Editors' Suggestion

Optimal Probabilistic Work Extraction beyond the Free Energy Difference with a Single-Electron Device

Olivier Maillet,[1,*] Paolo A. Erdman,[2] Vasco Cavina,[2] Bibek Bhandari,[2] Elsa T. Mannila,[1] Joonas T. Peltonen,[1] Andrea Mari,[2] Fabio Taddei,[2] Christopher Jarzynski,[3] Vittorio Giovannetti,[2] and Jukka P. Pekola[1]

[1]*QTF Centre of Excellence, Department of Applied Physics, Aalto University School of Science, P.O. Box 13500, 00076 Aalto, Finland*

[2]*NEST, Scuola Normale Superiore and Instituto Nanoscienze-CNR, I-56127 Pisa, Italy*

[3]*University of Maryland, College Park, Maryland 20742, USA*

(Received 15 October 2018; published 18 April 2019)

We experimentally realize protocols that allow us to extract work beyond the free energy difference from a single-electron transistor at the single thermodynamic trajectory level. With two carefully designed out-of-equilibrium driving cycles featuring kicks of the control parameter, we demonstrate work extraction up to large fractions of $k_B T$ or with probabilities substantially greater than $1/2$, despite the zero free energy difference over the cycle. Our results are explained in the framework of nonequilibrium fluctuation relations. We thus show that irreversibility can be used as a resource for optimal work extraction even in the absence of feedback from an external operator.

DOI: 10.1103/PhysRevLett.122.150604

The ongoing miniaturization of physical systems, together with advances in techniques for the conception and manipulation of small biological objects, has made the investigation of devices with few degrees of freedom possible. In such systems, fluctuations of physical quantities become comparable with or larger than their mean values. This property, in particular, has led to the theoretical [1,2] and experimental [3–5] development of stochastic thermodynamics [6], which considers single realizations of work and heat relative to a given transformation rather than averaged quantities over an ensemble of realizations, as for the case of macroscopic systems. While the first law of thermodynamics (energy conservation) remains untouched, the second law (entropy increase over time) does not apply at the level of a single realization because of the stochastic nature of heat and work. Experimental platforms for stochastic thermodynamics include colloids [4,7], single-electron boxes [8], electronic double dots which allow entropy production measurements [9,10], and recently experiments attained the quantum regime [11] with, e.g., NMR setups [12] and superconducting circuits [13,14]. In this context, work and heat must be addressed in terms of probability distributions [6]. In particular, work fluctuations obey the equality [1]

$$\langle e^{-W/k_B T}\rangle = e^{-\Delta F/k_B T}. \tag{1}$$

Here W is the work *performed on a system* during a single realization of the process, ΔF is the free energy difference between the system's initial and final states, k_B is Boltzmann's constant, T is the temperature of the heat bath to which the system is connected, and angular brackets denote an ensemble average over realizations. From this equality the second law of thermodynamics is recovered, $\langle W\rangle \geq \Delta F$. Additionally, Eq. (1) implies that for some realizations $W < \Delta F$; i.e., the extracted work $(-W)$ exceeds the decrease in free energy $(-\Delta F)$. Equation (1) places no limits on the magnitude of such "violations" of the second law, nor on the net likelihood of observing these violations. Therefore, it is interesting to consider how to design a process to maximize the amount of work that might be extracted during a single realization, or alternatively to maximize the net probability to extract work beyond the free energy difference.

With the exception of recent applications of one-shot methods in this context [15,16], until now optimal control for a system coupled to a single heat bath has been mostly concerned with the trade-off between minimizing either fluctuations or average work [17,18]. Recently, it has been shown with a quantum jump approach [19] that with a suitable far-from-equilibrium driving sequence, one can instead take advantage of fluctuations to force work extraction from a system by an arbitrarily large value with a nonzero probability while still obeying Eq. (1). In particular, Ref. [19] discusses how to perform this task in the most efficient way, finding an optimal sequence that relies on two quasistatic tuning steps of the control parameter, separated by the sudden change of its energy level spacing, also referred to as a "quench." Such a protocol maximizes the probability of extracting work beyond a given quantity (i.e., $W \leq W^-$, where $W^- < \Delta F$ is fixed) while ensuring that we never perform work exceeding a selected threshold W^+.

In this Letter, using a single-electron transistor (SET) [20], we experimentally demonstrate a significant probability of extracting work arbitrarily bigger than the free energy difference in a single protocol realization. We first

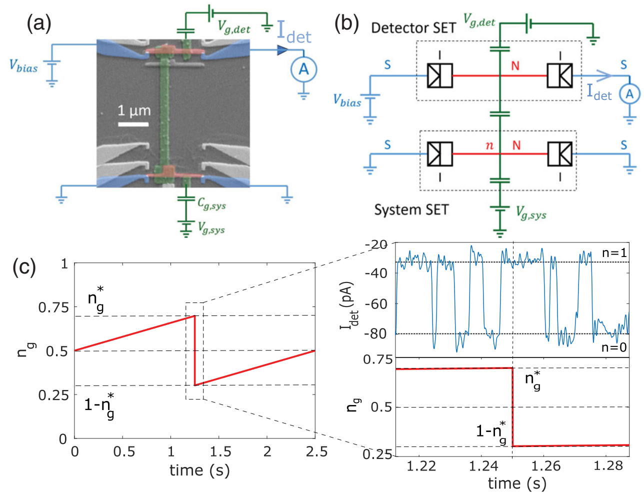

FIG. 1. (a) Scanning electron micrograph of the single-electron transistor (SET) capacitively coupled to a voltage-biased detector SET. Leads (blue) made of superconducting aluminum are coupled through oxide (tunnel) barriers to the copper (red) island. (b) Electrical circuit representation. (c) Protocol used to maximize work extraction, with a zoom on the detector SET output current under system driving, around the quench event.

show, in a simple symmetric configuration of the proposed protocol, that the resulting work probability distribution follows the bounds derived in Ref. [19], thus being optimal in the sense defined above. Building on this experimental proof, we arrange the protocol in such a way that the probability of extracting work just above the free energy difference is maximized, regardless of the energy cost in case of failure. We thus observe a probability *significantly greater than* $1/2$ of extracting work above the free energy difference, up to 65%, with the second law requirement $\langle W\rangle \geq \Delta F$ always satisfied. Quantitative agreement is found with both the nonequilibrium fluctuation relation [Eq. (1)] and predictions obtained from a master equation. These results are obtained without using the information on the system's state, unlike in a "Maxwell's demon" [21,22] experiment.

The system [see Fig. 1(a) for a micrograph and Fig. 1(b) for a full circuit representation] is an SET fabricated through multilayer shadow evaporation [23], made of a copper island of dimensions $2000 \times 200 \times 25\ \mathrm{nm}^3$, weakly coupled through oxide tunnel barriers to superconducting aluminum leads, under zero bias. Tunnel barriers allow electron quasiparticle transport in and out of the island. Heat is carried by these electrons, and electron-electron and electron-phonon interactions take place in the island at a much faster rate than tunneling events, ensuring that a constant electronic temperature T can be defined at any time [24]. The number n of excess charges in the island is our relevant degree of freedom, and the inverse tunneling rate sets the typical timescale of the system. The oxide barrier is opaque enough (the estimated tunneling resistance is $R_T \simeq 5\mathrm{M}\Omega$ for each junction, the sum of both capacitances being $C_\Sigma \approx 0.7$ fF) so that its combination with superconducting reservoirs leads to low tunneling rates at zero bias, enabling measurements with a low-frequency apparatus. The electrostatic energy of the island can be tuned by an external gate voltage $V_{g,\mathrm{sys}}$ through a gate electrode, which is patterned under the island and separated from it by a 50 nm oxide layer, forming a capacitance $C_{g,\mathrm{sys}} = 0.08\ \mathrm{fF} \ll C_\Sigma$. In this configuration, the Hamiltonian of the system takes a simple form [8],

$$H(n, n_g) = E_C(n - n_g)^2, \tag{2}$$

where $n_g = C_{g,\mathrm{sys}} V_{g,\mathrm{sys}}/e$ is the reduced gate voltage and $E_C \approx e^2/2C_\Sigma$ is the charging energy—i.e., the energy cost of adding one electron to the island due to Coulomb interaction, which sets the energy scale of the problem. The sample is cooled down to millikelvin temperatures in a dilution refrigerator; thus, the ratio $E_C/k_B = 1.3$ K is high enough so that we can restrict our analysis to two states $n = 0, 1$ [25], and the tunneling resistance is high enough to consider a sequential tunneling description. The system SET is capacitively coupled via a bottom gate electrode to another SET used as an electrometer monitoring tunneling events, and hence $n(t)$. The detector SET is biased with low enough voltage so that we can modulate its output current I_{det} with an external gate voltage $V_{g,\mathrm{det}}$ between zero and (typically) 100 pA. $V_{g,\mathrm{det}}$ is chosen to maximize the slope of current modulation $|dI_{\mathrm{det}}/dV_{g,\mathrm{det}}|$. This allows for maximum sensitivity to charge variation on the system island: due to the coupling gate electrode [green vertical element in Fig. 1(a)], electrons tunneling in or out of the system island at random times change the effective gate voltage seen by the detector SET, hence modulating its output current, which takes two values corresponding to the two charge states of the system. At charge degeneracy $n_g = 1/2$, where the states $n = 0$ and $n = 1$ are equiprobable (no charging energy cost), these tunneling events occur at a rate $\Gamma_d = 230$ Hz. This is slow enough for the detector [26], which has a bandwidth ~1 kHz limited by the low-pass filtering of a current amplifier. The two charge states' occupation probabilities satisfy the detailed balance relation with an effective electron temperature $T = 670$ mK [27]. From the Hamiltonian (2), we know the net heat transfer $\Delta E \equiv \Delta E_{0\to 1} = H(1, n_g) - H(0, n_g)$ for an electron tunneling onto the island,

$$\Delta E_{0\to 1}(n_g) = E_C(1 - 2n_g), \tag{3}$$

while the opposite heat transfer for an electron leaving the island is $\Delta E_{1\to 0}(n_g) = -\Delta E_{0\to 1}(n_g)$. By monitoring tunneling events during a driving cycle, and recording the corresponding jump times $\{t_k\}$ and gate voltage values $\{n_g(t_k)\}$, we experimentally determine the total heat absorbed by the system over the thermodynamic cycle:

$Q = \sum_k \Delta E[n_g(t_k)]\Delta n_k$, where $\Delta n_k = \pm 1$ depending on whether the electron jumps into or out of the island. The initial and final values of n_g are both set to $1/2$ so that we operate on a closed thermodynamic cycle. This way, the net energy change and the free energy difference ΔF over the entire cycle are both zero, and energy conservation ensures that $W = -Q$. Thus, we can directly infer the experimental value of the work at the end of the cycle based on the record of the transitions over the full cycle; see Fig. 1(c).

We first realize the driving sequence $n_g(t)$ depicted in Fig. 1(c), referred to as protocol, over a time t_f. For a given choice of W^- and W^+ satisfying $W^- < \Delta F < W^+$, the protocol [19] is designed to maximize the probability to observe a work value $W \leq W^-$ (successful event), while ensuring that we never observe $W^+ \geq \Delta F$ (failure events). For the sake of simplicity, we consider the symmetric case, i.e., $W^- = -W^+$. First, we prepare the system at charge degeneracy, i.e., $n_g(0) = 1/2$, at thermal equilibrium. Then we drive the system with a quasistatic ramp over a time $t_1 \gg \Gamma_d^{-1}$ up to a value $n_g^* \equiv n_g(t_1) = 1/2 + \Delta n_g$, with $0 < \Delta n_g < 1/2$. Next, a rapid swap of the energy splitting is operated by suddenly driving the system to a value $1 - n_g^*$. This "quench" must be realized over a time $\Delta t_q \ll \Gamma_d^{-1}$ so that no tunneling occurs in this time interval. Finally, we return the system to charge degeneracy through a quasistatic ramp, over a time t_1, such that $2t_1 + \Delta t_q = t_f$ and $n_g(t_f) = 1/2$. The total work output at the end of one cycle, obtained theoretically in the ideal quasistatic limit, is [27]

$$W(\bar{n}) = (1 - 2\bar{n})\Delta E(n_g^*), \quad (4)$$

where $\bar{n} \equiv n(t_1)$ is the charge state at the quench onset, and $\Delta E(n_g^*) < 0$. Therefore, W is a stochastic variable taking two values $W^{\mp} = \pm\Delta E(n_g^*)$. Its distribution $P(W) = p^*\delta(W - W^+) + (1 - p^*)\delta(W - W^-)$ with $1/2 < p^* < 1$ [19] is solely dictated by the equilibrium occupation probabilities of the two charge states before the quench, which obey the Gibbs ensemble: the ground state (one extra electron on the island) has a probability $p^* = (1 + e^{\Delta E(n_g^*)/k_B T})^{-1}$, while the excited state (zero extra electrons) has a probability $1 - p^* = (1 + e^{-\Delta E(n_g^*)/k_B T})^{-1}$. The outcome is simple to interpret physically: as the two ramps are quasistatic, the amount of work performed during those segments can be considered merely in terms of the equilibrium occupation probabilities at each instant, and it is here equal to zero because of the protocol's symmetry. On the other hand, the work performed during the quench does depend on the charge state at the quench onset: if the system is in the ground state $\bar{n} = 1$, the quench turns it into an energetically unfavorable state [since $\Delta E(1 - n_g^*) > 0$], and thus positive work has to be provided by the gate voltage source during the quench. If instead the system is in the excited state before the quench, the latter

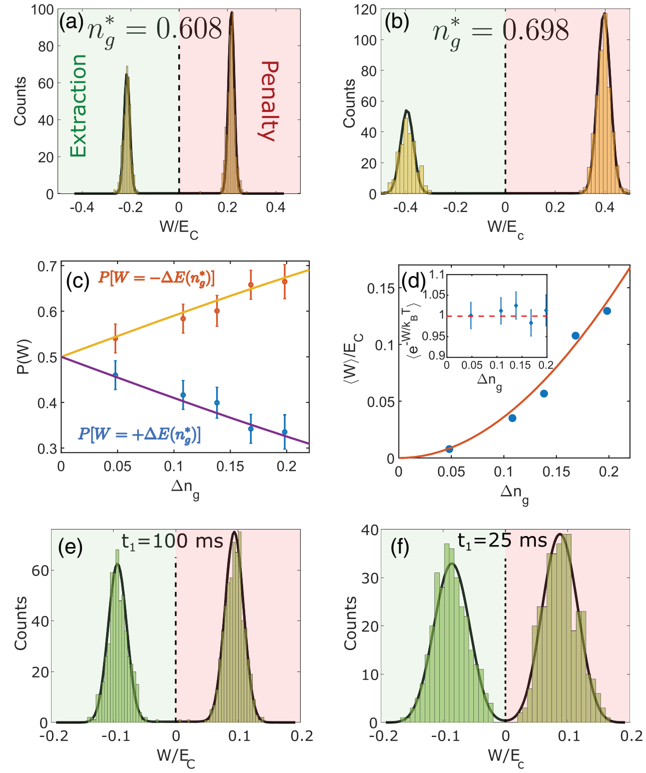

FIG. 2. (a),(b) Work histograms obtained for (a) $n_g^* = 0.608$ and (b) $n_g^* = 0.698$, with the same ramp time $t_1 = 1.25$ s. (c) Probability for $W = -\Delta E(n_g^*)$ (orange dots, mind the sign) and $W = \Delta E(n_g^*)$ (blue dots) events as a function of the quench amplitude. Solid lines are fits of Fermi functions (see text) with $E_C = 110$ μeV and $T = 670$ mK. Error bars are calculated from the number of protocol realizations. (d) Work performed on the system averaged over all outcomes as a function of the quench peak amplitude $\Delta n_g = n_g^* - 1/2$. The solid line is obtained from Eq. (5). *Inset:* Verification of Eq. (1) for all values of n_g. (e), (f) Work histograms obtained for the same quench amplitude $\Delta n_g = 0.048$, but with ramp times (e) $t_1 = 0.1$ s and (f) $t_1 = 0.025$ s, much shorter than in (a) and (b). In (a), (b), (e), and (f), solid lines are obtained by numerically solving the master equation [27]. All work values are normalized to E_C.

turns it into the ground state; thus, energy is released by the system as work, since there is no heat exchange during the quench. Thus, counterintuitively, the quench allows us to realize $W < \Delta F = 0$ by a possibly large amount by deliberately introducing irreversibility.

The protocol is repeated many times (∼1000) to experimentally map the work distribution. Because of the stray capacitance associated with the electrical setup, line filtering limits the quench time interval to $\Delta t_q = 0.3$ ms, still well below Γ_d^{-1}. Work histograms obtained for two different values of Δn_g (quench amplitudes) with the same ramp time are shown in Figs. 2(a) and 2(b). We indeed observe two peaks with maxima located at $\pm\Delta E(n_g^*)$. Their imbalance increases with the quench amplitude following Gibbs statistics, as seen in Fig. 2(c). This is expected, since the probability $1 - p^*$ to be in the excited state decreases as n_g^* gets further away from charge degeneracy. Namely, the ratio between the weights of the two peaks follows the detailed balance condition for the two energy states $\pm\Delta E(n_g^*)$: $P[W = \Delta E(n_g^*)]/P[W = -\Delta E(n_g^*)] = e^{\Delta E(n_g^*)/k_B T}$. Irreversibility, introduced by the quench, can be quantified by computing the work $\langle W\rangle = \int P(W) W dW$ performed on the system, averaged over all realizations:

$$\langle W\rangle = \Delta E(n_g^*) \tanh\left(\frac{\Delta E(n_g^*)}{2k_B T}\right). \quad (5)$$

Indeed, $\langle W\rangle \geq 0$, as expected from the second law of thermodynamics. In Fig. 2(d), we see that the experimental averaged work is positive and increasing with the quench amplitude, in good agreement with Eq. (5). The inset of Fig. 2(d) shows that our work histograms obey the nonequilibrium work relation [Eq. (1)].

Note that, in contrast to the theoretical situation [19], the peaks have a finite width in our experiment, which owes to the fact that a realistic ramp cannot be truly quasistatic, since one would need enough tunneling events between two infinitesimally close instants so that thermal equilibrium is properly defined at each instant t. Thus, the degree of reversibility is determined by the slope of the ramp with respect to the typical tunneling time, i.e., by $\Gamma_d^{-1}|dn_g/dt|$. For higher quench amplitudes but with the same ramping time, the residual irreversibility produces broader peaks [8], as Figs. 2(a) and 2(b) clearly show. We also run the protocol with constant quench amplitude but different ramp times. In Figs. 2(e) and 2(f), work histograms for two different ramp times unambiguously demonstrate that a shorter ramp time results in a broadened distribution, as captured through a master equation approach [8,27]. Indeed, we see in Fig. 2 that the obtained histograms are very well reproduced by the theoretical expectation, which validates this approach.

Next, building on this demonstration, we exhibit a protocol where the goal is to maximize the probability of exceeding the second law prescription (i.e., $W < \Delta F$), without any constraint on the energy cost of the failure events. This can be achieved with the protocol depicted in Fig. 3(a): we start at charge degeneracy, in thermal equilibrium, and ramp quasistatically the gate voltage up to a value $n_{g,a} > 1/2$. Then, in contrast with the previous protocol, we apply a quench such that the energy splitting is increased rather than reversed: over the quench time Δt_q, n_g is suddenly brought to $n_{g,b} > n_{g,a}$. In the last step, the system is brought back quasistatically to charge degeneracy. With this protocol, the work performed on the system over the cycle is [27],

$$W(\bar{n}) = k_B T(\Delta S)_q + (p_a - \bar{n})\Delta E(n_{g,a}) - (p_b - \bar{n})\Delta E(n_{g,b}), \quad (6)$$

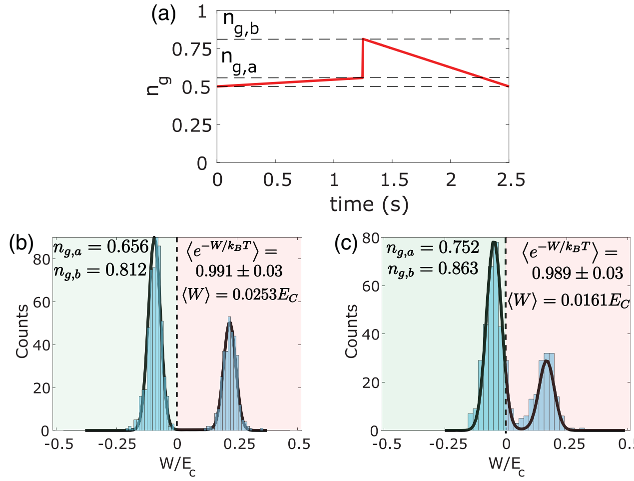

FIG. 3. (a) Protocol used to extract work with high probability (see text). (b),(c) Work histograms and experimental values of work and exponentiated work averages obtained for (b) $n_{g,a} = 0.656$, $n_{g,b} = 0.698$ and (c) $n_{g,a} = 0.752$, $n_{g,b} = 0.863$, with ramp time $t_1 = 1.25$ s. The vertical dashed line sets the zero free energy difference to guide the eye. Solid lines are obtained by numerically solving the master equation [27].

where p_a (p_b) is the $n = 1$ state equilibrium occupation probability right before (after) the quench. $(\Delta S)_q = S(p_b) - S(p_a)$ is the Shannon entropy difference between the equilibrium configurations before and after the quench, and $S(x) = -x \ln x - (1 - x) \ln(1 - x)$. S decreases during the quench, because the splitting and occupation asymmetry become larger. For $n_{g,b} > n_{g,a} > 1/2$, the sign of the work performed on the system is fully determined by the charge state at the quench onset [27]: $W(\bar{n} = 1) < 0 < W(\bar{n} = 0)$. Therefore, in this configuration, the probability of having $W < 0$ events is determined by the ground-state occupation probability $> 1/2$. Indeed, if the system is in the ground state at the quench onset, the entropy decrease associated with the quench is enough to have $W < 0$. In the opposite case, it is overwhelmed by the additional work required to further maintain the system in an even more unfavorable configuration. Note that this is not in contradiction with the second law of thermodynamics: from Eq. (6), one recovers again $\langle W\rangle > 0$, as confirmed experimentally together with Eq. (1), see Figs. 3(b) and 3(c). In Fig. 3(b), an example of a work histogram for such a protocol is shown. Here we indeed obtain more $W < 0$ events, but such events feature small work values, while $W > 0$ events result in large values of work performed on the system.

In principle, there is no bound strictly below 1 to the probability of having realizations with $W < 0$, since we can obtain a ground-state occupation probability arbitrarily close to 1 by ramping up the gate voltage towards the Coulomb blockade regime, i.e., $n_g \to 1$ (of course, in this case the work extracted is infinitesimally small). However, achieving this is difficult in practice, because for such n_g, the tunneling rates from the excited to the ground state are comparable with or larger than the detector's bandwidth [26]. In addition, for reasonable ramp times, driving up to higher n_g dissipates more energy. As a consequence, the peak containing $W < 0$ events, which is located close to 0, broadens until the events located at the right tail of the peak are transformed into $W > 0$ events, as shown in Fig. 3(c). For such events, the irreversibility associated with an imperfect quasistatic ramp overcomes the entropy decrease due to the quench. Despite these constraints, we were able to observe a probability of 65% for achieving $W < 0$, still significantly greater than $1/2$ [see Fig. 3(b)].

In conclusion, we have demonstrated that a substantial amount of work can be extracted with a non-negligible probability from a two-level system *coupled to a single heat bath*, using a SET driven far from equilibrium with a rapid quench. The driving cycle is designed to maximize either the work or the probability of extracting work from the system on one trajectory, by strongly amplifying work fluctuations rather than minimizing them, which represents a new paradigm for work extraction in mesoscopic engines. Our experimental results satisfy the nonequilibrium work relation and agree with a master equation approach which takes into account the irreversibility associated with finite time driving. We stress that even though work extraction can be favored, an external intervention (e.g., a Maxwell's demon [21]) would still be required to select only the extraction events: it is thanks to this absence that the second law remains valid, as we see experimentally. Appealing applications are foreseen if one optimizes the device: with a larger charging energy and bandwidth, using, e.g., a radio frequency detecting SET [28], it should be easier to obtain either very large work extraction or work extraction probabilities very close to 1. Moreover, the deviation from the quasistatic hypothesis leaves open the question of optimizing the protocol with respect either to the work fluctuations (i.e., the peak widths) or to the average values (the peak centers). Such a problem has received a lot of theoretical attention recently; for example, it has been shown that there is an analogy with first-order phase transitions between the protocols minimizing the two quantities [18]. Finally, the absence of quantum coherence in our system leaves open the question of probabilistic work extraction in the presence of quantum fluctuations and measurements [13,29], which could be addressed experimentally using, e.g., superconducting quantum bit circuits [13].

We thank L. B. Wang for technical help, as well as S. Ciliberto, R. Fazio, S. Singh, and I. M. Khaymovich for helpful discussions. This work was performed as part of the Academy of Finland Centre of Excellence program (Project No. 310257). This work has also been supported by the SNS-WIS joint lab QUANTRA, and by the CNR-CONICET cooperation program Energy Conversion in Quantum, Nanoscale, Hybrid devices. We acknowledge the provision of facilities by Aalto University at OtaNano-Micronova Nanofabrication Centre. C. J. acknowledges financial support from the U.S. Army Research Office under Contract No. W911NF-13-1-0390.

*olivier.maillet@aalto.fi

[1] C. Jarzynski, Nonequilibrium Equality for Free Energy Differences, Phys. Rev. Lett. **78**, 2690 (1997).

[2] G. E. Crooks, Entropy production fluctuation theorem and the nonequilibrium work relation for free energy differences, Phys. Rev. E **60**, 2721 (1999).

[3] G. M. Wang, E. M. Sevick, E. Mittag, D. J. Searles, and D. J. Evans, Experimental Demonstration of Violations of the Second Law of Thermodynamics for Small Systems and Short Time Scales, Phys. Rev. Lett. **89**, 050601 (2002).

[4] A. Bérut, 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] D. Collin, F. Ritort, C. Jarzynski, S. B. Smith, I. Tinoco, Jr., and C. Bustamante, Verification of the Crooks fluctuation theorem and recovery of RNA folding free energies, Nature (London) **437**, 231 (2005).

[6] U. Seifert, Stochastic thermodynamics, fluctuation theorems and molecular machines, Rep. Prog. Phys. **75**, 126001 (2012).

[7] É Roldán, I. A. Martínez, J. M. R. Parrondo, and D. Petrov, Universal features in the energetics of symmetry breaking, Nat. Phys. **10**, 457 (2014).

[8] O.-P. Saira, Y. Yoon, T. Tanttu, M. Möttönen, D. V. Averin, and J. P. Pekola, Test of the Jarzynski and Crooks Fluctuation Relations in an Electronic System, Phys. Rev. Lett. **109**, 180601 (2012).

[9] B. Küng, C. Rössler, M. Beck, M. Marthaler, D. S. Golubev, Y. Utsumi, T. Ihn, and K. Ensslin, Irreversibility on the Level of Single-Electron Tunneling, Phys. Rev. X **2**, 011001 (2012).

[10] S. Singh, É. Roldán, I. Neri, I. M. Khaymovich, D. S. Golubev, V. F. Maisi, J. T. Peltonen, F. Jülicher, and J. P. Pekola, Extreme reductions of entropy in an electronic double dot, Phys. Rev. B **99**, 115422 (2019).

[11] M. Campisi, P. Hänggi, and P. Talkner, Colloquium: Quantum fluctuation relations: Foundations and applications, Rev. Mod. Phys. **83**, 771 (2011).

[12] T. B. Batalhão, A. M. Souza, R. S. Sarthour, I. S. Oliveira, M. Paternostro, E. Lutz, and R. M. Serra, Irreversibility and the Arrow of Time in a Quenched Quantum System, Phys. Rev. Lett. **115**, 190601 (2015).

[13] N. Cottet, S. Jezouin, L. Bretheau, P. Campagne-Ibarcq, Q. Ficheux, J. Anders, A. Auffèves, R. Azouit, P. Rouchon, and B. Huard, Observing a quantum Maxwell demon at work, Proc. Natl. Acad. Sci. U.S.A. **114**, 7561 (2017).

[14] M. Naghiloo, J. J. Alonso, A. Romito, E. Lutz, and K. W. Murch, Information Gain and Loss for a Quantum Maxwell's Demon, Phys. Rev. Lett. **121**, 030604 (2018).

[15] D. Egloff, O. C. O. Dahlsten, R. Renner, and V. Vedral, A measure of majorization emerging from single-shot statistical mechanics, New J. Phys. **17**, 073001 (2015).

[16] N. Y. Halpern, A. J. P. Garner, O. C. O. Dahlsten, and V. Vedral, Introducing one-shot work into fluctuation relations, New J. Phys. **17**, 095003 (2015).

[17] T. Schmiedl and U. Seifert, Optimal Finite-Time Processes in Stochastic Thermodynamics, Phys. Rev. Lett. **98**, 108301 (2007).

[18] A. P. Solon and J. M. Horowitz, Phase Transition in Protocols Minimizing Work Fluctuations, Phys. Rev. Lett. **120**, 180605 (2018).

[19] V. Cavina, A. Mari, and V. Giovannetti, Optimal processes for probabilistic work extraction beyond the second law, Sci. Rep. **6**, 29282 (2016).

[20] G.-L. Ingold and Yu. V. Nazarov, Charge Tunneling Rates in Ultrasmall Junctions, in *Single Charge Tunneling: Coulomb Blockade Phenomena in Nanostructures*, edited by H. Grabert and M. H. Devoret, NATO ASI Series (Springer, Boston, MA, 1992), pp. 21–107.

[21] J. V. Koski, V. F. Maisi, J. P. Pekola, and D. V. Averin, Experimental realization of a Szilard engine with a single electron, Proc. Natl. Acad. Sci. U.S.A. **111**, 13786 (2014).

[22] T. Sagawa and M. Ueda, Generalized Jarzynski Equality under Nonequilibrium Feedback Control, Phys. Rev. Lett. **104**, 090602 (2010).

[23] T. A. Fulton and G. J. Dolan, Observation of Single-Electron Charging Effects in Small Tunnel Junctions, Phys. Rev. Lett. **59**, 109 (1987).

[24] F. Giazotto, T. T. Heikkilä, A. Luukanen, A. M. Savin, and J. P. Pekola, Opportunities for mesoscopics in thermometry and refrigeration: Physics and applications, Rev. Mod. Phys. **78**, 217 (2006).

[25] P. Lafarge, H. Pothier, E. R. Williams, D. Esteve, C. Urbina, and M. H. Devoret, Direct observation of macroscopic charge quantization, Z. Phys. B **85**, 327 (1991).

[26] O. Naaman and J. Aumentado, Poisson Transition Rates from Time-Domain Measurements with a Finite Bandwidth, Phys. Rev. Lett. **96**, 100201 (2006).

[27] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.122.150604 for detailed calculations and characterization.

[28] R. J. Schoelkopf, P. Wahlgren, A. A. Kozhevnikov, P. Delsing, and D. E. Prober, The radio-frequency single-electron transistor (RF-SET): A fast and ultrasensitive electrometer, Science **280**, 1238 (1998).

[29] C. Elouard, D. Herrera-Martí, B. Huard, and A. Auffèves, Extracting Work from Quantum Measurement in Maxwell's Demon Engines, Phys. Rev. Lett. **118**, 260603 (2017).