}_{2^{\lambda_2}}))] \mathbb{1}_{2^{\lambda_2}} = \rho_{X'R}$. Also, $\mathcal{T}(\mathbb{1}) = \text{tr}_{B'}[\mathbb{1}_{2^{\lambda_2}} \tilde{\mathcal{E}}_{XB \rightarrow X'B'}(\mathbb{1}_X \otimes (2^{-\lambda_1} \mathbb{1}_{2^{\lambda_1}}))] \mathbb{1}_{2^{\lambda_2}} \leq 2^{-\lambda_1} \text{tr}_{B'}[\mathbb{1}_{2^{\lambda_2}} \tilde{\mathcal{E}}_{XB \rightarrow X'B'}(\mathbb{1}_{XA} \mathbb{1}_{2^{\lambda_1}})] \leq 2^{\lambda_2 - \lambda_1} \mathbb{1}_X$, because $\tilde{\mathcal{E}}$ is subunital.

(ii) \Rightarrow (iii): We have $\mathcal{E}(\Pi_X) = \mathcal{T}(\Pi_X)$ because the maps are equal on the support of ρ_X (alternatively, operate $\text{tr}_R[(\cdot) \rho_R^{-1}]$ on both sides of $\mathcal{T}(\sigma_{XR}) = \rho_{X'R} = \mathcal{E}(\sigma_{XR})$ noting that $\rho_R = \sigma_R$); then because $\Pi_X \leq \mathbb{1}_X$, we have $\|\mathcal{E}(\Pi_X)\|_\infty \leq \|\mathcal{T}(\Pi_X)\|_\infty \leq 2^{-(\lambda_1 - \lambda_2)}$.

(iii) \Rightarrow (i): Let $\tilde{\mathcal{E}}_{XA \rightarrow X'A'}(\cdot) = \mathcal{E}(\text{tr}_B[(\Pi_X \otimes \mathbb{1}_{2^{\lambda_1}})(\cdot)(\Pi_X \otimes \mathbb{1}_{2^{\lambda_1}})]) \otimes (2^{-\lambda_2} \mathbb{1}_{2^{\lambda_2}})$. Observe that $\tilde{\mathcal{E}}$ is subunital: $\tilde{\mathcal{E}}_{XA \rightarrow X'A'}(\mathbb{1}_{XA}) = \mathcal{E}(\text{tr}_B[\Pi_X \otimes \mathbb{1}_{2^{\lambda_1}}]) \otimes (2^{-\lambda_2} \mathbb{1}_{2^{\lambda_2}}) \leq 2^{\lambda_1 - \lambda_2} \mathcal{E}(\Pi_X) \otimes \mathbb{1}_{2^{\lambda_2}} \leq \mathbb{1}_{X'A'}$. Also, $\tilde{\mathcal{E}}_{XA \rightarrow X'A'}(\sigma_{XR} \otimes (2^{-\lambda_1} \mathbb{1}_{2^{\lambda_1}})) = \mathcal{E}(\sigma_{XR}) \otimes (2^{-\lambda_2} \mathbb{1}_{2^{\lambda_2}})$, because the input to $\tilde{\mathcal{E}}$ is inside the support of $\Pi_X \otimes \mathbb{1}_{2^{\lambda_1}}$. Hence, $\tilde{\mathcal{E}}$ satisfies the conditions of (i).

With these propositions, we can calculate straightforwardly and explicitly the minimization in the formulation of the main problem. It now reduces to the simple question of minimizing $\lambda_2 - \lambda_1$ subject to $\|\mathcal{E}(\Pi_X)\|_\infty \leq 2^{\lambda_2 - \lambda_1}$; we have thus proven (2).

Entropic form of the bound. Some basic facts about the smooth entropy framework are necessary to understand the rest of this section. For a more complete introduction on the smooth entropy framework, we refer to (Supplementary Note 3).

An equivalent definition of the Rényi-zero conditional entropy, also known as alternative max-entropy, for a bipartite state ρ_{AB} , is given as

$$H_0(A|B)_\rho = \log_2 \|\text{tr}_A \Pi_{AB}\|_\infty, \quad (12)$$

where Π_{AB} is the projector on the support of ρ_{AB} . For consistency with the standard literature, we will express our final result in terms of the max-entropy, which is related to the Rényi-zero entropy up to factors logarithmic in ϵ (ref. 34). The non-smooth conditional max-entropy can be defined as

$$H_{\max}(A|B) = \max_{\omega_B} \log F^2(\rho_{AB}, \mathbb{1}_A \otimes \omega_B), \quad (13)$$

where $F(\rho_1, \rho_2) = \|\rho_1^{1/2} \rho_2^{-1/2}\|_1$ is the fidelity between two quantum states³⁵, and where the optimization ranges over density operators on B . The smooth conditional max-entropy is defined by ‘smoothing’ the max-entropy on states that are ϵ -close to ρ_{AB} in fidelity distance:

$$H_{\max}^\epsilon(A|B)_\rho = \min_{\hat{\rho} \approx_\epsilon \rho} H_{\max}(A|B)_{\hat{\rho}}, \quad (14)$$

where the minimization ranges over all $\hat{\rho}$ such that $F^2(\rho_{AB}, \hat{\rho}_{AB}) \geq 1 - \epsilon^2$.

Let’s now return to our bound (2). Consider the Stinespring dilation of \mathcal{E} , given by an isometry $V_{X \rightarrow X'E}$ including an additional system $E : \mathcal{E}(\cdot) = \text{tr}_E(V_{X \rightarrow X'E}(\cdot)V^\dagger)$. Defining the pure state $\rho_{X'ER} = V\sigma_{XR}V^\dagger$ is obviously compatible with our previous definition of $\rho_{X'R}$, as $\text{tr}_E \rho_{X'ER} = \mathcal{E}(\sigma_{XR})$. It follows that $\mathbb{V}\Pi_X V^\dagger = \Pi_{X'E}$, where $\Pi_{X'E}$ is the projector on the support of $\rho_{X'E}$. Recalling (12), we have

$$\|\mathcal{E}(\Pi_X)\|_\infty = \left\| \text{tr}_E \mathbb{V}\Pi_X V^\dagger \right\|_\infty = \|\text{tr}_E \Pi_{X'E}\|_\infty = 2^{H_0(E|X')_\rho}, \quad (15)$$

and our bound (2) takes the form

$$W^{\epsilon=0} \geq kT \ln 2 \cdot H_0(E|X')_\rho. \quad (16)$$

Considering an ϵ -approximation. A ‘smooth’ version of the result is straightforward to obtain. In this case, we allow the actual process to not implement precisely \mathcal{E} , but only approximate it well. The best strategy to detect this inexactness is to prepare $|\sigma\rangle_{XR}$ and send σ_X into the process, and then perform a measurement on $\rho_{X'R}$. To ensure that the approximate process is not distinguishable from the ideal process with probability greater than ϵ , we require that the trace distance between the ideal output of the process $\rho_{X'R}$ and the actual output $\hat{\rho}_{X'R}$ must not exceed ϵ . We can apply our main result to the approximate process that brings σ to $\hat{\rho}$, and lower bound the work cost of that process by

$$\begin{aligned} W(\sigma \rightarrow \hat{\rho}) &\geq H_0(E|X')_{\hat{\rho}} \cdot kT \ln 2 \\ &\geq H_{\max}(E|X')_{\hat{\rho}} \cdot kT \ln 2, \end{aligned} \quad (17)$$

where the second inequality is shown in ref. 44. This relaxation of H_0 to H_{\max} is done for the sake of presentation and consistency with other results within the smooth entropy framework. When smoothing with a parameter ϵ , there is no significant difference with this relaxation: indeed, the two quantities are equivalent up to adjustment of the ϵ parameter and up to a logarithmic term in ϵ (Lemma 18 of ref. 44).

If we optimize (17) over all possible maps \mathcal{T} that output such $\hat{\rho}_{X'R}$, we obtain a bound on the work requirement of the ϵ -approximation,

$$\begin{aligned} W &\geq \min_{\hat{\rho}_{X'R} \approx_\epsilon \rho_{X'R}} H_{\max}(E|X')_{\hat{\rho}} \cdot kT \ln 2 \\ &\geq \min_{\hat{\rho}_{X'RE} \approx_\epsilon \rho_{X'RE}} H_{\max}(E|X')_{\hat{\rho}} \cdot kT \ln 2 \\ &= H_{\max}^\epsilon(E|X')_\rho \cdot kT \ln 2, \end{aligned} \quad (18)$$

where the first optimization ranges over all $\hat{\rho}_{X'R}$ such that the trace distance $\frac{1}{2} \|\hat{\rho}_{X'R} - \rho_{X'R}\|_1 \leq \epsilon$, and where the second optimization ranges over all $\hat{\rho}_{X'RE}$ such that $F^2(\rho_{X'RE}, \hat{\rho}_{X'RE}) \geq 1 - \bar{\epsilon}^2$, with $\bar{\epsilon} = \sqrt{2\epsilon}$, where $F(\rho, \hat{\rho}) = \|\sqrt{\rho} \sqrt{\hat{\rho}}\|_1$ is the fidelity between the quantum states ρ and $\hat{\rho}$.³⁵

Macroscopic limit: many independent repetitions. As we have seen in the introduction, considerable previous work has focused on the limit cases where many i.i.d. systems are provided. In such a case, the process $\mathcal{E}^{\otimes n}$ is applied on n independent copies of the input $\sigma^{\otimes n}$, and outputs $\rho^{\otimes n}$. A smoothing parameter $\epsilon > 0$ is chosen freely. We may simply apply our (smoothed) main result to get an expression for our bound on the work cost,

$$W \geq H_{\max}^\epsilon(E^n|X^n)_{\rho^{\otimes n}} \cdot kT \ln 2, \quad (19)$$

However, it is known that the smooth entropies converge to the von Neumann entropy in the i.i.d. limit⁴²,

$$\lim_{\epsilon \rightarrow 0} \lim_{n \rightarrow \infty} \frac{1}{n} H_{\max}^\epsilon(E^n|X^n)_{\rho^{\otimes n}} = H(E|X)_\rho, \quad (20)$$

which allows us to simplify the expression of the work cost per particle, or per repetition of the process, to

$$H(E|X)_\rho = H(EX)_\rho - H(X)_\rho = H(X)_\sigma - H(X)_\rho,$$

where the last equality holds because ρ_{EX} and σ_X have the same spectrum being both purifications of the same $\rho_R = \sigma_R$. We conclude that in the asymptotic i.i.d. case, the work cost is simply given by the difference of entropy between the initial and final state,

$$W \geq [H(\text{initial state}) - H(\text{final state})] kT \ln 2. \quad (21)$$

Here W is the average work cost per particle, or per repetition of the process. In the case for example of many independent particles undergoing a similar, independent process, the total work W required is obtained by considering the entropy of the full system of all particles in both terms in (21).

References

- Jaynes, E. T. Information theory and statistical mechanics. *Phys. Rev.* **106**, 620–630 (1957).
- Shannon, C. E. A mathematical theory of communication. *Bell Syst. Tech. J.* **27**, 379–423 (1948).

3. Szilard, L. Über die Entropieverminderung in einem thermodynamischen System bei Eingriffen intelligenter Wesen. *Zeitschrift für Physik* **53**, 840–856 (1929).
4. Landauer, R. Irreversibility and heat generation in the computing process. *IBM J. Res. Dev.* **5**, 183–191 (1961).
5. Bennett, C. H. The thermodynamics of computation—a review. *Int. J. Theor. Phys.* **21**, 905–940 (1982).
6. Bennett, C. H. Notes on Landauer's principle, reversible computation and Maxwell's demon. *Stud. Hist. Phil. Mod. Phys.* **34**, 501–510 (2003).
7. Bennett, C. H. Logical reversibility of computation. *IBM J. Res. Dev.* **17**, 525–532 (1973).
8. Shizume, K. Heat generation required by information erasure. *Phys. Rev. E* **52**, 3495–3499 (1995).
9. Maruyama, K., Nori, F. & Vedral, V. Colloquium: the physics of Maxwell's demon and information. *Rev. Mod. Phys.* **81**, 1–23 (2009).
10. Piechocinska, B. Information erasure. *Phys. Rev. A* **61**, 062314 (2000).
11. Feynman, R. P. *Lectures on Computation* (Westview Press, 1996).
12. Leff, H. S. & Rex, A. F. *Maxwell's Demon 2: entropy, classical and quantum information, computing* (Taylor & Francis, 2010).
13. Sagawa, T. & Ueda, M. in *Nonequilibrium Statistical Physics of Small Systems: Fluctuation Relations and Beyond*. (eds Klages, R., Just, W., Jarzynski, C. & Schuster, H. G.) 181–211 (Wiley-VCH, 2013).
14. Lloyd, S. Ultimate physical limits to computation. *Nature* **406**, 1047–1054 (2000).
15. Plenio, M. B. & Vitelli, V. The physics of forgetting: Landauer's erasure principle and information theory. *Contemp. Phys.* **42**, 25–60 (2001).
16. Popescu, S., Short, A. J. & Winter, A. Entanglement and the foundations of statistical mechanics. *Nat. Phys.* **2**, 754–758 (2006).
17. Gemmer, J., Michel, M. & Mahler, G. *Quantum thermodynamics: Emergence of thermodynamic behavior within composite quantum systems* (Springer Verlag, 2009).
18. Oppenheim, J., Horodecki, M., Horodecki, P. & Horodecki, R. Thermodynamical approach to quantifying quantum correlations. *Phys. Rev. Lett.* **89**, 180402 (2002).
19. Hänggi, P. & Marchesoni, F. Artificial brownian motors: Controlling transport on the nanoscale. *Rev. Mod. Phys.* **81**, 387–442 (2009).
20. Baugh, J., Moussa, O., Ryan, C. A., Nayak, A. & Laflamme, R. Experimental implementation of heat-bath algorithmic cooling using solid-state nuclear magnetic resonance. *Nature* **438**, 470–473 (2005).
21. Alicki, R., Horodecki, M., Horodecki, P. & Horodecki, R. Thermodynamics of quantum information systems—hamiltonian description. *Open Syst. Inf. Dyn.* **11**, 205–217 (2004).
22. Janzing, D. *Computer Science Approach to Quantum Control* (Habilitation, Universität Karlsruhe, 2006).
23. Linden, N., Popescu, S. & Skrzypczyk, P. How small can thermal machines be? the smallest possible refrigerator. *Phys. Rev. Lett.* **105**, 130401 (2010).
24. Dahlsten, O. C. O., Renner, R., Rieper, E. & Vedral, V. Inadequacy of von Neumann entropy for characterizing extractable work. *New J. Phys.* **13**, 53015 (2011).
25. del Rio, L., Åberg, J., Renner, R., Dahlsten, O. & Vedral, V. The thermodynamic meaning of negative entropy. *Nature* **474**, 61–63 (2011).
26. Egloff, D., Dahlsten, O. C. O., Renner, R. & Vedral, V. Laws of thermodynamics beyond the von Neumann regime. Preprint at <http://arxiv.org/abs/1207.0434> (2012).
27. Brandão, F. G. S. L., Horodecki, M., Oppenheim, J., Renes, J. M. & Spekkens, R. W. Resource theory of quantum states out of thermal equilibrium. *Phys. Rev. Lett.* **111**, 250404 (2013).
28. Åberg, J. Truly work-like work extraction via a single-shot analysis. *Nat. Commun.* **4**, 1925 (2013).
29. Horodecki, M. & Oppenheim, J. Fundamental limitations for quantum and nanoscale thermodynamics. *Nat. Commun.* **4**, 2059 (2013).
30. Skrzypczyk, P., Short, A. J. & Popescu, S. Extracting work from quantum systems. Preprint at <http://arxiv.org/abs/1302.2811> (2013).
31. Reeb, D. & Wolf, M. M. An improved Landauer principle with finite-size corrections. *N. J. Phys.* **16**, 103011 (2014).
32. Brandão, F., Horodecki, M., Ng, N., Oppenheim, J. & Wehner, S. The second laws of quantum thermodynamics. *Proc. Natl Acad. Sci. USA* **112**, 201411728 (2015).
33. Renner, R. *Security of Quantum Key Distribution*. Ph.D. thesis, ETH Zürich (2005). Security of Quantum Key Distribution. Preprint at <http://arxiv.org/abs/quant-ph/0512258> (2005).
34. Tomamichel, M. *A Framework for Non-Asymptotic Quantum Information Theory*. Ph.D. thesis, ETH Zurich (2012). A Framework for Non-Asymptotic Quantum Information Theory Preprint at <http://arxiv.org/abs/1203.2142> (2012).
35. Nielsen, M. A. & Chuang, I. L. *Quantum Computation and Quantum Information* (Cambridge University Press, 2000).
36. König, R., Renner, R. & Schaffner, C. The operational meaning of min- and max-entropy. *IEEE Trans. Inf. Theory* **55**, 4337–4347 (2009).
37. Horodecki, M., Oppenheim, J. & Winter, A. Partial quantum information. *Nature* **436**, 673–676 (2005).
38. Sagawa, T. & Ueda, M. Minimal energy cost for thermodynamic information processing: Measurement and information erasure. *Phys. Rev. Lett.* **102**, 250602 (2009).
39. Buscemi, F., Hayashi, M. & Horodecki, M. Global information balance in quantum measurements. *Phys. Rev. Lett.* **100**, 210504 (2008).
40. Jacobs, K. Quantum measurement and the first law of thermodynamics: The energy cost of measurement is the work value of the acquired information. *Phys. Rev. E* **86**, 040106 (2012).
41. Earman, J. & Norton, J. D. Exorcist XIV: The wrath of Maxwell's demon. part II. from Szilard to Landauer and beyond. *Stud. Hist. Phil. Sci.* **30**, 1–40 (1999).
42. Tomamichel, M., Colbeck, R. & Renner, R. A fully quantum asymptotic equipartition property. *IEEE Trans. Inf. Theory* **55**, 5840–5847 (2009).
43. Bhatia, R. *Matrix Analysis* (Springer, 1997).
44. Tomamichel, M., Schaffner, C., Smith, A. & Renner, R. Leftover hashing against quantum side information. *IEEE Trans. Inf. Theory* **57**, 5524–5535 (2011).

Acknowledgements

We thank Johan Åberg, Francesco Buscemi, Lea Krämer Gabriel Joe Renes, Lídia del Rio and Paul Skrzypczyk for discussions. P.F., F.D. and R.R. were supported by the Swiss National Science Foundation (SNSF) through the National Center of Competence in Research 'Quantum Science and Technology', through grant No. 200020-135048, and by the European Research Council through grant No. 258932. F.D. was also supported by the SNSF through grants PP00P2-128455 and 20CH21-138799, as well as by the German Science Foundation (grant CH 843/2-1). J.O. is funded by the Royal Society of London. This work was also supported by the COST Action MP1209.

Author contributions

The main ideas were developed by all authors. P. F. wrote the manuscript.

Additional information

Supplementary Information accompanies this paper at <http://www.nature.com/naturecommunications>

Competing financial interests: The authors declare no competing financial interests.

Reprints and permission information is available online at <http://npg.nature.com/reprintsandpermissions/>

How to cite this article: Faist, P. *et al.* The minimal work cost of information processing. *Nat. Commun.* 6:7669 doi: 10.1038/ncomms8669 (2015).



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>