

Probabilistic Work Extraction on a Classical Oscillator Beyond the Second Law

Nicolas Barros¹, Sergio Ciliberto², and Ludovic Bellon^{3*}

¹Université de Lyon, ENS de Lyon, CNRS, Laboratoire de Physique, F-69342 Lyon, France

 (Received 29 February 2024; accepted 25 June 2024; published 31 July 2024)

We demonstrate experimentally that, applying optimal protocols that drive the system between two equilibrium states characterized by a free energy difference ΔF , we can maximize the probability of performing the transition between the two states with a work W smaller than ΔF . The second law holds only on average, resulting in the inequality $\langle W \rangle \geq \Delta F$. The experiment is performed using an underdamped oscillator evolving in a double-well potential. We show that with a suitable choice of parameters the probability of obtaining trajectories with $W \leq \Delta F$ can be larger than 95%. Very fast protocols are a key feature to obtain these results, which are explained in terms of the Jarzynski equality.

DOI: 10.1103/PhysRevLett.133.057101

Introduction—Numerous experimental platforms that act on the micro and nano scales allow us to explore the laws of thermodynamics for systems with few degrees of freedom coupled with thermal baths [1]. Typical examples are experiments in colloids [2–7], electric circuits [8], single electron transistors [9–11], mechanical devices [12,13], and single molecules [14,15]. In such systems, fluctuations of physical quantities play a fundamental role, contrary to classical thermodynamics, which mostly considers averaged quantities. Stochastic thermodynamics provides the suited framework to describe single realizations of thermodynamic transformation, and the associated work and heat distribution. In particular, the second law does not apply at the level of a single realization, and one can observe local “violations” due to the stochastic nature of the system, where the system can extract work, or gain free energy, at no cost to the operator and with no information feedback (illustration in Fig. 1) [16–19].

In a system driven by an external control parameter λ from an initial to a final equilibrium state resulting in a free energy difference ΔF between the two states, the probability distribution of the work performed when changing λ is constrained by Jarzynski’s equality [20]

$$\langle e^{-\beta W} \rangle = e^{-\beta \Delta F}, \quad (1)$$

where $\beta = 1/k_B T$, with T the temperature of the system and k_B the Boltzmann constant. The expression of the second law can be recovered from the convexity of the exponential function: $\langle W \rangle \geq \Delta F$. We are interested here in realizations where the free energy of the system can be increased spending an amount of work $W < \Delta F$. The free energy stored in the system can be later used, resulting in probabilistic work extraction from the device.

These realizations have been theoretically [10] and experimentally studied [9] with a single electron transistor

with discrete energy levels, obtaining a probability of extracting work up to 65%. Using a mechanical oscillator, we want to illustrate this behavior using a fully classical continuous system. Our goal is to experimentally obtain an arbitrary large probability of having realizations with $W < \Delta F$. To this aim, we propose a protocol that tends to the optimal work distribution predicted by Refs. [10,11]: it should consist of only two peaks, with the most frequent one below ΔF .

Experimental setup and protocol—The experimental setup is a microcantilever, which behaves in the absence of external forces as an underdamped harmonic oscillator of stiffness k , resonance frequency $f_0 = 1200$ Hz and quality factor $Q = 10$. The deflection x of the cantilever is measured by interferometry [21]. The oscillator is in equilibrium with the surrounding air at room temperature T and subject to thermal fluctuations. The variance of x is

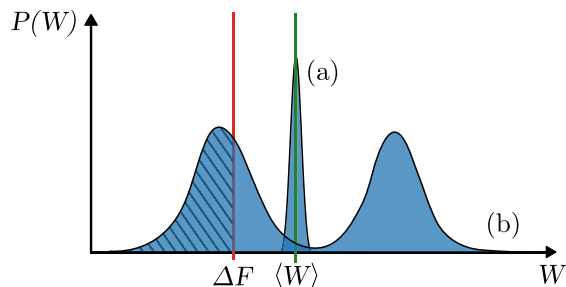


FIG. 1. Probability distribution function of the work $P(W)$ during a transformation. (a) In classical macroscopic thermodynamics, fluctuations are usually Gaussian and negligible compared with the work mean value $\langle W \rangle$ (green vertical line). (b) However, in stochastic thermodynamics, such fluctuations can have a more complex distribution (such as the bimodal one sketched here), and lead to local “violations” of the second law: several realizations (hatched area) of a transformation can be performed with a work W smaller than the free energy difference ΔF (red vertical line), even if in average $\langle W \rangle \geq \Delta F$. Note that in this sketch, the horizontal axis is rescaled between (a) and (b).

*Contact author: ludovic.bellon@ens-lyon.fr

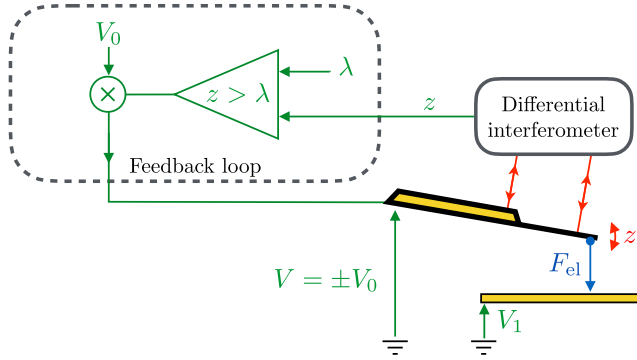


FIG. 2. Sketch of the experimental setup. We record the deflection z of a conductive cantilever measured with high precision by a differential interferometer [21]. Depending on the comparison of z with a tunable threshold λ , a constant voltage $V = \pm V_0$ is applied by a fast feedback loop, creating an electrostatic force F_{el} on the cantilever toward a nearby electrode at voltage $V_1 \gg V_0$.

$\sigma^2 = \langle x^2 \rangle = k_B T / k \sim 1 \text{ nm}^2$. σ is used as the unit length, and from now on all positions are expressed as dimensionless quantities $z = x/\sigma$, and energies in units of $k_B T$ [hence taking $\beta = 1$ in Jarzynski's equality, Eq. (1)].

To tune the potential experienced by the resonator, we use an electrostatic force acting on the cantilever. A time-dependent virtual potential can be created by a fast feedback loop [12,22] that adjusts the voltage V applied to the cantilever (thus the electrostatic force) depending on the measured position z (see Fig. 2). We implement the following simple rule: $V = -V_0$ if the position is below a threshold λ , and $V = +V_0$ otherwise. This creates an asymmetric double well, illustrated in Fig. 3. Two parameters are available to tune the potential: λ sets the barrier position and V_0 sets the centers of the two wells in $\pm z_0 \propto \pm V_0$. Theoretically, the potential energy constructed by this feedback is

$$U(z, \lambda, z_0) = \frac{1}{2} (z - S(z - \lambda)z_0)^2 + \lambda z_0 (S(z - \lambda) + S(\lambda)), \quad (2)$$

where S is the sign function: $S(z - \lambda) = -1$ if $z < \lambda$ and $S(z - \lambda) = 1$ if $z > \lambda$. To illustrate the validity of this model for U , we record a long time trace of the position in a static potential to evaluate the probability distribution function (PDF) of the position $P(z)$. We then reconstruct the double well using Boltzmann's prescription at equilibrium,

$$P(z) = \frac{1}{Z(\lambda, z_0)} e^{-U(z, \lambda, z_0)}, \quad (3a)$$

$$Z(\lambda, z_0) = \int_{-\infty}^{+\infty} e^{-U(z, \lambda, z_0)} dz, \quad (3b)$$

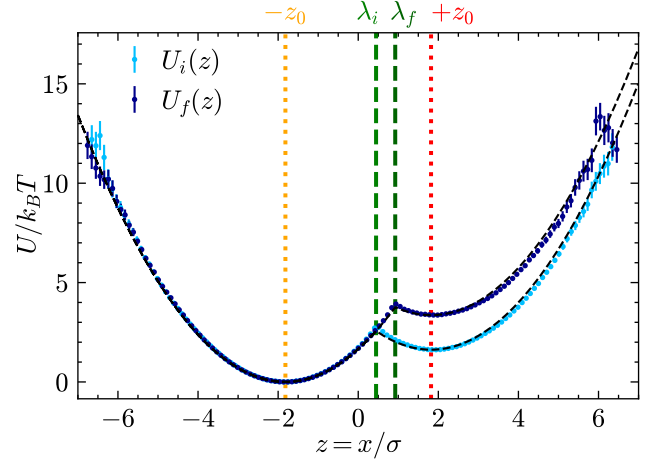


FIG. 3. Double-well potentials. The potential energy of the cantilever is the juxtaposition of two harmonic wells centered in $\pm z_0 = \pm 1.8$, the switch between the two occurring when $z = \lambda$. The two examples plotted here correspond to $\lambda_i = 0.45$ and $\lambda_f = 0.92$. The measured potential $U(z)$ data is inferred from Eq. (3) and the PDF of the experimental positions z during a long acquisition. The black dashed lines are best fit to Eq. (2), leading to the aforementioned values of z_0 and λ . The transformation we apply corresponds to a step of λ from λ_i to λ_f , leaving the lower well untouched while raising the upper one by $\Delta U = 2(\lambda_f - \lambda_i)x_0$.

with Z the partition function. As plotted in Fig. 3, the model is an excellent description of the effective potential evaluated through $U(z) = -\ln\{P(z)/P[-S(\lambda)z_0]\}$.

To observe local “violations” of the second law, we design the following protocol between an initial state (all quantities labeled by the subscript i) and a final state (subscript f). First, the cantilever evolves at equilibrium in an initial double-well potential $U_i(z) = U(z, \lambda_i, z_0)$. Then, we instantaneously increase the threshold λ between the two wells, from λ_i to λ_f . Finally, the cantilever is left in the final potential $U_f(z) = U(z, \lambda_f, z_0)$. The well centers are left unchanged during the protocol, z_0 keeping the same value. An example of the time trace of one realization of such protocol is plotted in Fig. 4.

Initial and final potentials are represented in Fig. 3. Each potential is measured from the equilibrium PDF $P(z)$ in the initial and final state, then fitted with very good accuracy using Eq. (2). The experimental values of the parameters λ_i , λ_f and z_0 are deduced from the fit. Potential wise, the protocol amounts to leaving the lower well unaffected while raising the upper one. The work W performed will be 0 most of the time (anytime the system is in the lower well, which is likely), while the free energy difference is positive [since $U(x)$ is globally increasing]: the probability of observing $W < \Delta F$ should thus be high.

The values of $z_0 = 1.8$ and $\lambda_f = 0.92$ are the same for all experiments. We tune the initial threshold λ_i with values going from 0 to 0.83. For each one of them, the protocol detailed in Fig. 4 is repeated around $N \sim 2500$ times.

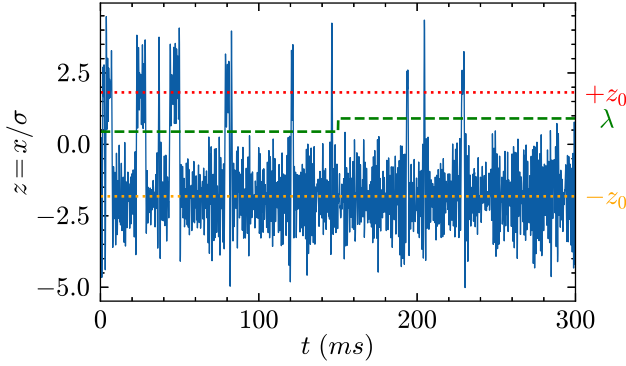


FIG. 4. Trajectory of an underdamped oscillator evolving in a time-dependent asymmetric double-well potential. The position z is expressed in units of standard deviation σ in a single well. The center of the wells is kept constant at $\pm z_0 = \pm 1.8$, but the commutation threshold λ is changed in less than $5 \mu\text{s}$ from $\lambda_i = 0.45$ to $\lambda_f = 0.92$ at $t = 150$ ms.

This allows us to obtain a good enough statistics to estimate the PDF $P(z)$, and further on the work distribution.

Analysis—Following the classical convention of stochastic thermodynamics [17], we define the work W done by the operator through the variations of the external parameters tuning the potential $U(z, \lambda, z_0)$,

$$W = \int \frac{\partial U}{\partial \lambda} \lambda dt. \quad (4)$$

Since z_0 is kept constant in all protocols, we do not consider the term in \dot{z}_0 , which is always zero here. W is computed using the recorded trajectories $z(t)$ and Eq. (2). Since the variation of $\lambda(t)$ is a step function, the work can be equivalently computed as $W = U_f(z) - U_i(z)$ using the value of z at the moment of the switch. From this expression we can also infer the theoretical expectation for $\langle W \rangle$,

$$\langle W \rangle = \int_{-\infty}^{+\infty} [U_f(z) - U_i(z)] \frac{1}{Z_i} e^{-U_i(z)} dz, \quad (5a)$$

$$= \frac{2z_0 e^{-2\lambda_i z_0}}{Z_i} [f(\lambda_f - z_0) - f(\lambda_i - z_0)], \quad (5b)$$

where f is defined by

$$f(u) = \sqrt{\pi/2} u \operatorname{erfc}(-u/\sqrt{2}) - e^{-u^2/2}. \quad (6)$$

$Z_i = Z(\lambda_i, z_0)$ is the partition function in the initial state, which can be computed from Eq. (3b) as

$$Z(\lambda, z_0) = \sqrt{\frac{\pi}{2}} \left[2 - \operatorname{erfc}\left(\frac{\lambda + z_0}{\sqrt{2}}\right) + e^{-2\lambda z_0} \operatorname{erfc}\left(\frac{\lambda - z_0}{\sqrt{2}}\right) \right]. \quad (7)$$

As shown in the Supplemental Material [23], the full distribution of the work $P(W)$ can actually be computed in this playground.

The free energy difference in the system $\Delta F = F_f - F_i$ can be computed in two different ways. For a given protocol, the work distribution obeys Jarzynski's equality [Eq. (1)], thus $\Delta F = -\ln\langle e^{-W} \rangle$ can be deduced from the experimental work distribution with good enough statistics. The second approach is to use the partition function Z at equilibrium: since the free energy of the system is $F = -\ln(Z)$, the free energy difference $\Delta F = \ln(Z_i/Z_f)$ can be theoretically directly calculated from the parameters of the initial and final potentials.

Results—The protocol is repeated $N \sim 2500$ times for ten values of λ_i ranging from 0 to 0.83. For each λ_i , we compute the work distribution. An example for $\lambda_i = 0.45$ is plotted in Fig. 5(a), where we also report the mean work $\langle W \rangle$ performed by the driving, and the difference in free energy ΔF . The work distribution consists mostly of two narrow peaks. The first one, for $W = 0$, corresponds to the cases where the cantilever is in the lower well ($z < \lambda_i$) of the potential during the step of λ . Since there is no change in the lower part of the potential, there is no energy cost. The second peak comes from the cases where $z > \lambda_f$: in this area, the potential is shifted by $\Delta U = W_{\max} = 2z_0(\lambda_f - \lambda_i)$, and driving the parameter λ implies this amount of work. We also observe some intermediate values, corresponding to $\lambda_i < z < \lambda_f$: the cantilever is initially in the upper well but ends in the lower one due to the step in λ , resulting in an intermediate energy change $\Delta U = 2z_0(z - \lambda_i)$, where z is the value of the deflection at the time of the switch.

As illustrated in Fig. 5(b), the second law is always satisfied ($\langle W \rangle > \Delta F$), but we manage to observe an almost arbitrary large proportion of transient violations ($W < \Delta F$). Indeed, by tuning the initial asymmetry of the potential, we can increase the probability of being in the lower well ($z < \lambda_i$) during the switch. The probability of observing a transient violation of the second law $P(W < \Delta F)$ can thus be arbitrary large. With our choice of parameters, for values of λ_i very close to the final value of λ_f , we manage to measure values of $P(W < \Delta F)$ of 95%. Our experimental results are in excellent agreement with the theoretical expectation, whose analytical expression is given in the Supplemental Material [23]. However, the free energy ΔF decreases when increasing λ_i : work extraction is more and more likely, but the gain with respect to ΔF decreases.

Another theoretical result that we can probe with our experiment is the inequality [24]

$$P(W \leq 0) \leq e^{-\Delta F}. \quad (8)$$

In our case, $P(W \leq 0) = P(W = 0)$ since W cannot be negative. Moreover, this probability is very close to $P(W \leq \Delta F)$ since the work distribution consists mainly of two peaks: the one in 0 and the other in W_{\max} , above ΔF . Some

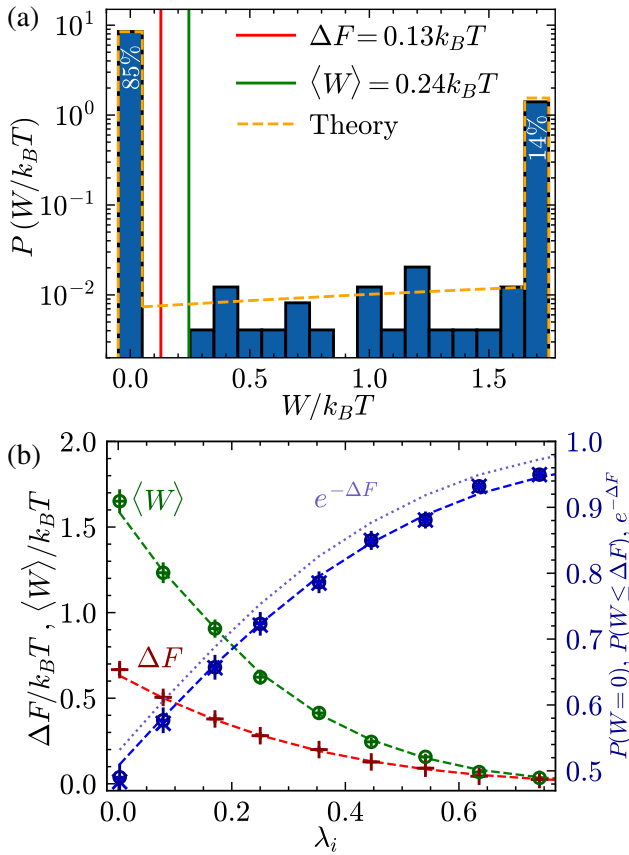


FIG. 5. (a) Work distribution for $N = 2450$ protocols, corresponding to the protocol in Fig. 3 ($\lambda_i = 0.45$, $\lambda_f = 0.92$). We also report the mean work $\langle W \rangle = 0.24 \pm 0.02$ and free energy difference $\Delta F = 0.13 \pm 0.01$ computed using Jarzynski’s equality. We observe a transient violation of the second law in 85% of the protocols (note the vertical log scale), when $W = 0$. The other values of the work are mostly $W = W_{\max}$ (right peak with 14% of the occurrences). The last 1% of the trajectories is distributed between the two extreme values. Although poorly sampled, it is compatible with the theoretical expectation [23] (dashed orange curve, including the two delta peaks plotted here using a bin width of 0.1). (b) Mean work $\langle W \rangle$ (green circle, top curve) and free energy difference ΔF (dark red plus, bottom curve) and for all initial conditions λ_i . Experiment (markers with error bars corresponding to 1 standard deviation of the statistical uncertainties) and theory (dashed lines) are in excellent agreement. The second law of thermodynamics $\langle W \rangle \geq \Delta F$ always holds, though the probability of observing the contrary on the single trajectory reaches 95% for the highest values of λ_i (blue circle marker, experiment, and dashed blue line, theory, right vertical scale). The probability of observing a zero work $P(W = 0)$ (blue cross) coincides with the one of transiently violating the second law, and is upper bounded by $e^{-\Delta F}$ (dotted blue line, right vertical scale), as predicted by Eq. (8) [24].

trajectories present a work in between the two peaks, but they are infrequent and poorly sampled in our experiment. The two probabilities $P(W \leq \Delta F)$ and $P(W = 0)$ therefore coincide within statistical uncertainties, as shown in Fig. 5(b). In this same figure, we plot the upper bound

given by Eq. (8), which is indeed confirmed in our experiment, and close to be saturated.

It is interesting to notice that in our system the dissipation (the quality factor Q) can be changed by controlling the pressure of the air surrounding the cantilever. However as we start from equilibrium and the work corresponds to an instantaneous change of the potentials, the dynamics of the system and the quality factor have no influence on the results plotted in Fig. 5: the work distribution and ΔF would be the same in an overdamped system.

Conclusion—We have shown, using a fully classical continuous mechanical system, how we can observe an arbitrary large number of apparent “violations” of the second law of thermodynamics, while being consistent with the rules of stochastic thermodynamics. We show a clear trade-off between this probability and the free energy gained during those events. We reach the theoretical limit of 95% probability (for our set of parameters) of having anomalous trajectories with $W < \Delta F$. This result is made possible by the specific way in which the potential is driven during the protocol: the center of the two wells does not change and only the minimum of the upper well is raised. We show in the Supplemental Material [23] that for constant stiffness wells, this protocol is the most efficient. Indeed it produces a work probability distribution with mainly two Dirac functions that matches the optimal distribution described in Refs. [10,11] to maximize the work extraction probability. One Dirac peak is centered in zero and corresponds to the trajectories that start in the lower well. The other is centered to a positive value of the work and corresponds to the trajectories starting on the upper well. In our experiment the relative amplitude of the two peaks, which determines the amount of “anomalous” trajectories, can be tuned by changing the minimum positions of the wells through the value of λ . In this way we have transformed for the transition properties a continuous classical system to a two level system using a protocol similar to the one of Ref. [11], which requires only a quench. In this context, the fast switch between the initial and final state is a key ingredient of the protocol, again as proposed in Refs. [10,11]. However, we do not use slow ramps as proposed in Ref. [10] and experimentally applied in Ref. [9]. Indeed a slow ramp would broaden the peaks of the work distribution, resulting in a situation similar to Fig. 1 and those described in the Supplemental Material [23]. For symmetric distributions as well, where the mean and the median are equal, observing $W < \Delta F$ is unlikely: the probability is smaller than 50%, this limit being reached in the reversible limit.

Let us conclude by saying that in spite of the fact there is an energy gain for 95% of the trajectories, the total mean remains positive. In order to use this energy surplus one should introduce a demon that selects the good trajectories, i.e., those starting on the lower well. Of course such a

demon would need some energy to elaborate the information gathered from the dynamics, and the second law of thermodynamics will still hold overall. However the advantage of this demon is that this energy loss could be spent remotely (in space or time) with respect to that of the standard operation of the system. It therefore decouples the transformation of the system from the necessary energy consumption that can be spent elsewhere or at some other time. This could, for instance, allow for a chemical reaction to stay cool during an exothermic transformation, or model some enzyme behavior in biological processes.

Acknowledgments—This work has been supported by project ANR-22-CE42-0022. We thank C. Jarzynski for enlightening physical discussions.

Data availability—The data supporting this study are openly available in Ref. [25].

-
- [1] S. Ciliberto, Experiments in stochastic thermodynamics: Short history and perspectives, *Phys. Rev. X* **7**, 021051 (2017).
- [2] V. Blickle, T. Speck, L. Helden, U. Seifert, and C. Bechinger, Thermodynamics of a colloidal particle in a time-dependent nonharmonic potential, *Phys. Rev. Lett.* **96**, 070603 (2006).
- [3] M. Gavrilov and J. Bechhoefer, Erasure without work in an asymmetric, double-well potential, *Phys. Rev. Lett.* **117**, 200601 (2016).
- [4] A. Béruit, A. Arakelyan, A. Petrosyan, S. Ciliberto, R. Dillenschneider, and E. Lutz, Experimental verification of Landauer’s principle linking information and thermodynamics, *Nature (London)* **483**, 187 (2012).
- [5] I. A. Martínez, E. Roldán, L. Dinis, D. Petrov, and R. A. Rica, Adiabatic processes realized with a trapped Brownian particle, *Phys. Rev. Lett.* **114**, 120601 (2015).
- [6] Y. Jun, M. Gavrilov, and J. Bechhoefer, High-precision test of Landauer’s principle in a feedback trap, *Phys. Rev. Lett.* **113**, 190601 (2014).
- [7] S. Toyabe, T. Sagawa, M. Ueda, M. Muneyuki, and M. Sano, Experimental demonstration of information-to-energy conversion and validation of the generalized Jarzynski equality, *Nat. Phys.* **6**, 988 (2010).
- [8] S. Ciliberto, A. Imparato, A. Naert, and M. Tanase, Heat flux and entropy produced by thermal fluctuations, *Phys. Rev. Lett.* **110**, 180601 (2013).
- [9] O. Maillet, P. A. Erdman, V. Cavina, B. Bhandari, E. T. Mannila, J. T. Peltonen, A. Mari, F. Taddei, C. Jarzynski, V. Giovannetti, and J. P. Pekola, Optimal probabilistic work extraction beyond the free energy difference with a single-electron device, *Phys. Rev. Lett.* **122**, 150604 (2019).
- [10] V. Cavina, A. Mari, and V. Giovannetti, Optimal processes for probabilistic work extraction beyond the second law, *Sci. Rep.* **6**, 29282 (2016).
- [11] C. Streibnig and H. Kantz, Apparent violations of the second law in two-level systems, *Phys. Rev. E* **100**, 052116 (2019).
- [12] S. Dago, J. Pereda, S. Ciliberto, and L. Bellon, Virtual double-well potential for an underdamped oscillator created by a feedback loop, *J. Stat. Mech.* (2022) 053209.
- [13] S. Joubaud, B. Garnier, N., and S. Ciliberto, Fluctuation theorems for harmonic oscillators, *J. Stat. Mech.* (2007) P09018.
- [14] D. Collin, F. Ritort, C. Jarzynski, S. Smith, I. Tinoco, and C. Bustamante, Verification of the Crooks fluctuation theorem and recovery of RNA folding free energies, *Nature (London)* **437**, 231 (2005).
- [15] M. Ribezzi-Crivellari and F. Ritort, Large work extraction and the Landauer limit in a continuous Maxwell demon, *Nat. Phys.* **15**, 660 (2019).
- [16] C. Jarzynski, Equalities and inequalities: Irreversibility and the second law of thermodynamics at the nanoscale, *Annu. Rev. Condens. Matter Phys.* **2**, 329 (2011).
- [17] K. Sekimoto, *Stochastic Energetics*, Lect. Notes Phys. Vol. 799 (Springer, Berlin, Heidelberg, 2010), 10.1007/978-3-642-05411-2.
- [18] U. Seifert, Stochastic thermodynamics, fluctuation theorems and molecular machines, *Rep. Prog. Phys.* **75**, 126001 (2012).
- [19] T. Sagawa, Thermodynamic and logical reversibilities revisited, *J. Stat. Mech.* (2014) P03025.
- [20] C. Jarzynski, Nonequilibrium equality for free energy differences, *Phys. Rev. Lett.* **78**, 2690 (1997).
- [21] P. Paolino, F. Aguilar Sandoval, and L. Bellon, Quadrature phase interferometer for high resolution force spectroscopy, *Rev. Sci. Instrum.* **84**, 095001 (2013).
- [22] S. Dago, N. Barros, J. Pereda, S. Ciliberto, and L. Bellon, Virtual potential created by a feedback loop: Taming the feedback demon to explore stochastic thermodynamics of underdamped systems, in *Crossroad of Maxwell Demon*, edited by X. Bouju and C. Joachim (Springer Nature Switzerland, Cham, 2024), pp. 115–135, 10.1007/978-3-031-57904-2_6.
- [23] See Supplemental Material at <http://link.aps.org/supplemental/10.1103/PhysRevLett.133.057101> for details on the computation of the probability distribution function of the work W . This computation is presented for the simple protocol used in the Letter, extended to a more general instantaneous transformation of the biquadratic potential energy, and explains why the protocol of the Letter is close to optimal.
- [24] C. Jarzynski, Nonequilibrium work relations: Foundations and applications, *Eur. Phys. J. B* **64**, 331 (2008).
- [25] N. Barros and L. Bellon, Dataset: Probabilistic work extraction on a classical oscillator beyond the second law, Zenodo, 10.5281/zenodo.10721406 (2024).