Maxwell's daemon in the 21st century

Oscar Dahlsten

with Aaberg, del Rio, Egloff, Renner, Rieper and Vedral

oscar.dahlsten@physics.ox.ac.uk

oscar.dahlsten@nus.edu.sg

Grenoble, 29 June 2012

ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

UNIVERSITY OF OXFORD

NUS
National University of Singapore

Why I study Maxwell's daemon

1. Maxwell's daemon uses information to extract work from a heat bath. The tension between the subjectiveness of information and the objectiveness of energy has a paradoxical feel. Paradoxes are a good place to start research.

2. Understanding the fundamental limits to heat engines is of immense importance for the wider world.

   The power densities of typical integrated circuits are approaching those of a light bulb filament ($\sim 100 \frac{W}{cm^2}$). Removal of the heat generated by an integrated circuit has become perhaps *the* crucial constraint on the performance of modern electronics*.

3. Quantum information can bring a rather clear and sophisticated understanding of entropy to the table here, and entropy is crucial in thermodynamics.

4. Landauers and Bennett's early arguments we will soon recap are indeed used to guide nano-electronics*.

* *MIT Open course on nano-electronics.*

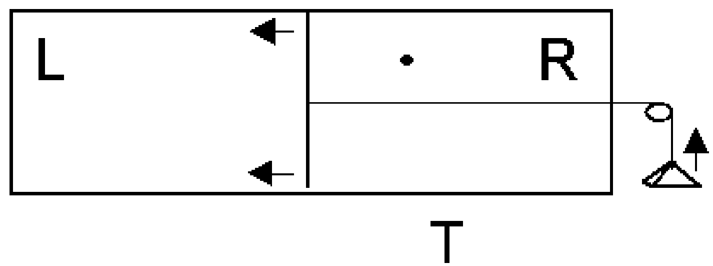
Overview

- Ideal gas Szilard engine Maxwell's daemon
- 2-level quantum Szilard engine.
- Demise of von Neumann entropy
- Single-shot information theory
- Single-shot thermodynamics
- Thermodynamic meaning of negative entropy

The daemon in the 20th century

Ideal gas Szilard engine

- There is a single particle in a box, and heat bath at temperature T.



- The daemon (also called agent) inserts a divider in the middle of the box, measures the position of the divider and hooks up the weight accordingly.

- It can extract work isothermally:

$$W = \int_V^{2V} p dV = \int_V^{2V} \frac{kT}{V} dV = kT \ln 2,$$

where we used the ideal gas equation $pV = NkT$.

- Here we used up one bit to gain $kT \ln 2$ of work from heat bath. The inverse process is **Landauer's principle**: *it costs at least $kT \ln 2$ of work to erase (reset) one bit* (just change integration limits around).

Role of entropy in Szilard's engine

- More general than ideal gas equation:

$$W = \Delta(U - TS).$$

  U is average internal energy of the working medium (e.g. particle), S is the standard entropy, and T the temperature of the heat bath.

- In the case of the Szilard engine $U = K.E. = \frac{1}{2}kT$ (by equipartition theorem), so $\Delta U = 0$. Thus

$$W = T\Delta S = Tk \ln 2.$$

- If one puts a quantum particle in the box the energies actually change both when the divider is inserted/extracted as well as when it is moved, so some care has to be taken here.

- It is often convenient to do calculations in terms of the partition function $Z = \sum_i \exp(-E_i\beta)$ where $\{E_i\}_{i=0}^{i_{\max}}$ are the energy values. ($\beta = 1/kT$). One can show

$$W = \Delta(U - TS) = kT \ln \frac{Z_f}{Z_i}.$$

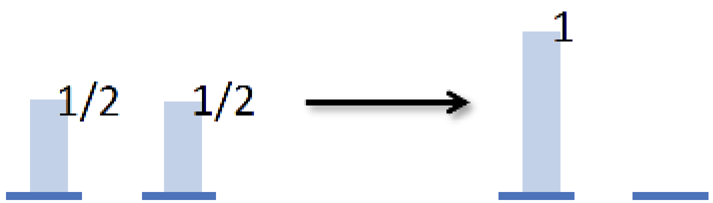
Daemon violates second law?

Kelvin's version of second law.

> No process is possible in which the sole result is the absorption of heat from a reservoir and its complete conversion into work.

- After the Szilard engine work extraction the working medium is restored to its original state, seemingly violating Kelvin's law, **unless** there is a hidden work cost not accounted for.
- Bennett points out that the daemon may correlate its memory with the particle by a reversible interaction (CNOT gate) in principle at no energy cost. **Measurement does not need to cost work**.
- But, Bennett argues, the entropy of the memory is increased by 1 bit and by Landauer's principle it costs at least $kT \ln 2$ work to reset it. Bennett thus located the hidden work cost which saves Kelvin's law.

A neat two-level quantum Landauer



[Alicki et. al.]

$$W = \Delta U = 0 - T\Delta S = Tk\ln 2,$$

- How? Raise the second level to infinity quasistatically and isothermally, so that the system is constantly in its thermal state

$$\rho_T = \sum_i \frac{\exp(-\beta E_i)}{Z}|e_i\rangle\langle e_i|$$

- If level is occupied when raised by δE_2, it costs δE_2 energy.

$$\langle dW\rangle = p(E_1)0 + p(E_2)dE_2 = p(E_2)dE_2 = \frac{\exp(-\beta E_2)}{Z}dE_2.$$

$\Rightarrow W = \int_0^\infty \frac{\exp(-\beta E_2)}{Z}dE_2 = kT\ln 2$, (integration requires some few lines).

Many-cylinder Szilard engines and quantifying information

- Bennett also considers N Szilard engines operated together. The agent/daemon's knowledge is represented by a probability distribution over $\{L, R\}^N$.
- Bennett showed knowledge of correlations can indeed be exploited to extract work.
- Say N=2 and [p(LL), p(LR), p(RL), p(RR)]=[1/2, 0, 0, 1/2]. Now p(L)=1/2 on each engine. But if the agent does a CNOT it gets [1/2, 0, 1/2, 0].
- Now it can extract work from the second bit, which is definitely L. The randomness has been compressed onto S bits, and one may think more generally that $W = (N - S)kT \ln 2$.

within Bennett's framework. Suppose we have a tape with *N* bits. We *define* the information, I, in the tape by the formula:

$$\textit{Fuel value of tape} = (N-I).kT \log 2. \quad (5.12)$$

[Feynman lectures on computation]

Mysteries for pub conversations

- I treat entropy as subjective in that different agents may have different amounts of knowledge about a system. Entropy is an observer-system quantity, not an objective quantity associated with a system.
- Thus, according to the above expressions, the *extractable* work is subjective. (The *extracted* work is not).
- A funny objection I read to the idea that entropy is subjective even in thermodynamics is "but $\delta S = \delta Q/T$ so if entropy is subjective I can change the temperature of a system by looking at it, this is absurd." Do you think this objection makes sense?
- To my knowledge one can very consistently think of entropy as subjective also in thermodynamics (see arguments by Jaynes for more on this long-running debate).

The daemon in the light of 21st century information theory

The demise of the von Neumann entropy

- The previous examples posed no challenge to the Shannon/von Neumann entropy $S_v N = \sum_i p_i \log_2 \frac{1}{p_i}$ being the appropriate entropy for thermodynamics (in quantum case $p_i \to \lambda_i$). They all involved a distribution which was flat on some events and zero on the rest.

- For such distributions many entropy measures coincide. In particular the *Renyi entropies*,

$$S_\alpha := \frac{\log \sum_i p_i^\alpha}{1 - \alpha}$$

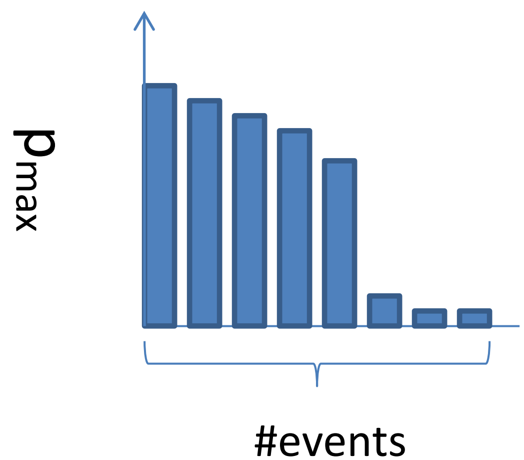
  all give the same value for such distributions, namely $S = \log d$ where d is the number of events with non-zero probability.

- Note $S_{vN} = S_1$. Two other crucial ones are S_0 and S_∞.
  Because $S_\infty \leq S_\alpha \leq S_0 \, \forall \alpha$, $S_0 := S_{\max}$, and $S_\infty := S_{\min}$.

- In quantum information theory Renner et. al. have shown that it is (modified versions of) $S_{\max}$ and $S_{\min}$ which capture the operational meanings one wants, and S_{vN} is only relevant when it happens to coincide with them.

Single-shot information theory, operational meaning of entropy

- One can show $S_0 := S_{\max} = \log(\sharp\text{events possible})$, and $S_\infty := S_{\min} = \log(\frac{1}{p_{\max}})$
- An important operational meaning of S_0 is that it is the size a memory needs to have to reliable retain information. Not on average in some sense but in a given experiment. It is called a single-shot entropy.

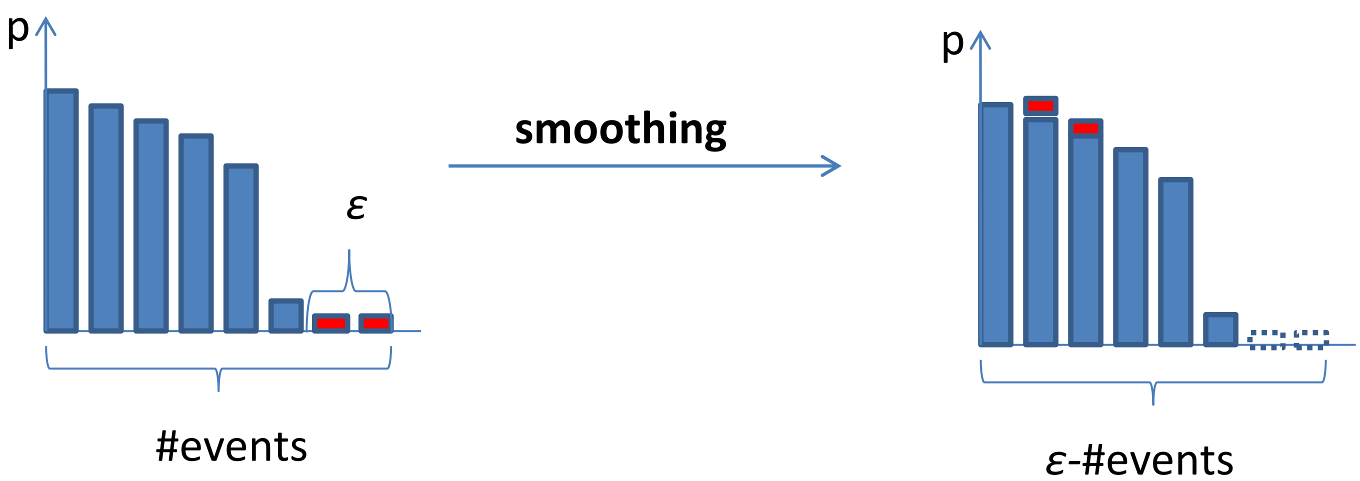


Smooth entropies

- An important technique is to **smooth** a distribution before evaluating the entropy. This gives the **smooth entropy**

  $H^{\varepsilon}_{\max}(\rho) := \min H_{\max}(\rho') \, | d(\rho, \rho') \leq \varepsilon$

  If d(.,.) is the trace distance the smoothing looks like this.



- Interpretation: the entropy is effectively $H^{\varepsilon}_{\max}(\rho)$.
- One reason smoothing important: $\lim(n \to \infty, \varepsilon \to 0) H^{\varepsilon}_{\max}(\rho^{\otimes n}) = nH(\rho)$

Single shot thermodynamics -the basic idea

What is the optimal work gain (or cost) W to effect a given change of a system's state, in any given realization, for a maximum error probability ε? (This is not a question about average work.)

$$W^{\varepsilon} = ?$$

Think of lifting a weight onto a table. We want to know with what probability it will reach the table in every extraction, not whether it will on average reach the table.

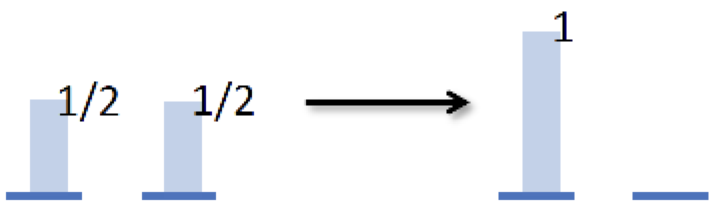
Single-shot statistical mechanics is heavily inspired by single-shot information theory, also known as the smooth entropy approach.

We have used (generalised) Szilard engines as a bridge between information theory and statistical mechanics.

Papers to date on this approach

Title	Authors	arXiv	Journal
Inadequacy of von Neumann entropy for characterizing extractable work	Dahlsten, Renner, Rieper, Vedral	0908.0424	NJP
Thermodynamic meaning of negative entropy	del Rio, Aaberg, Renner, Dahlsten, Vedral	1009.1630	Nature
Truly work-like work extraction	Aaberg	1110.6121	
Fundamental limitations for quantum and nano thermodynamics	Horodecki, Oppenheim	1111.3834	
The laws of thermodynamics beyond the von Neumann Regime	Egloff, Dahlsten, Renner and Vedral	1207.0434	

2-level case



- To understand the idea of W^ε it is instructive to again consider that protocol in the 2-level case.
- Once the second level is very high the probability of being in level 2 is very low, call it ε.
- Except with probability ε it then suffices to input W^ε of work in order to take the initial state to the final.
- If you calculate this you will find $W^\varepsilon = -kT(\ln 2 + \ln(1-\varepsilon)) \leq W^0$. W^0 is the 0-risk work cost of this protocol.

Most general statement to date-the game

- I like to think of work extraction as a game: given the rules of the game you want to find the strategy which optimises the work output.

- A very general game which all the things I have discussed fit into is to say that you have some specified initial state ρ and final state σ. There is also an initial energy spectrum E and final spectum F.

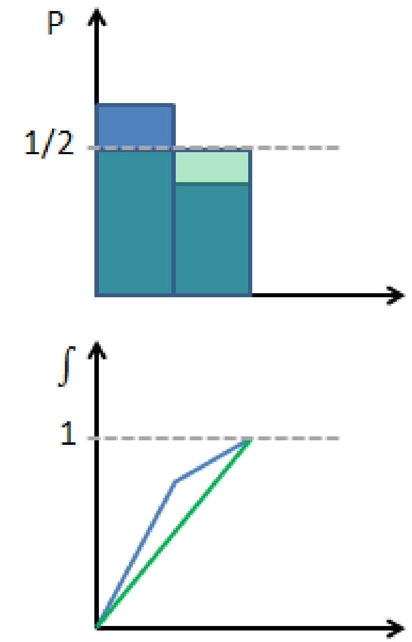
- The only restriction we will make on these is that ρ and σ are diagonal in the energy basis.

- One is allowed two things:
  1. Couple the system to a heat bath. This moves the probabilities towards the thermal state. No work cost/gain.
  2. Shift energy levels. There is work gain/cost if the level is occupied.

- For a given strategy $\mathcal{S}$ and ε probability of failure, what is $W_{\mathcal{S}}^{\varepsilon}$?

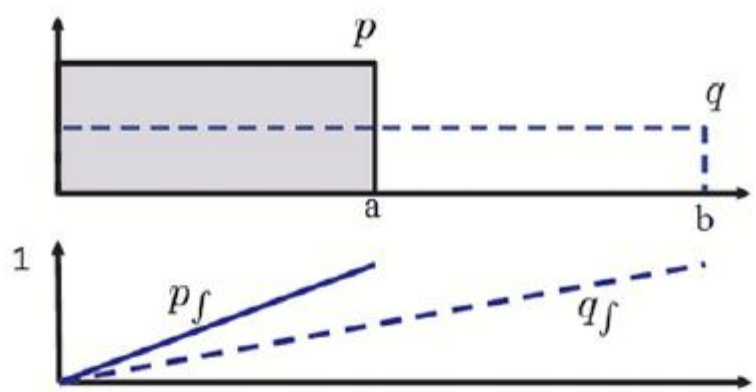
Def. 1. Relative Mixedness

- To give the answer we need two definitions. Firstly we define a measure of how much more random a distribution is than another.



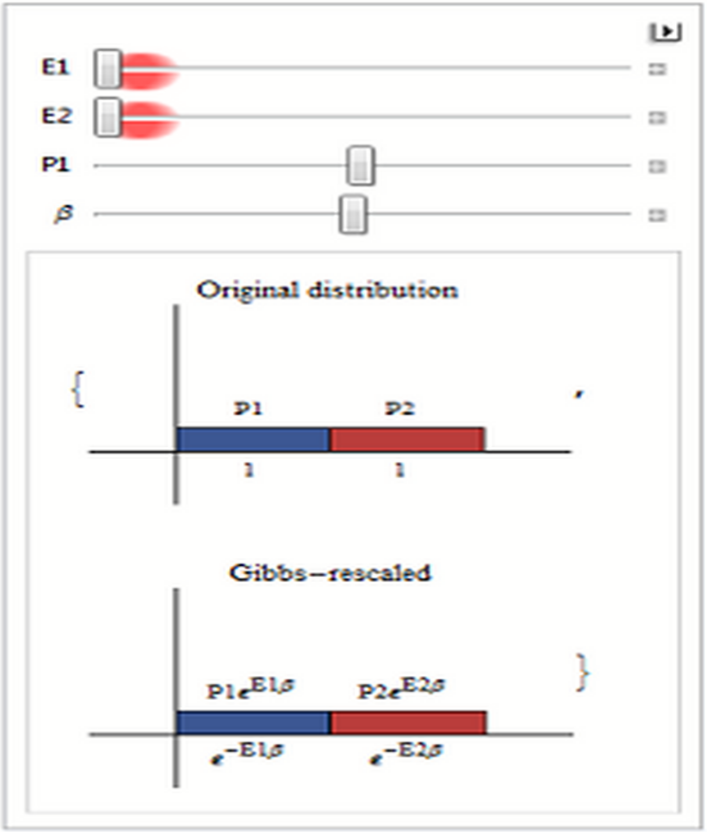
- It is closely related to the idea of majorisation. This is a very important alternative notion of randomness to entropy. A distribution majorises another if its integral (when both distributions are in descending order) always upper bounds the other.

- The relative mixedness of two states, $M(\rho||\sigma)$ we define as the maximum stretching factor by which one can stretch the spectrum of one such that it still upper bounds the other. In simple case to the right M=b/a.



Definition 2: Gibb's rescaling

Definition 1 (Gibbs-rescaling). *An operation on the eigenvalue spectrum* $\{\lambda_i\}$. *Firstly we transform* $\{\lambda_i\}$ *into the associated step-function. Then we take each block, rescale its height as* $\lambda_i \mapsto \frac{\lambda_i}{e^{-\frac{E_i}{kT}}}$, *and its width* $l = 1 \mapsto e^{-\frac{E_i}{kT}}$. *We write this operation applied to a density matrix* ρ *as* $G^T(\rho)$.



This trick is adapted from [Ruch, Meade et. al. (1970's)] Also used by Horodecki and Oppenheim.

Most general statement to date

Having defined the relative mixedness $M(.||.)$ and Gibbs-rescaling $G^T(.)$ we can now give the main result of Egloff et. al..

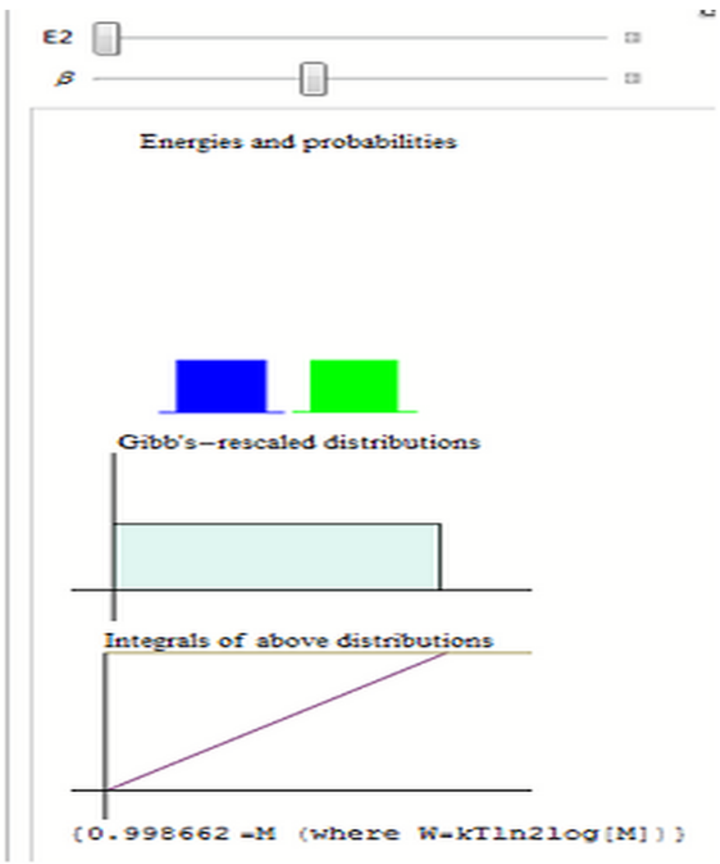
Theorem 1. *In the work extraction game defined above, consider an initial density matrix* $\rho = \sum_i \lambda_i |e_i\rangle\langle e_i|$ *and final density matrix* $\sigma = \sum_j \nu_j |f_j\rangle\langle f_j|$ *with* $\{|e_i\rangle\}$, $\{|f_j\rangle\}$ *the respective energy eigenstates and both* ρ *and* σ *having finite rank. Let the work one can extract except with a probability* ε *of failure using strategy* $\mathcal{S}$ *be denoted* $W^{\varepsilon}_{\mathcal{S}}(\rho \to \sigma)$. *For any strategy this respects* $W^{\varepsilon}_{\mathcal{S}}(\rho \to \sigma) \leq W^{\varepsilon}(\rho \to \sigma)$, *where*

$$W^{\varepsilon}(\rho \to \sigma) = kT \ln \left(M \left(\frac{G^T(\rho)}{1-\varepsilon} || G^T(\sigma) \right) \right).$$

(One may, under a certain proviso, construct a strategy that saturates the bound)

The expression reduces to standard expressions within single-shot information theory in certain limits, and to the standard free energy difference in the von Neumann limit.

Application to 2-level case



Quantitative Second law needs tightening

- A standard expression is that

$$\Delta\left(S_{vN} - \beta U\right) \geq 0, \tag{1}$$

where U is expected internal energy of a system interacting with a heat-bath with inverse temperature β.

- In our model this is necessary but not *sufficient* to guarantee that an evolution is possible. Instead a state change $\rho \rightarrow \rho'$ due to a thermalisation with a heat-bath at temperature T is possible if and only if

$$W^0(\rho \rightarrow \rho') \geq 0. \tag{2}$$

- There are processes that respect Eq.1 but violate Eq.2. We show such evolutions could be used to violate Kelvin's second law deterministically!

[Inspired also by Ruch, Meade et. al. (1970's)]

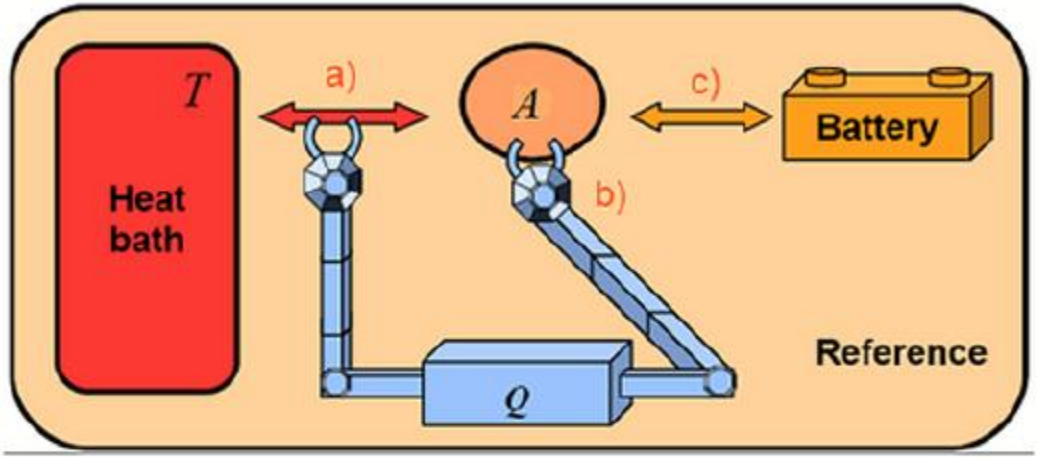
Negative entropy

- For Shannon entropy H, conditional entropy defined by

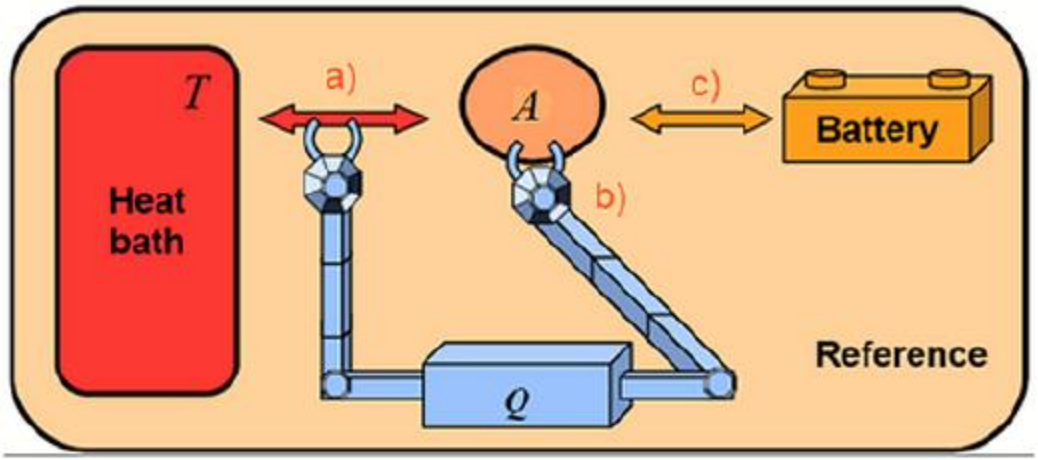
$$S(A|Q)=S(AQ)-H(Q).$$

- Example: $|00\rangle_{AQ}+|11\rangle_{AQ} \rightarrow S(AQ)=0, S(Q)=1 \rightarrow \mathbf{S(A|Q)=-1}$. (!)
- Want to interpret such negative entropy in the Szilard/Landauer setting.

- **We include the observer's memory Q** explicitly in the description of the erasure of system A.



Protocol erasing S with cost *S(A|Q)kTln2*:example



A simple example illustrates the general protocol:
$|\psi\rangle_{SQ} = |00\rangle + |11\rangle$, $S(A|Q) = -1$.

(i) Extract $Wout = 2kT\ln 2$ work from **both** A and Q.

(ii) Reset A to $|0\rangle$ by using $Win = kT\ln 2$ work.

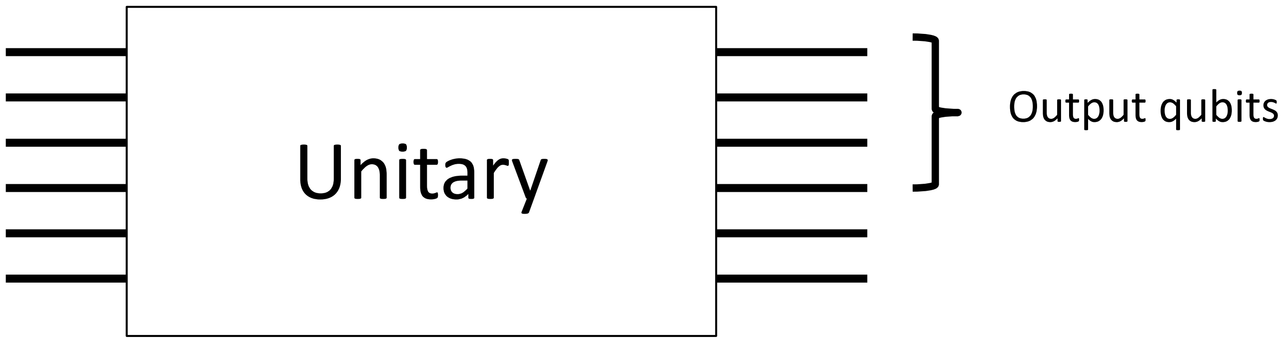
Net result: A reset, reduced state on Q unchanged:

$$W = Win - Wout = -kTln2 = S(A|Q)kTln2.$$

Application for cooling computers

- Extract work from correlations between the output qubits and the rest. Reduced state on output invariant.

- Consider circuit model computation, e.g. Shor's algorithm.



- Not all qubits are measured in the end to get the output.

- The energy extracted comes from the computer and its surroundings, so the computer is cooled.

Outlook

- These are still first steps on this approach.
- We need to engage with experimentalists (this has begun to some extent) to test the statements and to give us guidance about which directions of further development are more important. Need *analogue of steam engine* to guide this.
- A natural way forwards is also to study links with other approaches, such as Jarzynski's equality and fluctuation theorems as well as the typical entanglement approach.
- Thermodynamics is widely used in many areas such as chemistry and biology, it remains to apply these ideas there.

Thank you

Acknowledgements: *Discussions with J. Aaberg, J. Baez, B. Fong, J. Vicary and P. Perinotti. Funding from the Swiss National Science Foundation (grant No. 200020-135048), the European Research Council (grant No. 258932) and the National Research Foundation (Singapore).*

ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

UNIVERSITY OF OXFORD

NUS
National University of Singapore