Information and thermodynamics: experimental verification of Landauer's Erasure principle

Antoine Bérut, Artyom Petrosyan, Sergio Ciliberto

Information and thermodynamics: Experimental verification of Landauer's erasure principle

Antoine Bérut, Artyom Petrosyan and Sergio Ciliberto

Université de Lyon, Ecole Normale Supérieure de Lyon,
Laboratoire de Physique , C.N.R.S. UMR5672,
46, Allée d'Italie, 69364 Lyon, France

March 22, 2015

Abstract

We present an experiment in which a one-bit memory is constructed, using a system of a single colloidal particle trapped in a modulated double-well potential. We measure the amount of heat dissipated to erase a bit and we establish that in the limit of long erasure cycles the mean dissipated heat saturates at the Landauer bound, i.e. the minimal quantity of heat necessarily produced to delete a classical bit of information. This result demonstrates the intimate link between information theory and thermodynamics. To stress this connection we also show that a detailed Jarzynski equality is verified, retrieving the Landauer's bound independently of the work done on the system. The experimental details are presented and the experimental errors carefully discussed

Contents

1 A link between information theory and thermodynamics

The Landauer's principle was first introduced by Rolf Landauer in 1961 [10]. It states that any logically irreversible transformation of classical information is necessarily accompanied by the dissipation of at least $k_{\mathrm{B}}T \ln 2$ of heat per lost bit, where k_{B} is the Boltzmann constant and T is the temperature. This quantity represents only $\sim 3 \times 10^{-21}\,\mathrm{J}$ at room temperature ($300\,\mathrm{K}$) but is a general lower bound, independent of the specific kind of memory system used.

An operation is said to be logically irreversible if its input cannot be uniquely determined from its output. Any Boolean function that maps several input states onto the same output state, such as AND, NAND, OR and XOR, is therefore logically irreversible. In particular, the erasure of information, the RESET TO ZERO operation, is logically irreversible and leads to an entropy increase of at least $k_{\mathrm{B}} \ln 2$ per erased bit.

A simple example can be done with a 1-bit memory system (*i.e.* a systems with two states, called 0 and 1) modelled by a physical double well potential in contact with a single heat bath. In the initial state, the bistable potential is considered to be at equilibrium with the heat bath, and each state (0 and 1) have same probability to occur. Thus, the entropy of the system is $S = k_{\mathrm{B}} \ln 2$, because there are two states with probability $1/2$. If a RESET TO ZERO operation is applied, the system is forced into state 0. Hence, there is only one accessible state with probability 1, and the entropy vanishes

$S = 0$. Since the Second Law of Thermodynamics states that the entropy of a closed system cannot decrease on average, the entropy of the heat bath must increase of at least $k_{\mathrm{B}} \ln 2$ to compensate the memory system's loss of entropy. This increase of entropy can only be done by an heating effect: the system must release in the heat bath at least $k_{\mathrm{B}} T \ln 2$ of heat per bit erased[1].

For a reset operation with efficiency smaller than 1 (*i.e.* if the operation only erase the information with a probability $p < 1$), the Landauer's bound is generalised:

$$\langle Q \rangle \geq k_{\mathrm{B}} T \left[\ln 2 + p \ln(p) + (1-p) \ln(1-p) \right] \tag{1}$$

The Landauer's principle was widely discussed as it could solve the paradox of Maxwell's "demon" [2, 4, 14]. The demon is an intelligent creature able to monitor individual molecules of a gas contained in two neighbouring chambers initially at the same temperature. Some of the molecules will be going faster than average and some will be going slower. By opening and closing a molecular-sized trap door in the partitioning wall, the demon collects the faster (hot) molecules in one of the chambers and the slower (cold) ones in the other. The temperature difference thus created can be used to run a heat engine, and produce useful work. By converting information (about the position and velocity of each particle) into energy, the demon is therefore able to decrease the entropy of the system without performing any work himself, in apparent violation of the Second Law of Thermodynamics. A simpler version with a single particle, called Szilard Engine [20] has recently been realised experimentally [21], showing that information can indeed be used to extract work from a single heat bath. The paradox can be resolved by noting that during a full thermodynamic cycle, the memory of the demon, which is used to record the coordinates of each molecule, has to be reset to its initial state. Thus, the energy cost to manipulate the demon's memory compensate the energy gain done by sorting the gas molecules, and the Second Law of Thermodynamics is not violated any more.

More information can be found in the two books [11, 12], and in the very recent review [13] about thermodynamics of information based on stochastic thermodynamics and fluctuation theorems.

In this article, we describe an experimental realization of the Landauer's information erasure procedure, using a Brownian particle trapped with optical tweezers in a time-dependent double well potential. This kind of system was theoretically [19] and numerically [6] proved to show the Landauer's bound

[1] It it sometimes stated that the cost is $k_{\mathrm{B}} T \ln 2$ per bit written. It is actually the same operation as the RESET TO ZERO can also be seen to store one given state (here state 0), starting with an unknown state.

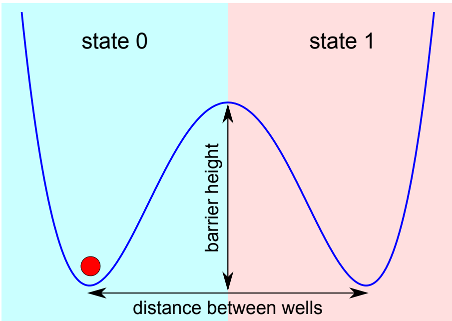


Figure 1: Schematic representation of the one-bit memory system, made of one particle trapped in a double well potential.

$k_B T \ln 2$ for the mean dissipated heat when an information erasure operation is applied. The results described in this article were partially presented in two previous articles [3, 5], and were later confirmed by two independent experimental works [8, 15].

The article is organized as follows. In section 2 we describe the experimental set-up and the one bit memory. In section 3 the experimental results on the Landauer's bound and the dissipated heat are presented. In section 4 we analyze the experimental results within the context of the Jarzinsky equality. Finally we conclude in section 5.

2 Experimental set-up

2.1 The one-bit memory system

The one-bit memory system, is made of a double well potential where one particle is trapped by optical tweezers. If the particle is in the left-well the system is in the state "0", if the particle is in the right-well the system in the state "1" (see fig. 1).

The particles are silica beads (radius $R = 1.00 \pm 0.05\,\mu\text{m}$), diluted a low concentration in bidistilled water. The solution is contained in a disk-shape cell. The center of the cell has a smaller depth ($\sim 80\,\mu\text{m}$) compared to the rest of the cell ($\sim 1\,\text{mm}$), see figure 2. This central area contains less particles than the rest of the cell and provides us a clean region where one particle can be trapped for a long time without interacting with other particles.

The double well potential is created using an Acousto-Optic Deflector which allows us to switch very rapidly (at a rate of $10\,\text{kHz}$) a laser beam (wavelength $\lambda = 1024\,\text{nm}$) between two positions (separated by a fixed dis-

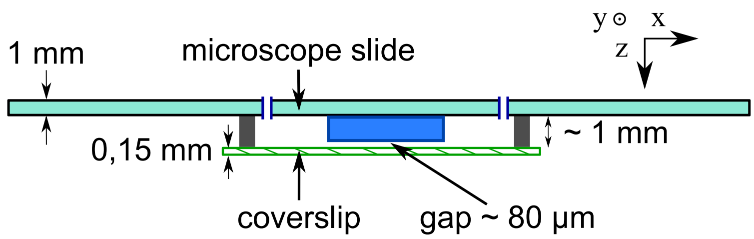


Figure 2: Schematic representation of the cell used to trap particles dispersed in water (view from the side). The central part has a smaller gap than the rest of the cell.

tance $d \sim 1\,\mu\mathrm{m}$). These two positions become for the particle the two wells of the double well potential. The intensity of the laser I can be controlled from 10 mW to more than 100 mW, which enables us to change the height of the double well potential's central barrier [2]. A NanoMax closed-loop piezo-electric stage from Thorlabs® with high resolution (5 nm) can move the cell with regard to the position of the laser. Thus it allows us to create a fluid flow around the trapped particle. The position of the bead is tracked using a fast camera with a resolution of 108 nm per pixel, which after treatment gives the position with a precision greater than 5 nm. The trajectories of the bead are sampled at 502 Hz. See figure 3.

The beads are trapped at a distance $h = 25\,\mu\mathrm{m}$ from the bottom of the cell. The double well potential must be tuned for each particle, in order to be as symmetrical as possible and to have the desired central barrier. The tuning is done by adjusting the distance between the two traps and the time that the laser spend on each trap. The asymmetry can be reduced to $\sim 0.1\,k_{\mathrm{B}}T$. The double well potential $U_0(x, I)$ (with x the position and I the intensity of the laser) can simply be measured by computing the equilibrium distribution of the position for one particle in the potential:

$$P(x, I) \propto \exp(-\frac{U(x, I)}{k_{\mathrm{B}}T}). \tag{2}$$

One typical double well potential is shown in figure 4.

2.2 The information erasure procedure

We perform the erasure procedure as a logically irreversible operation. This procedure brings the system initially in one unknown state (0 or 1 with same

[2]The values are the power measured on the beam before the microscope objective, so the "real" power at the focal point should be smaller, due to the loss in the objective.

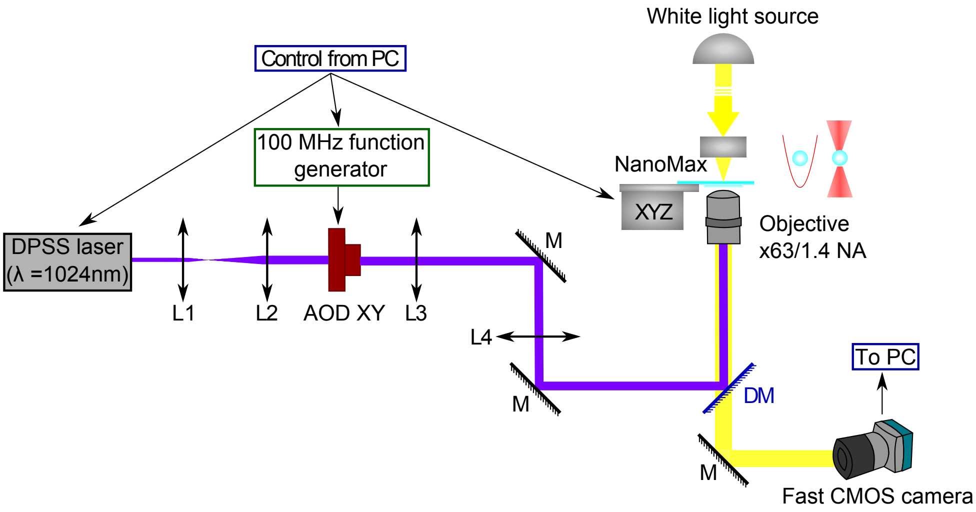


Figure 3: Schematic representation of optical tweezers set-up used to trap one particle in a double well potential. The Acousto-Optic Deflector (AOD) is used to switch rapidly the trap between two positions. The NanoMax piezo stage can move the cell with regard to the laser, which creates a flow around the trapped particle. "M" are mirrors and "DM" is a dichroic mirror.

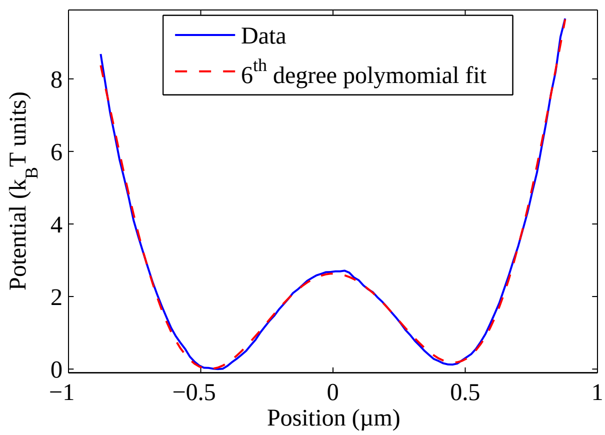


Figure 4: Double well potential measured by computing the equilibrium distribution of one particle's positions, with $I_{\text{laser}} = 15\,\text{mW}$. Here the distribution is computed on $1.5e6$ points sampled at $502\,\text{Hz}(i.e.a50\,\text{minlong measure-}$ment). The double well potential is well fitted by a 6^{th} degree polynomium.

probability) in one chosen state (we choose 0 here). It is done experimentally in the following way (and summarised in figure 5a):

- At the beginning the bead must be trapped in one well-defined state (0 or 1). For this reason, we start with a high laser intensity ($I_{\mathrm{high}} = 48\,\mathrm{mW}$) so that the central barrier is more than $8\,k_{\mathrm{B}}T$. In this situation, the characteristic jumping time (Kramers Time) is about 3000 s, which is long compared to the time of the experiment, and the equivalent stiffness of each well is about 1.5 pN/µm. The system is left 4 s with high laser intensity so that the bead is at equilibrium in the well where it is trapped [3]. The potential $U_0(x, I_{\mathrm{high}})$ is represented in figure 5b **1**.

- The laser intensity is first lowered (in a time $T_{\mathrm{low}} = 1\,\mathrm{s}$) to a low value ($I_{\mathrm{low}} = 15\,\mathrm{mW}$) so that the barrier is about $2.2\,k_{\mathrm{B}}T$. In this situation the jumping time falls to $\sim 10\,\mathrm{s}$, and the equivalent stiffness of each well is about 0.3 pN/µm. The potential $U_0(x, I_{\mathrm{low}})$ is represented figure 5b **2**.

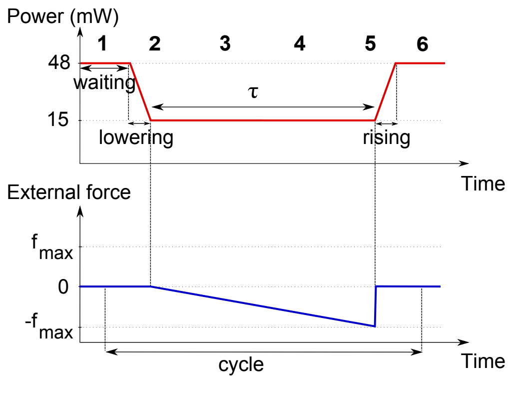
- A viscous drag force linear in time is induced by displacing the cell with respect to the laser using the piezoelectric stage. The force is given by $f = \gamma v$ where $\gamma = 6\pi R\eta$ (η is the viscosity of water) and v the speed of displacement. It tilts the double well potential so that the bead ends always in the same well (state 0 here) independently of where it started. See figures 5b **3** to **5**.

- At the end, the force is stopped and the central barrier is raised again to its maximal value (in a time $T_{\mathrm{high}} = 1\,\mathrm{s}$). See figure 5b **6**.

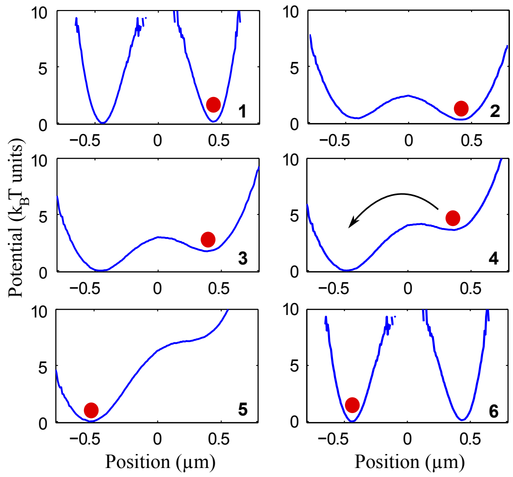
The total duration of the erasure procedure is $T_{\mathrm{low}}+\tau+T_{\mathrm{high}}$. Since we kept $T_{\mathrm{low}} = T_{\mathrm{high}} = 1\,\mathrm{s}$, a procedure is fully characterised by the duration τ and the maximum value of the force applied f_{max}. Its efficiency is characterized by the "proportion of success" P_S, which is the proportion of trajectories where the bead ends in the chosen well (state 0), independently of where it started.

Note that for the theoretical procedure, the system must be prepared in an equilibrium state with same probability to be in state 1 than in state 0. However, it is more convenient experimentally to have a procedure always starting in the same position. Therefore we separate the procedure in two sub-procedures: one where the bead starts in state 1 and is erased in state 0, and one where the bead starts in state 0 and is erased in state 0. The fact that

[3] The characteristic time for the particle trapped in one well when the barrier is high is 0.08 s



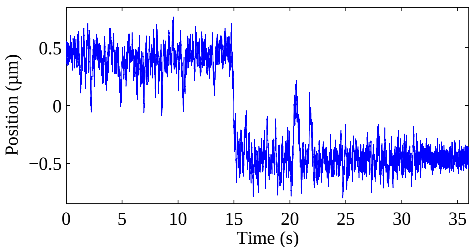
(a) Laser intensity and external drag force.



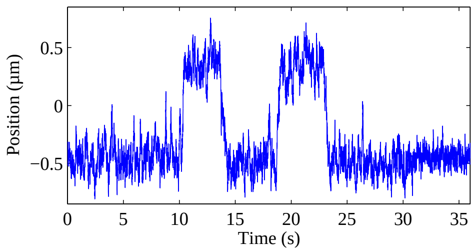
(b) Potential felt by the particle.

Figure 5: Schematical representation of the erasure procedure. The potential felt by the trapped particle is represented at different stages of the procedure (**1** to **6**). For **1** and **2** the potential $U_0(x, I)$ is measured. For **3** to **5** the potential is constructed from $U_0(x, I_{\text{low}})$ knowing the value of the applied drag force.

the position of the bead at the beginning of each procedure is actually known is not a problem because this knowledge is not used by the erasure procedure. The important points are that there is as many procedures starting in state 0 than in state 1, and that the procedure is always the same regardless of the initial position of the bead. Examples of trajectories for the two sub-procedures $1 \to 0$ and $0 \to 0$ are shown in figure 6.



(a) $1 \to 0$



(b) $0 \to 0$

Figure 6: Examples of trajectories for the erasure procedure. $t = 0$ corresponds to the time where the barrier starts to be lowered. The two possibilities of initial state are shown.

3 Landauer's bound for dissipated heat

3.1 Computing the dissipated heat

The system can be described by an over-damped Langevin equation:

$$\gamma \dot{x} = -\frac{\partial U_0}{\partial x}(x, I) + f(t) + \xi(t) \tag{3}$$

with x the position of the particle[4], $\dot{x} = \frac{\mathrm{d}x}{\mathrm{d}t}$ its velocity, $\gamma = 6\pi R\eta$ the friction coefficient (η is the viscosity of water), $U_0(x, I)$ the double well potential created by the optical tweezers, $f(t)$ the external drag force exerted by displacing the cell, and $\xi(t)$ the thermal noise which verifies $\langle \xi(t) \rangle = 0$ and $\langle \xi(t)\xi(t') \rangle = 2\gamma k_\mathrm{B} T \delta(t - t')$, where $\langle . \rangle$ stands for the ensemble average.

Following the formalism of the stochastic energetics [17], the heat dissipated by the system into the heat bath along the trajectory $x(t)$ between time $t = 0$ and t is:

$$Q_{0,t} = \int_0^t -\left(\xi(t') - \gamma \dot{x}(t')\right) \dot{x}(t')\, \mathrm{d}t'. \tag{4}$$

Using equation 3, we get:

$$Q_{0,t} = \int_0^t \left(-\frac{\partial U_0}{\partial x}(x, I) + f(t') \right) \dot{x}(t')\, \mathrm{d}t'. \tag{5}$$

For the erasure procedure described in 2.2 the dissipated heat can be decomposed in three terms:

$$Q_\mathrm{erasure} = Q_\mathrm{barrier} + Q_\mathrm{potential} + Q_\mathrm{drag} \tag{6}$$

Where:

- Q_barrier is the heat dissipated when the central barrier is lowered and risen ($f = 0$ during these stages of the procedure):

$$Q_\mathrm{barrier} = \int_0^{T_\mathrm{low}} \left(-\frac{\partial U_0}{\partial x}(x, I) \right) \dot{x}\, \mathrm{d}t' + \int_{T_\mathrm{low}+\tau}^{T_\mathrm{low}+\tau+T_\mathrm{high}} \left(-\frac{\partial U_0}{\partial x}(x, I) \right) \dot{x}\, \mathrm{d}t' \tag{7}$$

- $Q_\mathrm{potential}$ is the heat dissipated due to the force of the potential, during the time τ where the external drag force is applied (the laser intensity is constant during this stage of the procedure):

$$Q_\mathrm{potential} = \int_{T_\mathrm{low}}^{T_\mathrm{low}+\tau} \left(-\frac{\partial U_0}{\partial x}(x, I) \right) \dot{x}\, \mathrm{d}t' \tag{8}$$

[4] We take the reference $x = 0$ as the middle position between the two traps.

- Q_{drag} is the heat dissipated due the external drag force applied during the time τ (the laser intensity is constant during this stage of the procedure):

$$Q_{\text{drag}} = \int_{T_{\text{low}}}^{T_{\text{low}}+\tau} f\dot{x}\,\mathrm{d}t' \tag{9}$$

The lowering and rising of the barrier are done in a time much longer than the relaxation time of the particle in the trap ($\sim 0.1\,\mathrm{s}$), so they can be considered as a quasi-static cyclic process, and do not contribute to the dissipated heat in average. The complete calculation can also be done if we assume that the particle do not jump out of the well where it is during the change of barrier height. In this case, we can do a quadratic approximation: $U_0(x, I) = -k(I)x$ where k is the stiffness of the trap, which evolves in time because it depends linearly on the intensity of the laser. Then:

$$\langle Q_{\text{lowering}} \rangle = \left\langle \int_0^{T_{\text{low}}} -kx\dot{x}\,\mathrm{d}t' \right\rangle = \int_0^{T_{\text{low}}} -\frac{k}{2}\,\mathrm{d}\left(\langle x^2 \rangle\right). \tag{10}$$

Using the equipartition theorem (which is possible because the change of stiffness is assumed to be quasi-static), we get:

$$\langle Q_{\text{lowering}} \rangle = \int_0^{T_{\text{low}}} -k\frac{k_{\mathrm{B}}T}{2}\,\mathrm{d}\left(\frac{1}{k}\right) = \frac{k_B T}{2} \ln\left(\frac{k_{\text{low}}}{k_{\text{high}}}\right). \tag{11}$$

The same calculation gives $\langle Q_{\text{rising}} \rangle = \frac{k_B T}{2} \ln\left(\frac{k_{\text{high}}}{k_{\text{low}}}\right)$, and it follows directly that $\langle Q_{\text{barrier}} \rangle = \langle Q_{\text{lowering}} \rangle + \langle Q_{\text{rising}} \rangle = 0$.

The heat dissipated due to the potential when the force is applied is also zero in average. Indeed, the intensity is constant and U_0 becomes a function depending only on x. It follows that:

$$\langle Q_{\text{potential}} \rangle = \int_{T_{\text{low}}}^{T_{\text{low}}+\tau} -\frac{\mathrm{d}U_0}{\mathrm{d}x}\,\mathrm{d}x = \Big[U_0(x)\Big]_{x(T_{\text{low}}+\tau)}^{x(T_{\text{low}})}. \tag{12}$$

Since the potential is symmetrical, there is no change in U_0 when the bead goes from one state to another, and $\langle U_0(x(T_{\text{low}})) - U_0(x(T_{\text{low}}+\tau))\rangle = 0$.

Finally, since we are interested in the mean dissipated heat, the only relevant term to calculate is the heat dissipated by the external drag force:

$$Q_{\text{drag}} = \int_{T_{\text{low}}}^{T_{\text{low}}+\tau} \gamma v(t')\dot{x}\,\mathrm{d}t'. \tag{13}$$

Where γ is the known friction coefficient, $v(t)$ is the imposed displacement of the cell (which is not a fluctuating quantity) and $\dot{x}$ can be estimated simply:

$$\dot{x}(t + \delta t/2) = \frac{x(t+\delta t) - x(t)}{\delta t}. \tag{14}$$

We measured Q_{drag} for several erasure procedures with different parameters τ and f_{max}. For each set of parameters, we repeated the procedure a few hundred times in order to compute the average dissipated heat.

We didn't measure Q_{lowering} and Q_{rising} because it requires to know the exact shape of the potential at any time during lowering and rising of the central barrier. The potential could have been measured by computing the equilibrium distribution of one particle's positions for different values of I. But these measurements would have been very long since they require to be done on times much longer than the Kramers time to give a good estimation of the double well potential. Nevertheless, we estimated on numerical simulations with parameters close to our experimental ones that $\langle Q_{\text{lowering}} + Q_{\text{potential}} + Q_{\text{rising}} \rangle \approx 0.7\, k_{\text{B}}T$ which is only 10 % of the Landauer's bound.

3.2 Results

We first measured P_S the proportion of success for different set of τ and f_{max}. 'Qualitatively, the bead is more likely to jump from one state to another thanks to thermal fluctuations if the waiting time is longer. Of course it also has fewer chances to escape from state 0 if the force pushing it toward this state is stronger. We did some measurements keeping the product $\tau \times f_{\text{max}}$ constant. The results are shown in figure 7 (blue points).

The proportion of success is clearly not constant when the product $\tau \times f_{\text{max}}$ is kept constant, but, as expected, the higher the force, the higher P_S. One must also note that the experimental procedure never reaches a P_S higher than $\sim 95\,\%$. This effect is due to the last part of the procedure: since the force is stopped when the barrier is low, the bead can always escape from state 0 during the time needed to rise the barrier. This problem can be overcome with a higher barrier or a faster rising time T_{high}. It was tested numerically by Raoul Dillenscheider and Éric Lutz, using a protocol adapted from [6] to be close to our experimental procedure. They showed that for a high barrier of $8\, k_{\text{B}}T$ the proportion of success approaches only $\sim 94\,\%$, whereas for a barrier of $15\, k_{\text{B}}T$ it reaches $\sim 99\,\%$. Experimentally we define a proportion of success $P_{S\,\text{force}}$ by counting the number of procedures where the particle ends in state 0 when the force is stopped (*before* the rising of the barrier). It quantifies the efficiency of the pushing force, which is the relevant one since we have shown that the pushing force is the only contribution to the mean dissipated heat. Measured $P_{S\,\text{force}}$ are shown in figure 7 (red points). $P_{S\,\text{force}}$ is roughly always 5 % bigger than P_S and it reaches 100 % of success for high forces.

To reach the Landauer's bound, the force necessary to erase information

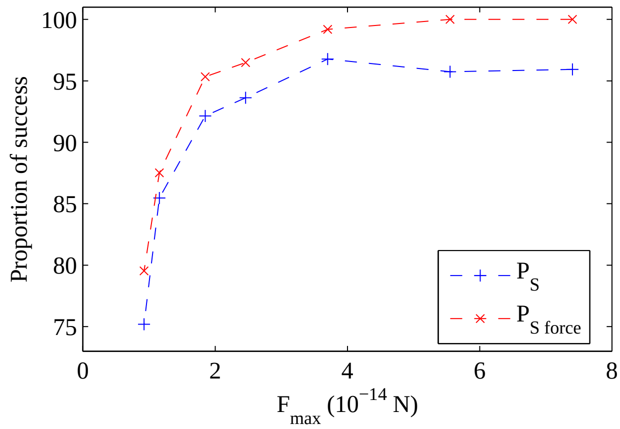


Figure 7: Proportion of success for different values of τ and $f_{\max}$, keeping constant the product $\tau \times f_{\max} \approx 0.4\,\mathrm{pN\,s}$ (which corresponds to $\tau \times v_{max} = 20\,\mu\mathrm{m}$). P_S quantifies the ratio of procedures which ends in state 0 after the barrier is risen. $P_{S\,\mathrm{force}}$ quantifies the ratio of procedures which ends in state 0 before the barrier is risen.

must be as low as possible, because it is clear that a higher force will always produce more heat for the same proportion of success. Moreover, the bound is only reachable for a quasi-static (*i.e.* $\tau \to \infty$) erasure procedure, and the irreversible heat dissipation associated with a finite time procedure should decrease as $1/\tau$ [18]. Thus we decided to work with a chosen τ and to manually[5] optimise the applied force. The idea was to choose the lowest value of $f_{\max}$ which gives a $P_{S\,\mathrm{force}} \geq 95\,\%$.

The Landauer's bound $k_{\mathrm{B}}T \ln 2$ is only valid for totally efficient procedures. Thus one should theoretically look for a Landauer's bound corresponding to each experimental proportion of success (see equation 1). Unfortunately the function $\ln 2 + p\ln(p) + (1-p)\ln(1-p)$ quickly decreases when p is lower than 1. To avoid this problem, we made an approximation by computing $\langle Q \rangle_{\to 0}$ the mean dissipated heat for the trajectories where the memory is erased (*i.e* the ones ending in state 0). We consider that $\langle Q \rangle_{\to 0}$ mimics the mean dissipated heat for a procedure with $100\,\%$ of success. This approximation is reasonable as long as $P_{S\,\mathrm{force}}$ is close enough to $100\,\%$, because the negative contributions which reduce the average dissipated heat are mostly due to the rare trajectories going against the force (*i.e* ending in state 1). Of course,

[5]The term "manually" refers to the fact that the optimisation was only empirical and that we did not computed the theoretical best $f_{\max}$ for a given value of τ.

at the limit where the force is equal to zero, one should find $P_{S\,\mathrm{force}} = 50\,\%$ and $\langle Q \rangle_{\to 0} = 0$ which is different from $k_{\mathrm{B}}T \ln 2$. The mean dissipated heat for several procedures are shown in figure 8.

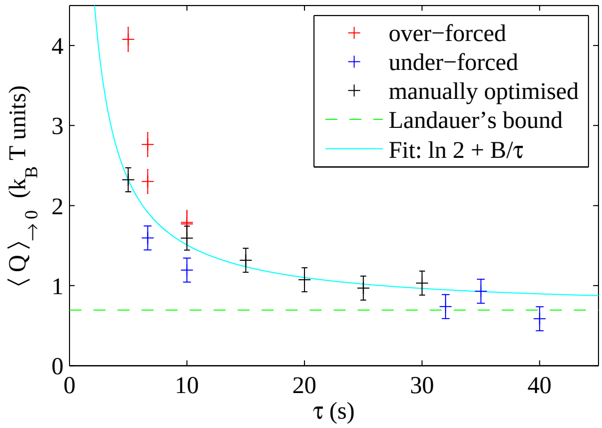


Figure 8: Mean dissipated heat for several procedures, with fixed τ and different values of f_{max}. The red points have a force too high, and a $P_{S\,\mathrm{force}} \geq 99\,\%$. The blue points have a force too low and $91\,\% \leq P_{S\,\mathrm{force}} < 95\,\%$ (except the last point which has $P_{S\,\mathrm{force}} \approx 80\,\%$). The black points are considered to be optimised and have $95\,\% \leq P_{S\,\mathrm{force}} < 99\,\%$. The errors bars are $\pm 0.15\,k_{\mathrm{B}}T$ estimated from the reproductibility of measurement with same parameters. The fit $\langle Q \rangle_{\to 0} = \ln 2 + B/\tau$ is done only by considering the optimised procedures.

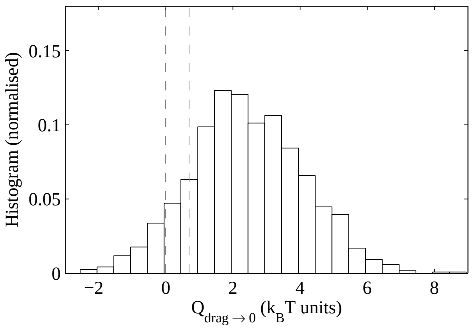
The mean dissipated heat decreases with the duration of the erasure procedure τ and approaches the Landauer's bound $k_{\mathrm{B}}T \ln 2$ for long times. Of course, if we compute the average on all trajectories (and not only on the ones ending in state 0) the values of the mean dissipated heat are smaller, but remain greater than the generalised Landauer's bound for the corresponding proportion of success p:

$$\langle Q \rangle_{\to 0} \geq \langle Q \rangle \geq k_{\mathrm{B}}T \left[\ln 2 + p \ln(p) + (1-p)\ln(1-p)\right] \tag{15}$$

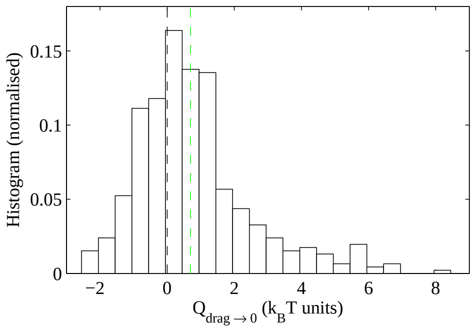
For example, the last point ($\tau = 40\,\mathrm{s}$) has a proportion of success $P_{S\,\mathrm{force}} \approx 80\,\%$, which corresponds to a Landauer's bound of only $\approx 0.19\,k_{\mathrm{B}}T$, and we measure $\langle Q \rangle_{\to 0} = 0.59\,k_{\mathrm{B}}T$ greater than $\langle Q \rangle = 0.26\,k_{\mathrm{B}}T$. The manually optimised procedures also seem to verify a decreasing of $\langle Q \rangle_{\to 0}$ proportional to $1/\tau$. A numerical least square fit $\langle Q \rangle_{\to 0} = \ln 2 + B/\tau$ is plotted in figure 8 and gives a value of $B = 8.15\,k_{\mathrm{B}}T\,\mathrm{s}$. If we do a fit with two free parameters

$\langle Q \rangle_{\to 0} = A + B/\tau$, we find $A = 0.72\, k_{\mathrm{B}}T$ which is close to $k_{\mathrm{B}}T \ln 2 \approx 0.693\, k_{\mathrm{B}}T$.

One can also look at the distribution of $Q_{\mathrm{drag}\to 0}$. Histograms for procedures going from 1 to 0 and from 0 to 0 are shown in figure 9. The statistics are not sufficient to conclude on the exact shape of the distribution, but as expected, there is more heat dissipated when the particle has to jump from state 1 to state 0 than when it stays in state 0. It is also noticeable that a fraction of the trajectories always dissipate less heat than the Landauer's bound, and that some of them even have a negative dissipated heat. We are able to approach the Landauer's bound in average thanks to those trajectories where the thermal fluctuations help us to erase the information without dissipating heat.



(a) $1 \to 0$



(b) $0 \to 0$

Figure 9: Histograms of the dissipated heat $Q_{\mathrm{drag}\to 0}$. (a) For one procedure going from 1 to 0 ($\tau = 10\,\mathrm{s}$ and $f_{\max} = 3.8 \times 10^{-14}\,\mathrm{N}$). (b) For one procedure going from 0 to 0 ($\tau = 5\,\mathrm{s}$ and $f_{\max} = 3.8 \times 10^{-14}\,\mathrm{N}$). The black vertical lines indicate $Q = 0$ and the green ones indicate the Landauer's bound $k_{\mathrm{B}}T \ln 2$.

4 Integrated Fluctuation Theorem applied on information erasure procedure

We have shown that the mean dissipated heat for an information erasure procedure applied on a 1 bit memory system approaches the Landauer's bound for quasi-static transformations. One may wonder if we can have direct access to the variation of free-energy between the initial and the final state of the system, which is directly linked to the variation of the system's entropy.

4.1 Computing the stochastic work

To answer to this question it seems natural to use the Integrated Fluctuation Theorem called the Jarzinsky equality [7] which allows one to compute the free energy difference between two states of a system, in contact with a heat bath at temperature T. When such a system is driven from an equilibrium state A to a state B through any continuous procedure, the Jarzynski equality links the stochastic work W_{st} received by the system during the procedure to the free energy difference $\Delta F = F_B - F_A$ between the two states:

$$\left\langle \mathrm{e}^{-\beta W_{\mathrm{st}}} \right\rangle = \mathrm{e}^{-\beta \Delta F} \tag{16}$$

Where $\langle . \rangle$ denotes the ensemble average over all possible trajectories, and $\beta = \frac{1}{k_{\mathrm{B}} T}$.

For a colloidal particle confined in one spatial dimension and submitted to a conservative potential $V(x, \lambda)$, where $\lambda = \lambda(t)$ is a time-dependent external parameter, the stochastic work received by the system is defined by [16]:

$$W_{\mathrm{st}}[x(t)] = \int_0^t \frac{\partial V}{\partial \lambda} \dot{\lambda} \, \mathrm{d}t' \tag{17}$$

Here the potential is made by the double-well and the tilting drag force[6]

$$V(x, \lambda) = U_0(x, I(t)) - f(t) x \tag{18}$$

and we have two control parameters: $I(t)$ the intensity of the laser and $f(t)$ the amplitude of the drag force.

Once again, we can separate two contributions: one coming from the lowering and rising of the barrier, and one coming from the applied external drag

[6] Since we are in one dimension, any external force can be written as the gradient of a global potential.

force. We again consider that the lowering and rising of the barrier should not modify the free-energy of the system, and that the main contribution is due to the drag force. Thus:

$$W_{\mathrm{st}} = \int_{T_{\mathrm{low}}}^{T_{\mathrm{low}}+\tau} -\dot{f}x \,\mathrm{d}t' \tag{19}$$

Noting that $f(t = T_{\mathrm{low}}) = 0 = f(t = T_{\mathrm{low}} + \tau)$, it follows from an integration by parts that the stochastic work is equal to the heat dissipated by the drag force:

$$W_{\mathrm{st}} = \int_{T_{\mathrm{low}}}^{T_{\mathrm{low}}+\tau} -\dot{f}x \,\mathrm{d}t' = \int_{T_{\mathrm{low}}}^{T_{\mathrm{low}}+\tau} f\dot{x} \,\mathrm{d}t' = Q_{\mathrm{drag}} \tag{20}$$

The two integrals have been calculated experimentally for all the trajectories of all the procedures tested and it was verified that the difference between the two quantity is completely negligible. In the following parts, we write W_{st} for theoretic calculations, and Q_{drag} when we apply the calculations to our experimental data.

4.2 Interpreting the free-energy difference

Since the memory erasure procedure is made in a cyclic way (which implies $\Delta U = 0$) and $\Delta S = -k_{\mathrm{B}} \ln 2$ it is natural to await $\Delta F = k_{\mathrm{B}} T \ln 2$. But the ΔF that appears in the Jarsynski equality is the difference between the free energy of the system in the initial state (which is at equilibrium) and the equilibrium state corresponding to the final value of the control parameter:

$$\Delta F_{\mathrm{Jarzynski}} = F(\lambda(t_{\mathrm{final}})) - F(\lambda(t_{\mathrm{initial}})) \tag{21}$$

Because the height of the barrier is always finite there is no change in the equilibrium free energy of the system between the beginning and the end of our procedure. Then $\Delta F_{\mathrm{Jarzynski}} = 0$, and we await $\left\langle \mathrm{e}^{-\beta W_{\mathrm{st}}} \right\rangle = 1$, which is not very interesting.

Nevertheless it has been shown [22] that, when there is a difference between the actual state of the system (described by the phase-space density ρ_t) and the equilibrium state (described by ρ_t^{eq}), the Jarzynski equality can be modified:

$$\left\langle \mathrm{e}^{-\beta W_{\mathrm{st}}(t)} \right\rangle_{(x,t)} = \frac{\rho^{\mathrm{eq}}(x, \lambda(t))}{\rho(x,t)} \mathrm{e}^{-\beta \Delta F_{\mathrm{Jarzynski}}(t)} \tag{22}$$

Where $\langle . \rangle_{(x,t)}$ is the mean on all the trajectories that pass through x at t.

In our procedure, selecting the trajectories where the information is actually erased is equivalent to fix the position x to the chosen final well (state 0 corresponds to $x < 0$) at the time $t = T_{\mathrm{low}} + \tau$. It follows that $\rho(x < 0, T_{\mathrm{low}} + \tau)$ is directly $P_{S\,\mathrm{force}}$, the proportion of success of the procedure, and $\rho^{\mathrm{eq}}(x < 0, \lambda(T_{\mathrm{low}}+\tau)) = 1/2$ since both wells have same probability at equilibrium[7]. Then:

$$\left\langle \mathrm{e}^{-\beta W_{\mathrm{st}}(T_{\mathrm{low}}+\tau)} \right\rangle_{\to 0} = \frac{1/2}{P_{S\,\mathrm{force}}} \tag{23}$$

Similarly for the trajectories that end the procedure in the wrong well (state 1) we have:

$$\left\langle \mathrm{e}^{-\beta W_{\mathrm{st}}(T_{\mathrm{low}}+\tau)} \right\rangle_{\to 1} = \frac{1/2}{1 - P_{S\,\mathrm{force}}} \tag{24}$$

Taking into account the Jensen's inequality, i.e. $\langle \mathrm{e}^{-x} \rangle \geq \mathrm{e}^{-\langle x \rangle}$, we find that equations 23 and 24 imply:

$$\begin{array}{l} \langle W_{\mathrm{st}} \rangle_{\to 0} \geq k_{\mathrm{B}} T \left[\ln(2) + \ln(P_{S\,\mathrm{force}}) \right] \\ \langle W_{\mathrm{st}} \rangle_{\to 1} \geq k_{\mathrm{B}} T \left[\ln(2) + \ln(1 - P_{S\,\mathrm{force}}) \right] \end{array} \tag{25}$$

Since that the mean stochastic work dissipated to realise the procedure is simply:

$$\langle W_{\mathrm{st}} \rangle = P_{S\,\mathrm{force}} \langle W_{\mathrm{st}} \rangle_{\to 0} + (1 - P_{S\,\mathrm{force}}) \langle W_{\mathrm{st}} \rangle_{\to 1} \tag{26}$$

it follows:

$$\langle W_{\mathrm{st}} \rangle \geq k_{\mathrm{B}} T \left[\ln(2) + P_{S\,\mathrm{force}} \ln(P_{S\,\mathrm{force}}) + (1 - P_{S\,\mathrm{force}}) \ln(1 - P_{S\,\mathrm{force}}) \right] \tag{27}$$

which is the generalization of the Landauer's bound for $P_{S\,\mathrm{force}} < 100\,\%$. Hence, the Jarzynski equality applied to the information erasure procedure allows one to find the complete Landauer's bound for the stochastic work received by the system.

Finally we experimentally compute ΔF_{eff} which is the logarithm of the exponential average of the dissipated heat for trajectories ending in state 0:

$$\Delta F_{\mathrm{eff}} = -\ln\left(\left\langle \mathrm{e}^{-\beta Q_{\mathrm{drag}}} \right\rangle_{\to 0} \right). \tag{28}$$

Data are shown in figure 10. The error bars are estimated by computing the average on the data set with 10% of the points randomly excluded, and taking the maximal difference in the values observed by repeating this operation 1000 times. Except for the first points[8] ($\tau = 5\,\mathrm{s}$), the values are very

[7] A more detailed demonstration is given in Appendix.

[8] We believe that the discrepancy can be explained by the fact that the values of $Q_{\mathrm{drag}\,\to 0}$ are bigger and that it is more difficult to estimate correctly the exponential average in this case.

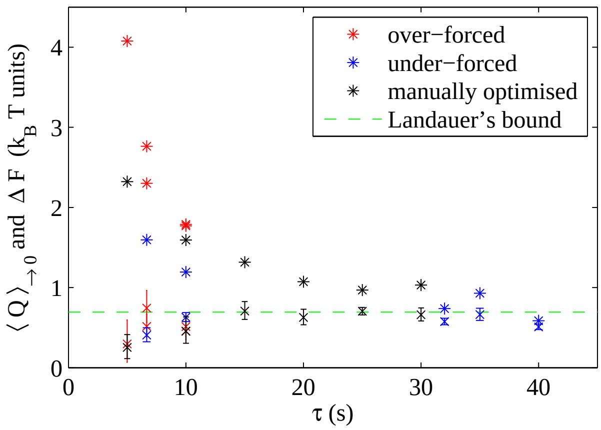


Figure 10: Mean dissipated heat ($*$) and effective free energy difference ($\times$) for several procedures, with fixed τ and different values of $f_{\max}$. The red points have a force too high, and a $P_{S\,\text{force}} \geq 99\,\%$. The blue points have a force too low and $91\,\% \leq P_{S\,\text{force}} < 95\,\%$ (except the last point which has $P_{S\,\text{force}} \approx 80\,\%$). The black points are considered to be optimised and have $95\,\% \leq P_{S\,\text{force}} < 99\,\%$.

close to $k_{\mathrm{B}}T \ln 2$, which is in agreement with equation 23, since $P_{S\,\text{force}}$ is close to $100\,\%$. Hence, we retrieve the Landauer's bound for the free-energy difference, for any duration of the information erasure procedure.

Note that this result is not in contradiction with the classical Jarzynski equality, because if we average over all the trajectories (and not only the ones where the information is erased), we should find:

$$\left\langle \mathrm{e}^{-\beta W_{\mathrm{st}}} \right\rangle = P_{S\,\text{force}} \left\langle \mathrm{e}^{-\beta W_{\mathrm{st}}} \right\rangle_{\to 0} + (1 - P_{S\,\text{force}}) \left\langle \mathrm{e}^{-\beta W_{\mathrm{st}}} \right\rangle_{\to 1} = 1. \tag{29}$$

However, the verification of this equality is hard to do experimentally since we have very few trajectories ending in state 1, which gives us not enough statistics to estimate $\left\langle \mathrm{e}^{-\beta W_{\mathrm{st}}} \right\rangle_{\to 1}$ properly.

4.3 Separating sub-procedures

To go further, we can also look at the two sub-procedures $1 \to 0$ and $0 \to 0$ separately. To simplify calculations, we make here the approximation that $P_{S\,\text{force}} = 100\,\%$.

We can compute the exponential average of each sub-procedure:

$$M_{1\to 0} = \left\langle \mathrm{e}^{-\beta W_{\mathrm{st}}} \right\rangle_{1\to 0} \quad \text{and} \quad M_{0\to 0} = \left\langle \mathrm{e}^{-\beta W_{\mathrm{st}}} \right\rangle_{0\to 0} \tag{30}$$

For each sub-procedure taken independently the classical Jarzynski equality does not hold because the initial conditions are not correctly tested. Indeed selecting trajectories by their initial condition introduces a bias in the initial equilibrium distribution. But it has been shown [9] that for a partition of the phase-space into non-overlapping subsets χ_j $(j = 1, ..., K)$ there is a detailed Jarzynski Equality :

$$\left\langle \mathrm{e}^{-\beta W_{\mathrm{st}}} \right\rangle_j = \frac{\tilde{\rho}_j}{\rho_j} \left\langle \mathrm{e}^{-\beta W_{\mathrm{st}}} \right\rangle \tag{31}$$

with:

$$\rho_j = \int_{\chi_j} \rho(t_a)\, \mathrm{d}x\mathrm{d}p \quad \text{and} \quad \tilde{\rho}_j = \int_{\tilde{\chi}_j} \tilde{\rho}(t_a)\, \mathrm{d}x\mathrm{d}p \tag{32}$$

where $\rho(t_a)$ and $\tilde{\rho}(t_a)$ are the phase-space densities of the system measured at the same intermediate but otherwise arbitrary point in time, in the forward and backward protocol, respectively. The backward protocol is simply the time-reverse of the forward protocol.

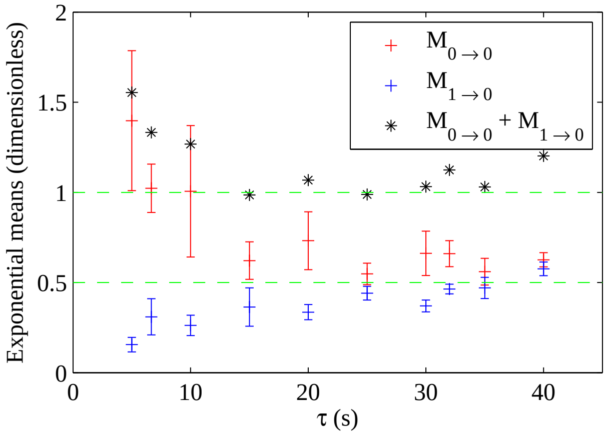


Figure 11: Exponential means computed on the sub-procedures, for several parameters, with fixed τ and manually optimised values of f_{max}. The error-bars are estimated by computing the exponential mean on the data set with 10% of the points randomly excluded, and taking the maximal difference in the values observed by repeating this operation 1000 times.

Here, we take only two subsets $j = \{0 \to 0, 1 \to 0\}$, defined by the position where the bead starts, and we choose $t_a = T_{\mathrm{low}}$ the starting point of the applied force. Then we have:

$$M_{i\to 0} = \frac{\tilde{\rho}_{i\to 0}}{\rho_{i\to 0}} \mathrm{e}^{-\beta \Delta F_{\mathrm{eff}}} = \frac{\tilde{P}_{0\to i}}{1/2}\,\frac{1}{2} \tag{33}$$

where $i = \{0, 1\}$ and $\tilde{P}_{0\to i}$ is the probability that the system returns into its initial state i under the time-reversed procedure (which always starts in state 0). Finally:

$$M_{1\to 0} = \tilde{P}_{0\to 1} \quad \text{and} \quad M_{0\to 0} = \tilde{P}_{0\to 0} \tag{34}$$

Experimental data are shown in figure 11. $M_{1\to 0}$ is an increasing function of τ whereas $M_{0\to 0}$ is decreasing with τ. Their sum is always close to 1 (which is the same result that gives $\Delta F_{\text{eff}} \approx k_B T \ln 2$), and $M_{1\to 0}$ is always smaller than $M_{0\to 0}$. These observations are intuitive because the work is higher when the bead jumps from state 1 to 0 than when it stays in state 0, and the work is higher when τ is smaller. The interpretation of eq. 34 gives a little more information. It is indeed reasonable to think that for time-reversed procedures the probability of returning to state 1 is small for fast procedures and increases when τ is bigger, whereas the probability of returning to state 0 increase when τ is smaller[9].

To be more quantitative one has to measure $\tilde{P}_{0\to 1}$ and $\tilde{P}_{0\to 0}$, but the time-reversed procedure cannot be realised experimentally, because it starts with a very fast rising of the force, which cannot be reached in our experiment. Thus, we performed numerical simulations, where it is possible to realise the corresponding time-reversed procedure and to compute $\tilde{P}_{0\to 1}$ and $\tilde{P}_{0\to 0}$. We simply integrate eq. 3 with Euler's method, for different set of parameters as close as possible to the experimental ones. The Gaussian white noise is generated by the "randn" function from Matlab® (normally distributed pseudorandom numbers). For each set of parameters we repeat the numerical procedure a few thousand of times. Some results are shown in table 1 (values are estimated with errorbar ± 0.02):

τ (s)	$f_{\max}$ (fN)	$M_{1\to 0}$	$\tilde{P}_{0\to 1}$	$M_{0\to 0}$	$\tilde{P}_{0\to 0}$	P_S (%)	$P_{S\,\text{force}}$ (%)
5	37.7	0.17	0.16	0.86	0.84	97.3	99.8
10	28.3	0.29	0.28	0.74	0.72	96.6	99.3
20	18.9	0.42	0.41	0.63	0.59	94	97.1
30	18.9	0.45	0.43	0.59	0.57	94.4	97.7

Table 1: Results for simple numerical simulations of the experimental procedure.

[9] Of course $\tilde{P}_{0\to 1} + \tilde{P}_{0\to 0} = 1$.

All the qualitative behaviours observed in the experimental data are retrieved, and the agreement between $M_{i\to 0}$ and $\tilde{P}_{0\to i}$ is correct. It was also verified that for proportions of success $< 100\,\%$, if one takes all the trajectories, and not only the ones where the bead ends in the state 0, the classical Jarzynski equality is verified: $\left\langle \mathrm{e}^{-\beta W_{\mathrm{st}}} \right\rangle = 1$. This result means that the small fraction of trajectories where the bead ends the erasure procedure where it shouldn't, which represent sometimes less than $1\,\%$ of all trajectories, is enough to retrieve the fact that $\Delta F_{\mathrm{Jarzynski}} = 0$.

5 Conclusion

In conclusion, we have realised an experimental information erasure procedure with a 1-bit memory system, made of a micro-particle trapped in a double well potential with optical tweezers. The procedure uses an external drag force to reset the memory of the system in one state, *i.e.* erase the knowledge of previous state and lose information. We measured the proportion of success of erasure procedures with different duration τ and amplitude of the force f_{max}. These data were used to manually optimised the procedure, *i.e.* for different τ we found the lowest force which gives a good erasure of information. By varying the duration of the information erasure procedure τ, we were able to approach the Landauer's bound $k_{\mathrm{B}}T \ln 2$ for the mean dissipated heat by the system $\langle Q \rangle$. We have also shown that $\langle Q \rangle$ seems to decrease as $1/\tau$, which is in agreement with the theoretical prediction for an optimal information erasure procedure [1], and was later confirmed experimentally with a more controlled experimental system [8].

We have computed the stochastic work received by the system during the procedure, which is in our particular case equal to the heat dissipated by the action of the external force. We used a modified version of the Jarzynski equality [22] for systems ending in a non- equilibrium state to retrieve the generalised Landauer's bound for any proportion of success on the mean stochastic work received by the system. This relation has been tested experimentally, and we have shown that the exponential average of the stochastic work, computed only on the trajectories where the information is actually erased, reaches the Landauer's bound for any duration of the procedure.

We also used a detailed version of the Jarzynski equality [9] to consider each sub-procedure where the information is erased ($1 \to 0$ and $0 \to 0$) independently. This relation allowed us to link the exponential average of stochastic work, computed only on a subset of the trajectories (corresponding to one of the sub-procedures), to the probability that the system returns to its initial state under a time- reversed procedure. We have shown that the experimental data are qualitatively in agreement with this interpretation. Finally, we used some very simple numerical simulations of our experimental procedure to compare quantitatively the partial exponential averages to the probabilities that the system returns in its initial state under time-reversed procedures.

Appendix

Equation 23 is obtained directly if the system is considered as a two state system, but it also holds if we consider a bead that can take any position in a continuous 1D double potential along the x-axis. We place the reference $x = 0$ at the center of the double potential.
Equation 22 states:

$$\left\langle e^{-\beta W_{\rm st}(t)} \right\rangle_{(x,t)} = \frac{\rho^{\rm eq}(x, \lambda(t))}{\rho(x,t)} e^{-\beta \Delta F_{\rm Jarzynski}(t)} \tag{35}$$

where $\langle . \rangle_{(x,t)}$ is the mean on all the trajectories that pass through x at t.
We choose $t = T_{\rm low} + \tau$ the ending time of the procedure, and we will not anymore write the explicit dependence upon t since it's always the same chosen time. We recall that $\Delta F_{\rm Jarzynski} = 0$ at $t = T_{\rm low} + \tau$ for our procedure.
We define the proportion of success, which is the probability that the bead ends its trajectory in the left half-space $x < 0$:

$$P_{S\,\rm force} = \rho(x < 0) = \int_{-\infty}^{0} {\rm d}x\, \rho(x) \tag{36}$$

The conditional mean is given by:

$$\left\langle e^{-\beta W_{\rm st}} \right\rangle_x = \int {\rm d}W_{\rm st}\, \rho(W_{\rm st}|x) e^{-\beta W_{\rm st}} \tag{37}$$

where $\rho(W_{\rm st}|x)$ is the conditional density of probability of having the value $W_{\rm st}$ for the work, knowing that the trajectory goes through x at the chosen time τ.
We recall from probability properties that:

$$\rho(W_{\rm st}|x) = \frac{\rho(W_{\rm st}, x)}{\rho(x)} \tag{38}$$

where $\rho(W_{\rm st}, x)$ is the joint density of probability of the value $W_{\rm st}$ of the work and the position x through which the trajectory goes at the chosen time τ.
We also recall:

$$\rho(W_{\rm st}|x < 0) = \frac{\int_{-\infty}^{0} {\rm d}x\, \rho(W_{\rm st}, x)}{\int_{-\infty}^{0} {\rm d}x\, \rho(x)} = \frac{\int_{-\infty}^{0} {\rm d}x\, \rho(W_{\rm st}, x)}{P_{S\,\rm force}} \tag{39}$$

Then by multiplying equation 35 by $\rho(x)$ and integrating over the left half-space $x < 0$ we have:

$$\int_{-\infty}^{0} {\rm d}x\, \rho(x) \left\langle e^{-\beta W_{\rm st}} \right\rangle_x = \int_{-\infty}^{0} {\rm d}x\, \rho^{\rm eq}(x) \tag{40}$$

Since the double potential is symmetric $\int_{-\infty}^{0} \mathrm{d}x \, \rho^{\mathrm{eq}}(x) = \frac{1}{2}$.
By applying definition 37 and equality 38, it follows:

$$\int_{-\infty}^{0} \mathrm{d}x \int \mathrm{d}W_{\mathrm{st}} \, \rho(W_{\mathrm{st}}, x) \mathrm{e}^{-\beta W_{\mathrm{st}}} = \frac{1}{2} \tag{41}$$

Then using equality 39:

$$P_{S \, \mathrm{force}} \int \mathrm{d}W_{\mathrm{st}} \, \rho(W_{\mathrm{st}} | x < 0) \mathrm{e}^{-\beta W_{\mathrm{st}}} = \frac{1}{2} \tag{42}$$

Finally we obtain:

$$\left\langle \mathrm{e}^{-\beta W_{\mathrm{st}}} \right\rangle_{x(T_{\mathrm{low}}+\tau)<0} = \frac{1/2}{P_{S \, \mathrm{force}}} \tag{43}$$

which is equation 23 of the main text.

References

[1] Erik Aurell, Krzysztof Gawędzki, Carlos Mejía-Monasterio, Roya Mohayaee, and Paolo Muratore-Ginanneschi. Refined second law of thermodynamics for fast random processes. *Journal of statistical physics*, 147(3):487–505, 2012.

[2] Charles H Bennett. The thermodynamics of computation—a review. *International Journal of Theoretical Physics*, 21(12):905–940, 1982.

[3] Antoine Bérut, Artak Arakelyan, Artyom Petrosyan, Sergio Ciliberto, Raoul Dillenschneider, and Eric Lutz. Experimental verification of landauer's principle linking information and thermodynamics. *Nature*, 483(7388):187–189, 2012.

[4] Leon Brillouin. *Science and information theory.* Academic Press, Inc., New York, 1956.

[5] A. Bérut, A. Petrosyan, and S. Ciliberto. Detailed jarzynski equality applied to a logically irreversible procedure. *EPL (Europhysics Letters)*, 103(6):60002, 2013.

[6] Raoul Dillenschneider and Eric Lutz. Memory erasure in small systems. *Phys. Rev. Lett.*, 102:210601, May 2009.

[7] C. Jarzynski. Nonequilibrium equality for free energy differences. *Phys. Rev. Lett.*, 78:2690–2693, Apr 1997.

[8] Yonggun Jun, Mom čilo Gavrilov, and John Bechhoefer. High-precision test of landauer's principle in a feedback trap. *Phys. Rev. Lett.*, 113:190601, Nov 2014.

[9] R. Kawai, J. M. R. Parrondo, and C. Van den Broeck. Dissipation: The phase-space perspective. *Phys. Rev. Lett.*, 98:080602, Feb 2007.

[10] Rolf Landauer. Irreversibility and heat generation in the computing process. *IBM journal of research and development*, 5(3):183–191, 1961.

[11] Harvey Leff and Andrew F Rex. *Maxwell's Demon 2 Entropy, Classical and Quantum Information, Computing.* CRC Press, 2002.

[12] Harvey S Leff and Andrew F Rex. *Maxwell's demon: entropy, information, computing.* Princeton University Press, 1990.

[13] Juan MR Parrondo, Jordan M Horowitz, and Takahiro Sagawa. Thermodynamics of information. *Nature Physics*, 11(2):131–139, 2015.

[14] Oliver Penrose. *Foundations of statistical mechanics: a deductive treatment*. Pergamon Press, Oxford, 1970.

[15] É Roldán, Ignacio A Martínez, Juan MR Parrondo, and Dmitri Petrov. Universal features in the energetics of symmetry breaking. *Nature Physics*, 10(6):457–461, 2014.

[16] Udo Seifert. Stochastic thermodynamics, fluctuation theorems and molecular machines. *Reports on Progress in Physics*, 75(12):126001, 2012.

[17] Ken Sekimoto. Langevin equation and thermodynamics. *Progress of Theoretical Physics Supplement*, 130:17–27, 1998.

[18] Ken Sekimoto and Shin-ichi Sasa. Complementarity relation for irreversible process derived from stochastic energetics. *Journal of the Physical Society of Japan*, 66(11):3326–3328, 1997.

[19] Kousuke Shizume. Heat generation required by information erasure. *Phys. Rev. E*, 52:3495–3499, Oct 1995.

[20] Leo Szilard. Über die entropieverminderung in einem thermodynamischen system bei eingriffen intelligenter wesen. *Zeitschrift für Physik*, 53(11-12):840–856, 1929. English translation "On the decrease of entropy in a thermodynamic system by the intervention of intelligent beings" in *Behavioral Science*, vol. 9, no. 4, pp. 301–310, 1964.

[21] Shoichi Toyabe, Takahiro Sagawa, Masahito Ueda, Eiro Muneyuki, and Masaki Sano. Experimental demonstration of information-to-energy conversion and validation of the generalized jarzynski equality. *Nature Physics*, 6(12):988–992, 2010.

[22] S. Vaikuntanathan and C. Jarzynski. Dissipation and lag in irreversible processes. *EPL (Europhysics Letters)*, 87(6):60005, 2009.