Maxwell Demon that Can Work at Macroscopic Scales

Nahuel Freitas and Massimiliano Esposito

Complex Systems and Statistical Mechanics, Department of Physics and Materials Science, University of Luxembourg, L-1511 Luxembourg, Luxembourg

Maxwell's demons work by rectifying thermal fluctuations. They are not expected to function at macroscopic scales where fluctuations become negligible and dynamics become deterministic. We propose an electronic implementation of an autonomous Maxwell's demon that indeed stops working in the regular macroscopic limit as the dynamics becomes deterministic. However, we find that if the power supplied to the demon is scaled up appropriately, the deterministic limit is avoided and the demon continues to work. The price to pay is a decreasing thermodynamic efficiency. Our Letter suggests that novel strategies may be found in nonequilibrium settings to bring to the macroscale nontrivial effects so far only observed at microscopic scales.

Introduction.—A famous thought experiment proposed by James C. Maxwell in 1867 introduced an agent with the ability to perceive the velocity of individual molecules in a gas, and open or close molecular valves based on its observations [1]. Such agent, or Maxwell's demon (MD), could in that way create gradients of temperature or pressure without expending any work, in apparent contradiction with the second law of thermodynamics. Indeed, the intention of Maxwell was to illustrate the statistical nature of this law, that according to him was only valid at macroscopic scales where we have "no power of handling or perceiving separate molecules" [1]. Later developments proved Maxwell to be partially wrong: the second law is statistical, but it is also valid at microscopic scales. Thus, a MD must necessarily dissipate energy and produce entropy, in such a way that the second law is still respected at a global level of description that includes the MD. Nevertheless, the resolution of the issue led to a revolution in our understanding of thermodynamics, that unveiled its deep connection with the abstract notion of information [2,3]. In modern terms, MDs or "information engines" are understood as active feedback control loops acting on a fluctuating system, and the second law has been successfully extended to those settings [4–11].

A first experimental realization of a MD was achieved 15 years ago in a molecular system powered by light [12]. After that, MDs were implemented in a great variety of systems such as trapped atoms [13–16], colloidal particles [17,18], single electron circuits [19–22], nuclear spins [23], electrophotonic systems [24], superconducting qubits [25–27], DNA hairpin pulling experiments [28], and cavity QED setups [29]. All those implementations share a common feature with the "very observant and neat-fingered" agent Maxwell imagined: they work at the microscopic level. They are able to detect fluctuations at the level of individual particles, molecules, electrons, or photons, depending on the setup. This is only natural, given that the task of a MD is to rectify thermal fluctuations, and those fluctuations are negligible at macroscopic scales (as is known from equilibrium thermodynamics, relative fluctuations of macroscopic observables decrease as $1/\sqrt{\Omega}$, where Ω is an adimensional scale parameter [30]). One exception are systems with continuous phase transitions, during which the symmetry of a system is spontaneously broken in favor of one of several metastable states that are macroscopically distinct. Since the final metastable state is selected by thermal fluctuations, continuous phase transitions can be employed to transfer information from the micro to the macro level [31], and macroscopic MDs can be devised based on this principle [32]. A simpler way to make the rectification effects of a MD detectable at the macroscopic level is to run a macroscopic number of individual MD in parallel. In fact, the implementation in [12] can be interpreted in this way.

In this Letter, we propose an alternative way to construct macroscopic MDs by realizing that thermodynamic resources can be employed to actively amplify thermal fluctuations from the micro to the macro level. Specifically, we propose a practical electronic implementation of an *autonomous* MD that, to the best of our knowledge, is the first one to have a well-defined macroscopic limit (which is nontrivial, in the sense that it cannot be considered a simple aggregation of individual devices). In addition, it only involves regular CMOS technology, the one powering modern computers. By only changing the width of the conduction channel of the MOS transistors, the proposed design can work from the single-electron regime all the way up to macroscopic regimes of operation. We find that when all additional parameters are fixed, the rectification effects associated with the action of the MD disappear above a
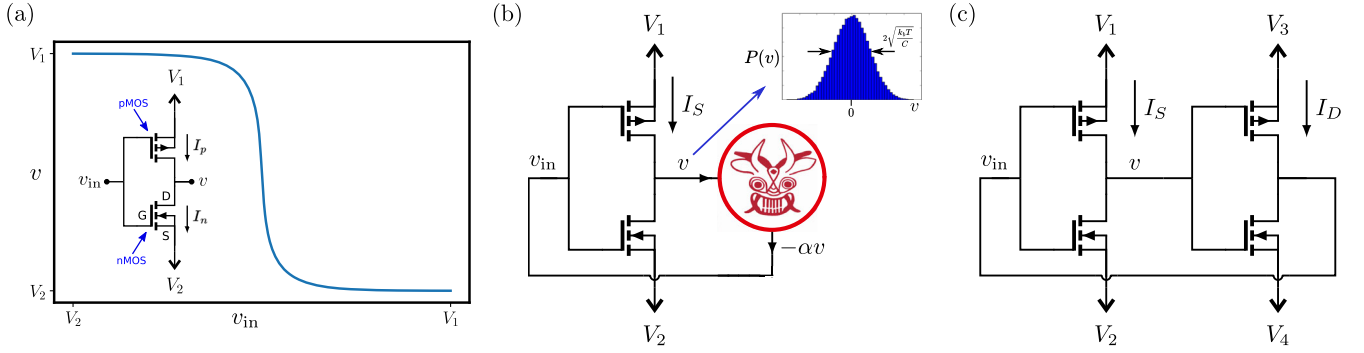
FIG. 1.   (a) Typical deterministic transfer function for the common implementation of a subthreshold CMOS inverter. (b) An external agent measures the fluctuating output voltage v of a CMOS inverter and adjusts the input voltage as $v_{\rm in} = -\alpha v$, with $\alpha > 0$. The histogram shows the Gaussian fluctuations at equilibrium ($V_1 = V_2 = 0$). (c) A similar feedback protocol is implemented with an additional CMOS inverter.

finite scale Ω^*. However, as we show analytically and numerically, when the power available to the MD is appropriately scaled, the rectification effects survive in the macroscopic limit. In that case, the price to pay is a decreasing thermodynamic efficiency, that we show scales as $\eta \sim 1/\Omega^2 \log(\Omega)$.

A full-CMOS Maxwell's demon.—Our design is based on the common implementation of a CMOS inverter, or NOT gate, shown in Fig. 1(a). It is composed of a nMOS transistor and a pMOS transistor operating in the subthreshold regime [33]. In a three-terminal nMOS (pMOS) transistor, electrical conduction between drain (D) and source (S) terminals is exponentially enhanced (reduced) for increasing gate (G) to source voltage. As a consequence, when a voltage bias is applied by connecting the drain terminals to voltage sources V_1 and $V_2 = -V_1$, the input-output transfer function of the inverter has the typical shape shown in Fig. 1(a). For positive values of the input, conduction through the nMOS transistor dominates and the output voltage approaches V_2. The situation is reversed for negative input voltages, and the output voltage approaches V_1. If no bias is applied, $V_1 = V_2 = 0$, the circuit attains thermal equilibrium and the steady state fluctuations in the output voltage v are given by $P_{\rm eq}(v) \propto e^{-\beta U(v)}$, where $\beta = 1/k_b T$ is the inverse temperature and $U(v) = v^2/2C$ the electrostatic energy. C is the output capacitance of the inverter, and is related to the physical scale of the transistors. Then, the mean output voltage is $\langle v \rangle = 0$, with a standard deviation $\sigma = \sqrt{k_b T/C}$. Let us consider the following feedback protocol, carried on by some external agent or device [see Fig. 1(b)]. At time t, the output voltage v is measured and a voltage $v_{\rm in} = -\alpha v$ is applied to the input, with $\alpha > 0$. If a positive fluctuation $v > 0$ is observed at time t, at subsequent times conduction through the upper pMOS transistor will be enhanced with respect to conduction through the nMOS transistor (since the input voltage at the gates will be negative). Therefore, the excess charge corresponding to the observed positive voltage

fluctuation will be most likely dissipated through the pMOS transistor, generating a net upward current. In a similar way, when a negative fluctuation $v < 0$ is observed, it will most probably be compensated by conduction events through the bottom nMOS transistor, also generating a net upward current. Thus, by the repeated application of the feedback protocol, charge can be made to flow in the upward direction. We can say that by controlling the input voltage, the MD opens or closes electronic doors according to the observed output voltage, forcing the electrons to flow in one direction, in full analogy to the original notion of a MD. A net upward electric current could also be observed even if a small downward bias $\Delta V_S = V_1 - V_2 > 0$ is applied. In that case, the entropy production rate $\dot{\Sigma}_S = \Delta V_S \langle I_S \rangle / T$ is negative ($\langle I_S \rangle$ is the average steady state electric current through the inverter, and therefore $\Delta V_S \langle I_S \rangle < 0$ is the rate of heat dissipation). According to the second law of thermodynamics, and our modern understanding of Maxwell's demons, that negative entropy production must be compensated by a positive entropy production in the device implementing the feedback control protocol (the MD). To study this, one must consider an explicit mechanism for the control device and turn to an autonomous picture that includes the MD. We do so by noticing that the operation $v_{\rm in} \to -\alpha v$ can be approximately implemented by an additional CMOS inverter, as shown in Fig. 1(c). This additional inverter is powered by a voltage bias $\Delta V_D = V_3 - V_4$, which generates a current I_D. If we break the left-right symmetry of the full circuit by choosing $\Delta V_D > \Delta V_S$, we can consider that the second inverter (the demon) continuously monitors the output of the first inverter, and adjusts its input accordingly. Whenever the demon produces an average current $\langle I_S \rangle$ against the bias ΔV_S, its efficiency is

$$\eta = -\dot{\Sigma}_S/\dot{\Sigma}_D \geq 0, \tag{1}$$

where $\Sigma_D = \Delta V_D \langle I_D \rangle / T$ is the rate of entropy production at the demon side. The total entropy production is $\dot{\Sigma} = \Sigma_S + \Sigma_D \geq 0$, which implies that $\eta \leq 1$.

We show next how to model the full circuit in Fig. 1(c) both at the deterministic and stochastic levels. Importantly, and at variance with the idealized feedback protocol described above, our modeling fully accounts for the intrinsic thermal noise at the demon side, and the delay between measurement and feedback.

Deterministic description.—All voltages will be expressed in units of the thermal voltage $V_T = k_b T/q_e$, where q_e is the absolute value of the electron charge. We consider powering voltages $V_1 = -V_2 = \Delta V_S/2$ and $V_3 = -V_4 = \Delta V_D/2$. The deterministic equations of motion for the voltages v and v_{in}—the 2 degrees of freedom in the circuit of Fig. 1(b)—are

$$C \, d_t v = I_p(v, v_{\text{in}}; \Delta V_S) - I_n(v, v_{\text{in}}; \Delta V_S),$$
$$C \, d_t v_{\text{in}} = I_p(v_{\text{in}}, v; \Delta V_D) - I_n(v_{\text{in}}, v; \Delta V_D), \quad (2)$$

where $I_{(n/p)}(v, v_{\text{in}}; \Delta V)$ are the electric currents through the (n/p)MOS transistors for given voltages v and v_{in} [of course, at steady state $I_n = I_p$ and we have single current through each inverter, see Fig. 1(a)]. In the subthreshold regime of operation, the pMOS current is given by [33]

$$I_p(v, v_{\text{in}}; \Delta V) = I_0 e^{(\Delta V/2 - v_{\text{in}} - V_{\text{th}})/n}(1 - e^{-(\Delta V/2 - v)}), \quad (3)$$

where I_0 (the specific current), V_{th} (the threshold voltage) and $n \geq 1$ (the slope factor), are phenomenological parameters characterizing the transistor. Assuming, for simplicity, that these parameters are the same for all transistors, the nMOS current is given by $I_n(v, v_{\text{in}}; \Delta V) = I_p(-v, -v_{\text{in}}; \Delta V)$. Close to thermal equilibrium (i.e., for low biases ΔV_S and ΔV_D), the previous deterministic dynamics has the unique fixed point $v = v_{\text{in}} = 0$, but the system becomes bistable when $\alpha^2 \equiv \alpha_S \alpha_D > 1$, where $\alpha_{S/D} \equiv (1 - e^{\Delta V_{S/D}/2})/n$. This is easily understood by noticing that, for small inputs, each inverter works as a linear amplifier with gain $\alpha_{S/D}$. The bistability can be exploited to store the value of a bit. Indeed, for symmetric powering $\Delta V_S = \Delta V_D = \Delta V$ the circuit in Fig. 1 is the typical CMOS implementation of SRAM memory cells. Note that the fixed points of Eqs. (2) always satisfy $V_1 > v > V_2$ and $V_3 > v_{\text{in}} > V_4$, implying that all electric currents are positive and thus that no rectification is possible at the deterministic level.

Stochastic description.—Conduction through the MOS transistors is not deterministic. The number of charges transported through the channel of a MOS transistor during a given time interval is a stochastic quantity that, in the subthreshold regime of operation, displays shot noise [34]. This can be described by an effective model that assigns two Poisson processes to every transistor, corresponding to elementary conduction events in the forward and backward directions. The two Poisson rates can be computed from the *I-V* curve characterization of the device and the requirement of thermodynamic consistency [35]. The dynamics

becomes a Markov jump process in the space $\{\boldsymbol{v}\}$ of possible circuit states and the probability distribution $P_t(\boldsymbol{v})$ over those states evolves according to the Markovian master equation

$$d_t P_t(\boldsymbol{v}) = \sum_\rho \lambda_\rho(\boldsymbol{v} - \boldsymbol{\Delta}_\rho v_e) P_t(\boldsymbol{v} - \boldsymbol{\Delta}_\rho v_e) - \lambda_\rho(\boldsymbol{v}) P_t(\boldsymbol{v}), \quad (4)$$

where $\boldsymbol{v} = (v, v_{\text{in}})$ is a vector with the 2 degrees of freedom of the circuit, $\lambda_\rho(\boldsymbol{v})$ are the Poisson jump rates assigned to the different transistors (that are indexed by ρ), and the vectors $\boldsymbol{\Delta}_\rho$ encode the change in $\boldsymbol{v}$ for each jump. For example, for forward transitions through the pMOS and nMOS transistors of the first inverter, we have $\boldsymbol{\Delta}_\rho = (1, 0)$ and $\boldsymbol{\Delta}_\rho = (-1, 0)$, respectively. The elementary voltage $v_e \equiv q_e/C$ is the magnitude by which the voltages change in each jump. The two Poisson rates for the pMOS I-V curve in Eq. (3) are

$$\lambda_+^p(v, v_{\text{in}}; \Delta V) = \tau_0^{-1} e^{(\Delta V/2 - v_{\text{in}})/n},$$
$$\lambda_-^p(v, v_{\text{in}}; \Delta V) = \lambda_+^p(v, v_{\text{in}}; \Delta V) e^{-(\Delta V/2 - v)} e^{-v_e/2}, \quad (5)$$

where $\tau_0 = (I_0 e^{-V_{\text{th}}/n}/q_e)^{-1}$ sets the system timescale. The explicit construction of the master equation for the circuit in Fig. 1(c) is explained in [36].

The steady state $P_{\text{ss}}(\boldsymbol{v})$ of Eq. (4) can be obtained by spectral methods or by generating stochastic trajectories via the Gillespie algorithm. From it, one can compute the average electric currents through the different transistors in the circuit. Typical steady state distributions are shown in the Supplemental Material [37].

Macroscopic limit.—We now study the stochastic behavior of the circuit as the physical dimensions of the transistors are increased. There are two relevant length scales: the width W and the length L of the conduction channel in each transistor [38]. For fixed L, both the parameter I_0 appearing in Eq. (3) (that enter the Poisson rates through the timescale τ_0) and the capacitance C are proportional to W. Thus, the Poisson rates $\lambda_\pm^{n/p}(\boldsymbol{v})$ increase as W, while the elementary voltage $v_e = q_e/C$ decreases as W^{-1} [36]. We therefore consider $v_e^{-1} \propto W$ as the scale parameter. From the previous observations, it follows that the probability distribution $P_t(\boldsymbol{v})$ satisfies a large deviations (LD) principle in the limit $v_e \to 0$. This means that fluctuations away from the deterministic behavior are exponentially suppressed in v_e^{-1}. Mathematically, the LD principle implies the existence of the limit $f(\boldsymbol{v}, t) = \lim_{v_e \to 0} - v_e \log[P_t(\boldsymbol{v})]$, also expressed as

$$P_t(\boldsymbol{v}) \asymp e^{-(f(\boldsymbol{v}, t) + o(v_e))/v_e}, \quad (6)$$

where $f(\boldsymbol{v}, t)$ is the rate function (i.e., the rate at which the probability of fluctuation v decreases with v_e^{-1}). Indeed, by plugging the previous ansatz in the master equation (4)

and keeping only the dominant terms in $v_e \to 0$, we find that the rate function must satisfy

$$d_t f(\mathbf{v}, t) = \sum_\rho \omega_\rho(\mathbf{v})[1 - e^{\boldsymbol{\Delta}_\rho \cdot \nabla f(\mathbf{v}, t)}], \qquad (7)$$

where $\omega_\rho(\mathbf{v}) = \lim_{v_e \to 0} v_e \lambda_\rho(\mathbf{v})$ are the scaled Poisson rates. The minima $\mathbf{v}(t)$ of the rate function $f(\mathbf{v}, t)$ at time t are the most probable values. They evolve according to the closed deterministic dynamics of Eq. (2) [39]. This shows how the deterministic dynamics emerges from the stochastic one.

Equation (7) cannot be solved exactly, even at steady state. In order to make analytical progress, we perform a Gaussian approximation around $\mathbf{v} = 0$, the most probable value in the monostable phase ($\alpha^2 < 1$). In this way we can compute the average steady state current through the first inverter to first nontrivial order in $v_e \to 0$ (see Supplemental Material [37]). While its full expression is complicated, to first order in ΔV_S it reads

$$\langle I_S \rangle \simeq \frac{q_e}{2\tau_0}[\Delta V_S - v_e(e^{\Delta V_D/2} - 1)]. \qquad (8)$$

The first term corresponds to the deterministic current $I_S = q_e \Delta V_S / 2\tau_0 + \mathcal{O}(\Delta V_S^2)$ which is incompatible with rectification (it is always positive for $\Delta V_S > 0$). The second term describes the effect of the fluctuations and the feedback control, and it is always negative. Then, we see that the minimum value of v_e for which rectification is possible (i.e., $\langle I_S \rangle < 0$) is $v_e^* = \Delta V_S / \alpha_D$. Thus, the rectification effects associated to the action of the MD disappear above a scale $1/v_e^*$, that increases with the amplification factor α_D of the MD, and decreases with the opposing bias ΔV_S. Similarly, the minimum value of ΔV_D necessary for rectification is

$$\Delta V_D^* = 2\log(1 + \Delta V_S / v_e). \qquad (9)$$

The previous expression is plotted with a dashed line in Fig. 2. The deviations from the exact $\langle I_S \rangle = 0$ boundary are due to non-Gaussian effects and the fact that Eq. (8) is only valid for first order in ΔV_S.

Scaling analysis.—Although we did not assume anything about the magnitude of ΔV_D, the Gaussian approximation leading to Eq. (8) is only valid for $\alpha^2 < 1$. This imposes a maximum value for ΔV_D given any nonzero value of ΔV_S. To take this into account, in the following we will consider a constant value of α^2, and rewrite Eq. (8) as

$$\langle I_S \rangle = \frac{q_e}{2\tau_0}\left[\Delta V_S - v_e \frac{\alpha^2}{e^{\Delta V_S/2} - 1}\right]. \qquad (10)$$

Recalling that the timescale τ_0 is proportional to v_e, we see that if we scale the system bias as $\Delta V_S = cv_e$ (for a
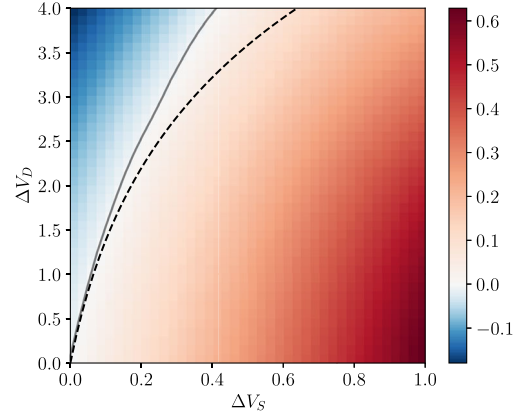

FIG. 2. Average electric current $\langle I_S \rangle$ (in units of q_e/τ_0) at steady state for the circuit of Fig. 1(c) as a function of the powering voltage biases ΔV_S and ΔV_D ($v_e = 0.1$, $n = 1$). The solid gray line marks the boundary between the $\langle I_S \rangle < 0$ (blue) and $\langle I_S \rangle > 0$ (red) regions, and the dashed line corresponds to the approximate analytical result of Eq. (9).

constant $c > 0$), the rectification power at the system side becomes independent of the scale

$$T\dot{\Sigma}_S = \Delta V_S \langle I_S \rangle \sim -\alpha^2 / c \quad \text{for } v_e \to 0. \qquad (11)$$

This result is remarkable since it implies that the mean value $\langle I_S \rangle$ remains negative in the macroscopic limit, and therefore does not reduce to the positive value obtained from the deterministic treatment. See the Supplemental Material [37] for an estimation of the order of magnitude of $\langle I_S \rangle$ for modern fabrication processes. The link between the macroscopic and deterministic limits is broken by the powering of the demon, needed to amplify thermal fluctuations. Indeed, in order to maintain the value of α^2 constant, ΔV_D must be $\Delta V_D = 2\log(1 + 2\alpha^2/cv_e)$. Then, the current through the demon inverter is $\langle I_D \rangle \simeq (q_e/\tau_0) \times (e^{\Delta V_D/2} - 1) = 2q_e\alpha^2/cv_e\tau_0$, and the dissipation power at the demon side scales as

$$T\dot{\Sigma}_D = \Delta V_D \langle I_D \rangle \sim \log(v_e^{-1})/v_e^2 \quad \text{for } v_e \to 0. \qquad (12)$$

As a consequence, the thermodynamic efficiency scales as $\eta \sim v_e^2/\log(v_e^{-1})$. This scaling is verified numerically by generating stochastic trajectories via the Gillespie algorithm and computing the average currents $\langle I_{S/D} \rangle$. The results are shown in Fig. 3. We see that a similar scaling is obtained even when the system is working at the critical point $\alpha^2 = 1$ or in the bistable phase $\alpha^2 > 1$. Thus, the rectification effects can survive the macroscopic limit if the power available to the MD is appropriately scaled, at the price of a decreasing efficiency. Finally, the scaling with temperature can be analyzed by recalling that all voltages are expressed in units of $V_T = k_b T/q_e$. Then, from Eq. (10) we see that $\langle I_S \rangle < 0$ only for $V_T > V_T^*$, where
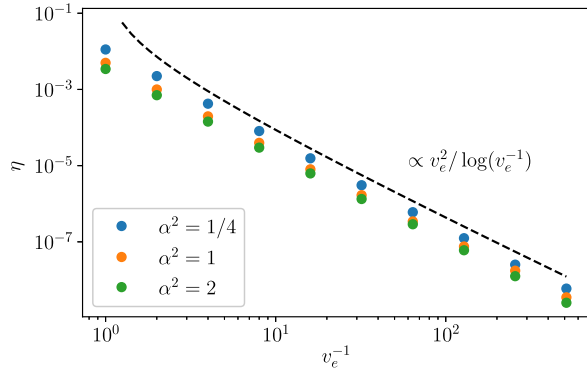
FIG. 3. Efficiency as a function of the scale parameter v_e^{-1} with the scaling strategy $\Delta V_S = c v_e$ and $\Delta V_D = 2 \log(1 + 2\alpha^2/cv_e)$ for $c = 0.1$ and different values of α^2. The results are obtained using the Gillespie algorithm. The dashed black line shows the scaling obtained from Eqs. (11) and (12).

$V_T^* = \Delta V_S / \log(1 + \alpha^2 v_e / \Delta V_S)$ sets the temperature below which the demon stops working.

Discussion.—Our results unveil a general mechanism via which the usual correspondence between the macroscopic and deterministic limits can be broken. As one increases the scale of a system, thermal fluctuations of macroscopic observables (in this case the voltage of a conductor) decrease, and this can be compensated by increasing the power dedicated to amplify them. In this way, microscopic thermal fluctuations can be transported to the macroscale where they can be employed to produce phenomena (in this case the inversion of a current) that are not allowed by the deterministic dynamics obtained from the macroscopic limit taken at fixed powering. In principle, the same mechanism could be implemented in other nonequilibrium systems, for example, chemical reaction networks. Finally, the scaling we obtained for the thermodynamic efficiency relies heavily on the details of the amplification step provided by the second inverter, which might not be optimal.

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

[2] Charles H. Bennett, The thermodynamics of computation-a review, Int. J. Theor. Phys. **21**, 905 (1982).

[3] Juan M. R. Parrondo, Jordan M. Horowitz, and Takahiro Sagawa, Thermodynamics of information, Nat. Phys. **11**, 131 (2015).

[4] Takahiro Sagawa and Masahito Ueda, Second Law of Thermodynamics with Discrete Quantum Feedback Control, Phys. Rev. Lett. **100**, 080403 (2008).

[5] Takahiro Sagawa and Masahito Ueda, Minimal Energy Cost for Thermodynamic Information Processing: Measurement and Information Erasure, Phys. Rev. Lett. **102**, 250602 (2009).

[6] F. J. Cao and M. Feito, Thermodynamics of feedback controlled systems, Phys. Rev. E **79**, 041118 (2009).

[7] Takahiro Sagawa and Masahito Ueda, Nonequilibrium thermodynamics of feedback control, Phys. Rev. E **85**, 021104 (2012).

[8] Massimiliano Esposito and Gernot Schaller, Stochastic thermodynamics for "Maxwell demon" feedbacks, Europhys. Lett. **99**, 30003 (2012).

[9] A. C. Barato and U. Seifert, Unifying Three Perspectives on Information Processing in Stochastic Thermodynamics, Phys. Rev. Lett. **112**, 090601 (2014).

[10] Jordan M. Horowitz and Massimiliano Esposito, Thermodynamics with Continuous Information Flow, Phys. Rev. X **4**, 031015 (2014).

[11] David Hartich, Andre C. Barato, and Udo Seifert, Stochastic thermodynamics of bipartite systems: Transfer entropy inequalities and a Maxwell's demon interpretation, J. Stat. Mech. (2014) P02016.

[12] Viviana Serreli, Chin-Fa Lee, Euan R. Kay, and David A. Leigh, A molecular information ratchet, Nature (London) **445**, 523 (2007).

[13] Gabriel N. Price, S. Travis Bannerman, Kirsten Viering, Edvardas Narevicius, and Mark G. Raizen, Single-Photon Atomic Cooling, Phys. Rev. Lett. **100**, 093004 (2008).

[14] Mark G. Raizen, Comprehensive control of atomic motion, Science **324**, 1403 (2009).

[15] S. Travis Bannerman, Gabriel N. Price, Kirsten Viering, and Mark G. Raizen, Single-photon cooling at the limit of trap dynamics: Maxwell's demon near maximum efficiency, New J. Phys. **11**, 063044 (2009).

[16] Aishwarya Kumar, Tsung-Yao Wu, Felipe Giraldo, and David S. Weiss, Sorting ultracold atoms in a three-dimensional optical lattice in a realization of Maxwell's demon, Nature (London) **561**, 83 (2018).

[17] 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, Nat. Phys. **6**, 988 (2010).

[18] Tushar K. Saha, Joseph N. E. Lucero, Jannik Ehrich, David A. Sivak, and John Bechhoefer, Maximizing power and velocity of an information engine, Proc. Natl. Acad. Sci. U.S.A. **118**, e2023356118 (2021).

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

[20] J. V. Koski, V. F. Maisi, T. Sagawa, and J. P. Pekola, Experimental Observation of the Role of Mutual Information in the Nonequilibrium Dynamics of a Maxwell Demon, Phys. Rev. Lett. **113**, 030601 (2014).

[21] J. V. Koski, A. Kutvonen, I. M. Khaymovich, T. Ala-Nissila, and J. P. Pekola, On-Chip Maxwell's Demon as an Information-Powered Refrigerator, Phys. Rev. Lett. **115**, 260602 (2015).

[22] Kensaku Chida, Samarth Desai, Katsuhiko Nishiguchi, and Akira Fujiwara, Power generator driven by Maxwell's demon, Nat. Commun. **8**, 15310 (2017).

[23] Patrice A. Camati, John P. S. Peterson, Tiago B. Batalhão, Kaonan Micadei, Alexandre M. Souza, Roberto S. Sarthour, Ivan S. Oliveira, and Roberto M. Serra, Experimental Rectification of Entropy Production by Maxwell's Demon in a Quantum System, Phys. Rev. Lett. **117,** 240502 (2016).

[24] Mihai D. Vidrighin, Oscar Dahlsten, Marco Barbieri, M. S. Kim, Vlatko Vedral, and Ian A. Walmsley, Photonic Maxwell's Demon, Phys. Rev. Lett. **116,** 050401 (2016).

[25] Nathanaël Cottet, Sébastien Jezouin, Landry Bretheau, Philippe Campagne-Ibarcq, Quentin Ficheux, Janet Anders, Alexia Auffèves, Rémi Azouit, Pierre Rouchon, and Benjamin Huard, Observing a quantum Maxwell demon at work, Proc. Natl. Acad. Sci. U.S.A. **114,** 7561 (2017).

[26] Y. Masuyama, K. Funo, Y. Murashita, A. Noguchi, S. Kono, Y. Tabuchi, R. Yamazaki, M. Ueda, and Y. Nakamura, Information-to-work conversion by Maxwell's demon in a superconducting circuit quantum electrodynamical system, Nat. Commun. **9,** 1291 (2018).

[27] 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).

[28] M. Ribezzi-Crivellari and F. Ritort, Large work extraction and the Landauer limit in a continuous Maxwell demon, Nat. Phys. **15,** 660 (2019).

[29] Baldo-Luis Najera-Santos, Patrice A. Camati, Valentin Métillon, Michel Brune, Jean-Michel Raimond, Alexia Auffèves, and Igor Dotsenko, Autonomous Maxwell's demon in a cavity QED system, Phys. Rev. Research **2,** 032025(R) (2020).

[30] R. K. Pathria and Paul Beale, *Statistical Mechanics* (Academic Press, Cambridge, MA, USA, 2021).

[31] É. Roldán, Ignacio A. Martinez, Juan M. R. Parrondo, and Dmitri Petrov, Universal features in the energetics of symmetry breaking, Nat. Phys. **10,** 457 (2014).

[32] Juan M. R. Parrondo, The Szilard engine revisited: Entropy, macroscopic randomness, and symmetry breaking phase transitions, Chaos **11,** 725 (2001).

[33] Alice Wang, Benton H. Calhoun, and Anantha P. Chandrakasan, *Sub-Threshold Design for Ultra Low-Power Systems* (Springer, New York, 2006), Vol. 95.

[34] Rahul Sarpeshkar, Tobias Delbruck, and Carver A. Mead, White noise in MOS transistors and resistors, IEEE Circuits Devices Mag. **9,** 23 (1993).

[35] Nahuel Freitas, Jean-Charles Delvenne, and Massimiliano Esposito, Stochastic Thermodynamics of Nonlinear Electronic Circuits: A Realistic Framework for Computing Around kT, Phys. Rev. X **11,** 031064 (2021).

[36] Nahuel Freitas, Karel Proesmans, and Massimiliano Esposito, Reliability and entropy production in nonequilibrium electronic memories, Phys. Rev. E **105,** 034107 (2022).

[37] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.129.120602 for additional details about (i) the computation of the steady state distribution, (ii) the analytical calculation under the Gaussian approximation, and (iii) estimations of orders of magnitude for modern fabrication processes.

[38] Yannis Tsividis, *Operation and Modeling of the MOS Transistor* (McGraw-Hill, Inc., New York, 1987).

[39] Nahuel Freitas, Gianmaria Falasco, and Massimiliano Esposito, Linear response in large deviations theory: A method to compute non-equilibrium distributions, New J. Phys. **23,** 093003 (2021).

Supplementary material for "A Maxwell demon that can work at macroscopic scales"

Nahuel Freitas[1] and Massimiliano Esposito[1]

[1] *Complex Systems and Statistical Mechanics, Department of Physics and Materials Science,*
University of Luxembourg, L-1511 Luxembourg, Luxembourg

I. DETERMINATION OF THE STEADY STATE DISTRIBUTION

To compute the steady state distribution for given parameters we first construct the generator M of the master equation given in the main text. For this, we truncate the state space to $(2N + 1)^2$ states by restricting v and v_{in} to the ranges $-N \leq v/v_e \leq N$ and $-N \leq v_{\text{in}}/v_e \leq N$. Then we encode the probability distribution $P_t(v, v_{\text{in}})$ into a vector $|P_t\rangle$, and construct the matrix M such that the master equation reads $d_t|P_t\rangle = M|P_t\rangle$. The steady state distribution is then obtained as the normalized right eigenvector of M with zero eigenvalue. Steady state distributions for different values of α^2 are shown in Figure 1.
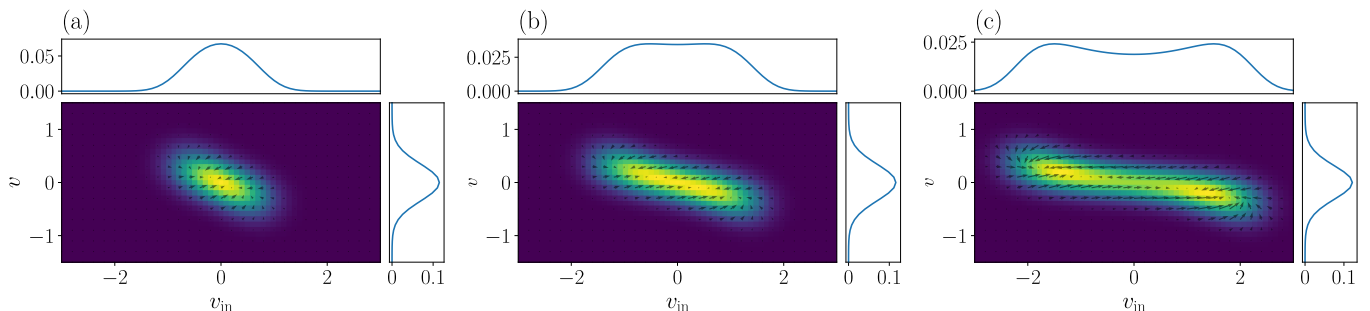

FIG. 1: Histograms of the steady state distribution $P_{\text{ss}}(\boldsymbol{v})$ for $v_e = 0.1$, $\Delta V_S = 0.4$ and: (a) $\Delta V_D = 2$ ($\alpha^2 < 1$, monostable phase), (b) $\Delta V_D \simeq 3.415$ ($\alpha^2 = 1$, critical point), and (c) $\Delta V_D = 5$ ($\alpha^2 > 1$, bistable phase). Lateral and upper plots show the partial distributions for v and v_{in}, respectively. The arrows indicate the direction and magnitude of the flow $\boldsymbol{u}(\boldsymbol{v}) = \sum_\rho \boldsymbol{\Delta}_\rho \, \omega_\rho(\boldsymbol{v}) P_{\text{ss}}(\boldsymbol{v})$ at each point in the state space.

II. GAUSSIAN APPROXIMATION OF THE STEADY STATE DISTRIBUTION AND COMPUTATION OF THE AVERAGE CURRENT

The steady state rate function satisfies:

$$0 = \sum_\rho \omega_\rho(\boldsymbol{v}) \left[1 - e^{\boldsymbol{\Delta}_\rho \cdot \nabla f_{\text{ss}}(\boldsymbol{v})} \right]. \tag{1}$$

We consider the second order expansion $f_{\text{ss}}(\boldsymbol{v}) = \boldsymbol{v}^T C^{-1} \boldsymbol{v}/2 + \mathcal{O}(|\boldsymbol{v}|^3)$ around $\boldsymbol{v} = 0$, the deterministic fixed point for $\alpha < 1$. Then, from the previous equation one can show that the matrix C must satisfy the Lyapunov equation

$$0 = CA + A^T C + B, \tag{2}$$

where the matrices A and B are given by $\{A\}_{ij} = \sum_\rho \partial_{v_i} \omega_\rho(0)(\boldsymbol{\Delta}_\rho)_j$ and $\{B\}_{ij} = \sum_\rho \omega_\rho(0)(\boldsymbol{\Delta}_\rho)_i(\boldsymbol{\Delta}_\rho)_j$ in terms of the transition rates. Eq. (2) can be solved explicitly for $n = 1$, obtaining

$$C_{11} = \frac{1 + e^{\Delta V_D/2} + e^{\Delta V_S} - e^{(\Delta V_D + \Delta V_S)/2}}{2(e^{\Delta V_D/2} + e^{\Delta V_S/2} - e^{(\Delta V_D + \Delta V_S)/2})}$$

$$C_{22} = \frac{1 + e^{\Delta V_D} + e^{\Delta V_S/2} - e^{(\Delta V_D + \Delta V_S)/2}}{2(e^{\Delta V_D/2} + e^{\Delta V_S/2} - e^{(\Delta V_D + \Delta V_S)/2})} \tag{3}$$

$$C_{12} = C_{21} = \frac{e^{(\Delta V_D + \Delta V_S)/2} - 1}{2(e^{(\Delta V_D + \Delta V_S)/2} - e^{\Delta V_S/2} - e^{\Delta V_D/2})}.$$

The large deviations principle $P_{\text{ss}}(\boldsymbol{v}) \asymp e^{-v_e^{-1} f_{\text{ss}}(\boldsymbol{v})}$ implies that for small v_e typical fluctuations are Gaussian with covariance matrix $\langle \boldsymbol{v}\boldsymbol{v}^T \rangle = v_e C$. Then, to first non-trivial order in $v_e \to 0$, the average steady state current through the first inverter can be computed as

$$\langle I_S \rangle \simeq \iint d^2\boldsymbol{v} \, \frac{e^{-v_e^{-1} f_{\text{ss}}(\boldsymbol{v})}}{Z_{\text{ss}}} \underbrace{q_e \left(\lambda_+^{\text{p/n}}(\boldsymbol{v}) - \lambda_-^{\text{p/n}}(\boldsymbol{v}) \right)}_{I(\boldsymbol{v})}$$

$$\simeq I(0) + \text{Tr}[\boldsymbol{H}_I(0)\langle \boldsymbol{v}\boldsymbol{v}^T \rangle]/2,$$

(4)

where $Z_{\text{ss}} = \int dv \exp(-v_e^{-1} f_{\text{ss}}(\boldsymbol{v}))$, $\lambda_{\pm}^p(\boldsymbol{v})$ are the Poisson rates of the (p/n)MOS transistor, and $\boldsymbol{H}_I(\boldsymbol{v})$ is the Hessian of $I(\boldsymbol{v})$. In this way we can obtain the following expression for $\langle I_S \rangle$:

$$\langle I_S \rangle = \frac{q_e}{\tau_0} \left(e^{\Delta V_S/2} - 1 \right) + \frac{q_e}{\tau_0} \frac{v_e \left(2e^{(\Delta V_D + \Delta V_S)/2} - e^{\Delta V_D + \Delta V_S/2} + e^{\Delta V_D/2 + \Delta V_S} - e^{\Delta V_D/2} + e^{\Delta V_D} - 2e^{\Delta V_S/2} \right)}{\left(e^{\Delta V_D/2} - 1 \right) \left(4e^{\Delta V_S/2} \right) - 4e^{\Delta V_D/2}} + O\left(v_e^2 \right),$$

(5)

which to lower order in ΔV_S reduces to the one given in the main text.

III. ESTIMATIONS FOR MODERN FABRICATION PROCESSES

In modern fabrication processes the gate capacitance can be as low as $C = 50$ aF [1, 2]. This corresponds to an elementary voltage $v_e = q/C \simeq 3.2$ mV. For room temperature, the thermal voltage is $V_T = k_b T/q_e \simeq 26$ mV. Thus, for those parameters the scale parameter is $\Omega = V_T/v_e \simeq 8$ (this motivates the value $\Omega = 10$, or equivalently $v_e = 0.1 V_T$, used for the numerical results in the main text). Also, for the same fabrication processes [1, 2], the factor $I_0 e^{-V_{\text{th}}/n}$ appearing in the I-V curve and Poisson rates of our model is in the order of $1 - 10$ nA. According to the results in the main text, this is the same order of magnitude of the rectified current $\langle I_S \rangle \leq 0$.

[1] Peng Zheng. Advanced MOSFET Structures and Processes for Sub-7 nm CMOS Technologies, 2016. [Online; accessed 13. Jul. 2022].

[2] Etienne Sicard and Lionel Trojman. Introducing 5-nm FinFET technology in Microwind, June 2021. [Online; accessed 13. Jul. 2022].