

Information flows in macroscopic Maxwell's demonsNahuel Freitas  and Massimiliano Esposito *Complex Systems and Statistical Mechanics, Department of Physics and Materials Science, University of Luxembourg, L-1511 Luxembourg, Luxembourg*

(Received 19 October 2022; accepted 23 December 2022; published 26 January 2023)

A CMOS-based implementation of an autonomous Maxwell's demon was recently proposed [[Phys. Rev. Lett. **129**, 120602 \(2022\)](#)] to demonstrate that a Maxwell demon can still work at macroscopic scales, provided that its power supply is scaled appropriately. Here we first provide a full analytical characterization of the nonautonomous version of that model. We then study system-demon information flows within generic autonomous bipartite setups displaying a macroscopic limit. By doing so, we can study the thermodynamic efficiency of both the measurement and the feedback process performed by the demon. We find that the information flow is an intensive quantity and that, as a consequence, any Maxwell's demon is bound to stop working above a finite scale if all parameters but the scale are fixed. However, this can be prevented by appropriately scaling the thermodynamic forces. These general results are applied to the autonomous CMOS-based demon.

DOI: [10.1103/PhysRevE.107.014136](https://doi.org/10.1103/PhysRevE.107.014136)**I. INTRODUCTION**

A Maxwell's demon is an agent or mechanism able to make a system behave in a way that seems to contradict the second law of thermodynamics [1]. It achieves that by gaining information about the system via measurements, based on which it performs a feedback on the system. In an ideal situation, the feedback control must not provide any energy to the system [2]. Maxwell's demon models come in two flavors: nonautonomous ones where the feedback is modeled as a parametric time dependence on the system's dynamics, and autonomous ones, where the mechanism of the demon coupled to the system is explicitly modeled and the system-demon dynamics is homogeneous in time.

Many implementations of these ideas have been provided over recent years [3–17]. An electronic implementation of an autonomous Maxwell's demon fully based on CMOS technology (the same used to build regular computers) was recently proposed in Ref. [18]. Its distinctive feature is that it can be made to work at different physical scales. This scale, Ω , is controlled by the size of the MOS transistors involved in the proposed circuit. It was shown that when all the other parameters are fixed, the rectification effects associated with the action of the demon disappear above a finite scale Ω^* . However, when the powering available to the demon to perform its functions is appropriately scaled with Ω , the demon can in principle continue to work at any scale. In that case, the price to pay is a decreasing thermodynamic efficiency, that was shown to scale as $\eta \propto 1/\Omega^2 \log(\Omega)$.

In Sec. II we start by studying a nonautonomous version of the CMOS demon proposed in Ref. [18]. This enables an exact analytical treatment of the rectification and captures the essential features of the demon mechanism. In Sec. III we revisit (in more mathematical detail) the autonomous implementation proposed in Ref. [18] and compare the results with

those of the nonautonomous version. In Sec. IV we turn to the most important part of the paper: The general analysis of autonomous Maxwell demons displaying a macroscopic limit in terms of the information thermodynamics theory developed in Refs. [19,20]. This allows us to study the information flows between the demon and the system and to define independent thermodynamic efficiencies for the measurement and feedback processes performed by the demon. We derive general scaling laws for the information flows and the thermodynamic efficiencies. In particular, we find that if all extra parameters are fixed, then the information flow is scale invariant for any autonomous bipartite system with a macroscopic limit and, as a consequence, any autonomous Maxwell's demon will stop operating above a finite scale. Scaling power is thus essential to produce macroscopic Maxwell demons.

II. NONAUTONOMOUS CMOS-DEMON

In this section we show how a simple feedback control protocol in the CMOS inverter can make the electrical current through it flow against the powering bias. As shown in Fig. 1, a CMOS inverter is composed of an nMOS transistor and a pMOS transistor. We will consider the subthreshold operation of these devices [21]. In a three-terminal nMOS transistor, electrical conduction between gate and source terminals is enhanced when the gate-source voltage is positive and reduced when that voltage is negative. This relation is reversed for pMOS transistors: conduction is enhanced (reduced) for negative (positive) gate-source voltages. As a consequence, when a voltage bias is applied by connecting the drain terminals to voltage sources $V_1 = -V_2 = \Delta V_S/2 > 0$, the input-output transfer function of the inverter has the typical shape shown in Fig. 1(b). 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

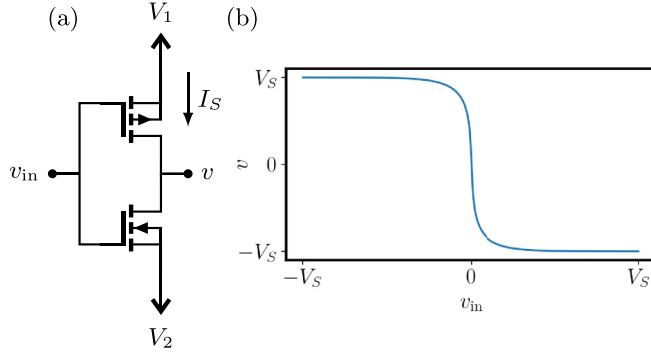


FIG. 1. (a) Common implementation of a NOT gate with CMOS technology. (b) Deterministic output voltage as a function of the input (for $V_1 = -V_2 = \Delta V_S/2 > 0$)

voltages, and the output voltage approaches V_1 . Let us first consider the situation where no bias is applied: $V_1 = V_2 = 0$. In that case the circuit will attain thermal equilibrium, and the steady-state fluctuations in the output voltage v will be given by the Gibbs distribution $P_{\text{eq}}(v) \propto e^{-\beta U(v)}$, where $\beta = 1/k_b T$ is the inverse temperature, and $U(v) = v^2/2C$ is the electrostatic energy for a given value of voltage (C is the output capacitance of the inverter). In the following, all voltages will be expressed in units of the thermal voltage $V_T = k_b T/q_e$, where q_e is the positive electron charge. Let us now consider the following feedback protocol: At time t the output voltage v is measured and a voltage $v_{\text{in}} = -\alpha v$ is applied to the input, with $\alpha > 0$. Thus, if a positive fluctuation $v > 0$ is observed at time t , then 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 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, in absence of a voltage bias. A net upward electric current could also be observed even if a small downward bias $V_1 - V_2 = \Delta V_S > 0$ is applied. This is indeed confirmed by numerical results shown in Fig. 2. We see that as we increase the biasing voltage V_S , a higher amplification factor α is required by the feedback protocol in order to reverse the direction of the current. When that happens, the entropy production rate $\dot{\Sigma} = \Delta V_S \langle I_S \rangle / T$ is actually negative ($\langle I_S \rangle$ is the average steady-state electric current through the inverter, and therefore $\Delta V_S \langle I_S \rangle$ is the rate of heat dissipation). According to the second law of thermodynamics, and the modern understanding of Maxwell's demons and information engines, that negative entropy production must be compensated by a positive entropy production in the system implementing the feedback control protocol. This was studied in the autonomous implementation proposed in Ref. [18], which is detailed in the next section. Before that, we review the deterministic description of the CMOS inverter in subthreshold operation and also its stochastic counterpart leading to the exact results in Fig. 2.

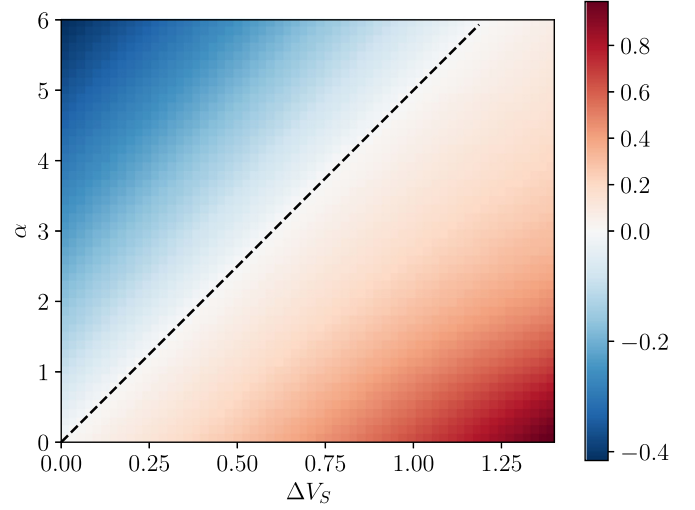


FIG. 2. Average electric current $\langle I_S \rangle$ at steady state for the feedback-controlled CMOS inverter as a function of the voltage bias ΔV_S and the feedback parameter α ($v_e = 0.1$, $n = 1$). Negative values (blue) correspond to a current flowing against the voltage bias or, equivalently, in the upward direction. The dashed line marks the boundary between the two phases and is given by Eq. (19).

A. Deterministic dynamics

The deterministic circuit analysis of the inverter in Fig. 1(a) leads to the following dynamical equation for the output voltage $v(t)$:

$$C d_t v(t) = I_p(v(t), v_{\text{in}}; \Delta V_S) - I_n(v(t), v_{\text{in}}; \Delta V_S), \quad (1)$$

where $I_{n/p}(v, v_{\text{in}})$ is the current through the n/p transistor for given values of v and v_{in} and C is the output capacitance of the inverter (or the effective capacitance including the one of the next stage to which the inverter is connected to). In the subthreshold regime of operation, the currents are given by:

$$I_p(v, v_{\text{in}}; \Delta V_S) = I_0 e^{(\Delta V_S/2 - v_{\text{in}} - v)/n} [1 - e^{-(\Delta V_S/2 - v)}] \quad (2)$$

and $I_n(v, v_{\text{in}}; \Delta V) = I_p(-v, -v_{\text{in}}; \Delta V)$. In the previous expression, V_{th} (the threshold voltage) and $n \geq 1$ (the slope factor) are phenomenological parameters characterizing the transistors. The deterministic output voltage v^* for a given v_{in} are found by solving $d_t v(t) = 0$ or, equivalently, $I_p(v^*, v_{\text{in}}) = I_n(v^*, v_{\text{in}})$. An explicit expression for $v^*(v_{\text{in}})$ can be obtained from (for $n = 1$):

$$e^{v^* - v_{\text{in}}} = \sqrt{1 + e^{\Delta V_S} [\cosh(v_{\text{in}}) - 1]} / 2 - e^{\Delta V_S/2} \sinh(v_{\text{in}}). \quad (3)$$

In general, we always have $V_1 > v^* > V_2$ and $v^* \rightarrow 0$ for $v_{\text{in}} \rightarrow 0$. From the first property it follows that both currents are positive for any value of v_{in} . Then, no inversion of the current is possible at the deterministic level. The second property justifies performing a series expansion in both v and v_{in} for $v_{\text{in}} \ll 1$. Doing that, it is easy to see that the gain for small inputs is

$$\alpha_S = \left. \frac{\partial v^*}{\partial v_{\text{in}}} \right|_{v_{\text{in}}=0} = (1 - e^{\Delta V_S/2})/n. \quad (4)$$

Also, we note that v^* is independent of the parameters I_0 and V_{th} , since they only fix the global timescale,

$$\tau_0 = (I_0 e^{-V_{th}/n} / q_e)^{-1}, \quad (5)$$

and do not affect the steady-state solution.

The deterministic output voltage v^* is always unique for a given fixed value of v_{in} . This is not anymore the case for the feedback protocol mentioned above. We assume that the repeated application of the protocol (that is, the measurement of the output voltage v and the subsequent adjustment of the input voltage as $v_{in} = -\alpha v$) is sufficiently fast compared to the timescale τ_0 . Then the steady-state output voltage can be obtained by solving $I_p(v^*, -\alpha v^*; \Delta V_S) = I_n(v^*, -\alpha v^*; \Delta V_S)$. This equation has $v^* = 0$ as the single solution for $\alpha \leq \alpha_c \equiv n / (e^{\Delta V_S/2} - 1)$. For $\alpha > \alpha_c$, two additional symmetric solutions v_{\pm}^* appear and the solution $v^* = 0$ becomes unstable. Thus, under the feedback control protocol, the system becomes bistable when the amplification factor is above the critical value α_c . Note that the condition $\alpha > \alpha_c$ can be interpreted in the following way: The product of the inverter gain $|\alpha_S|$ in Eq. (4) and the amplification factor α of the feedback protocol must be greater than 1. The electric current is also positive in any of the solutions of the bistable phase, so rectification is still not possible at the deterministic level.

B. Stochastic dynamics

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 [22]. 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 [23]. The rates that must be assigned to the two Poisson processes can be computed from the I - V curve characterization of the device and the requirement of thermodynamic consistency, as explained in Ref. [23]. Then the stochastic dynamics of a given circuit can be mapped to a Markov jump process in the space of possible circuit states, and the probability distribution over those states evolves according to a Markovian master equation. For the CMOS inverter, the master equation reads:

$$d_t P(v, t) = \sum_{\rho} \lambda_{\rho}(v - \Delta_{\rho} v_e) P(v - \Delta_{\rho} v_e, t) - \lambda_{\rho}(v) P(v, t), \quad (6)$$

where v is the output voltage (that can only take discrete values spaced by the elementary voltage $v_e = q_e / C$), and the index ρ runs over all the possible jump processes, with Poisson rates $\lambda_{\pm}^{n/p}(v, v_{in})$. For example, the rates $\lambda_{\pm}^p(v, v_{in})$ correspond to conduction events through the pMOS transistor, for which $v \rightarrow v \pm v_e$, while the rates $\lambda_{\pm}^n(v, v_{in})$ are associated conduction events through the nMOS transistor, where $v \rightarrow v \mp v_e$. The numbers Δ_{ρ} indicate the change in v/v_e for each jump, and therefore we have $\Delta_{\pm}^p = \pm 1$ and $\Delta_{\pm}^n = \mp 1$.

In order to ensure thermodynamic consistency, the Poisson rates must satisfy the local detailed balance (LDB) condition,

$$\log \frac{\lambda_{\rho}(v)}{\lambda_{-\rho}(v + \Delta_{\rho} v_e)} = \sigma_{\rho} / k_b, \quad (7)$$

according to which the log-ratio of forward and backward rates corresponding to a given transition equals the entropy production σ_{ρ} during that transition. For this kind of electronic circuits, the entropy production can be computed as $\sigma_{\rho} = -Q_{\rho}/T$, where $Q_{\rho} = U(v + \Delta_{\rho} v_e) - U(v) + W_{\rho}$ is the heat provided by the thermal environment during the transition [$U(v)$ is the electrostatic energy and W_{ρ} the nonconservative work performed by the voltage sources during the transition]. The Poisson rates $\lambda_{\pm}^{n/p}(v, v_{in})$ can be determined from the I - V curve characterization of the transistors in Eq. (2) and the LDB condition in Eq. (7). For the pMOS transistor, they read:

$$\begin{aligned} \lambda_{+}^p(v, v_{in}) &= \tau_0^{-1} e^{(\Delta V_S/2 - v_{in})/n} \\ \lambda_{-}^p(v, v_{in}) &= \lambda_{+}^p(v, v_{in}) e^{-(\Delta V_S/2 - v)} e^{-v_e/2}, \end{aligned} \quad (8)$$

while for the nMOS transistor we have $\lambda_{\pm}^n(v, v_{in}) = \lambda_{\pm}^p(-v, -v_{in})$.

We are interested in obtaining the steady-state solution of Eq. (6). Since this is a one-dimensional problem, the steady-state distribution is fully characterized by the condition that the total rate of transitions $v \rightarrow v + v_e$ must balance the total rate of the reverse transitions $v + v_e \rightarrow v$. This condition can be expressed as the recurrence relation

$$P_{ss}(v) = \frac{\lambda_{+}^p(v - v_e, v_{in}) + \lambda_{-}^n(v - v_e, v_{in})}{\lambda_{-}^p(v, v_{in}) + \lambda_{+}^n(v, v_{in})} P_{ss}(v - v_e), \quad (9)$$

that can be solved to find $P_{ss}(v)$. So far we have considered the input voltage v_{in} to be fixed. The steady state $P_{ss}^{fb}(v)$ attained under the action of the idealized feedback protocol described above can be found by just replacing $v_{in} \rightarrow -\alpha v$ in the previous expressions.

Once the steady state is determined, the average values of the net electric current through both transistors can be computed as

$$\langle I_{n/p} \rangle = q_e \sum_v P_{ss}^{fb}(v) [\lambda_{+}^{n/p}(v, -\alpha v) - \lambda_{-}^{n/p}(v, -\alpha v)]. \quad (10)$$

Since the two transistors are connected in series, at steady state we have $\langle I_n \rangle = \langle I_p \rangle \equiv \langle I_S \rangle$. This is how the results in Fig. 2 were obtained.

C. Macroscopic limit

We will now study what is the stochastic behavior of the CMOS inverter at different scales. For this, we will analyze how the parameters entering the problem are modified 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 [24]. For fixed L , both the parameter I_0 appearing in Eq. (2) (that enter the Poisson rates through the timescale τ_0) and the output capacitance C are proportional to W . Thus, the Poisson rates $\lambda_{\pm}^{n/p}(v)$ increase as W , while the elementary voltage $v_e = q_e / C$ decreases as W^{-1} . From these two observations it follows that the probability distribution $P(v, t)$

satisfies a large deviations (LD) principle in the limit $W \rightarrow \infty$ [25,26]. This means that fluctuations away from the deterministic behavior are exponentially suppressed in W . Taking $\Omega \equiv v_e^{-1} \propto W$ as a scale parameter, the LD property is expressed mathematically as the existence of the limit $f(v, t) = \lim_{v_e \rightarrow 0} -v_e \log[P(v, t)]$ or, equivalently,

$$P(v, t) \underset{v_e \rightarrow 0}{\simeq} e^{-[f(v, t) + o(v_e)]/v_e}, \quad (11)$$

where $f(v, t)$ is the rate function (it gives the rate at which the probability of fluctuation v decreases with v_e^{-1}). Indeed, plugging the previous ansatz in the master equation (6) and keeping only the dominant terms in $v_e \rightarrow 0$, we find that the rate function must satisfy the evolution equation:

$$d_t f(v, t) = \sum_{\rho} \omega_{\rho}(v) [1 - e^{\Delta_{\rho} d_v f(v, t)}], \quad (12)$$

where $\omega_{\rho}(v) = \lim_{v_e \rightarrow 0} v_e \lambda_{\rho}(v)$ are the scaled Poisson rates. The minimum $v(t)$ of the rate function $f(\cdot, t)$ at time t is the most probable value, and it can be seen that it evolves according to the closed deterministic dynamics of Eq. (1). In this way we see how the deterministic dynamics emerges from the stochastic one.

From Eq. (12), the steady-state rate function $f_{ss}(v)$ satisfies:

$$e^{-d_v f_{ss}(v)} = \frac{\omega_{+}^p(v, v_{in}) + \omega_{-}^n(v, v_{in})}{\omega_{+}^p(v, v_{in}) + \omega_{+}^n(v, v_{in})}. \quad (13)$$

Note that this equation can also be obtained as the $v_e \rightarrow 0$ limit of the recursive relation in Eq. (9). Again, the steady-state rate function $f_{ss}^{\text{fb}}(v)$ corresponding to the application of the feedback control protocol satisfies a similar equation where v_{in} is replaced by $-\alpha v$. Solving for $d_v f_{ss}^{\text{fb}}(v)$ and integrating in v , it is possible to obtain the following explicit expression (up to an arbitrary constant):

$$f_{ss}^{\text{fb}}(v) = \frac{v^2 + v \Delta V_S}{2} + \frac{n}{n + 2\alpha} [L(v, \Delta V_S) - L(v, \Delta V_S)], \quad (14)$$

where $L(v, \Delta V) = \text{Li}_2[-\exp[\Delta V/2 + v(1 + 2\alpha/n)]]$ and $\text{Li}_2(\cdot)$ is the polylogarithm function of second order.

We can use the previous result to compute the average currents according to Eq. (10), which for $v_e \ll 1$ can be approximated as:

$$\langle I_S \rangle \simeq \int dv \frac{e^{-v_e^{-1} f_{ss}^{\text{fb}}(v)}}{Z_{ss}} \underbrace{q_e [\lambda_{+}^p(v, -\alpha v) - \lambda_{-}^p(v, -\alpha v)]}_{I(v)}, \quad (15)$$

where $Z_{ss} = \int dv \exp[-v_e^{-1} f_{ss}^{\text{fb}}(v)]$. For $v_e \ll 1$ only small fluctuations around the minimum of the rate function will contribute to the integral, and for $\alpha < \alpha_c$ that minimum is $v = 0$. Expanding the function $I(v)$ defined above around $v = 0$, we obtain (using the fact that the distribution of v is symmetric around $v = 0$):

$$\langle I_S \rangle \simeq I(0) + I''(0) \langle v^2 \rangle / 2. \quad (16)$$

Also, a Gaussian approximation to the steady state for $\alpha < \alpha_c$ can be obtained by expanding the rate function to second order

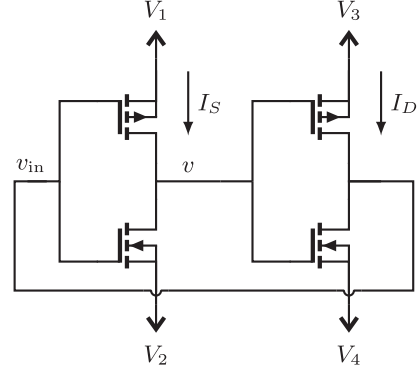


FIG. 3. CMOS implementation of an autonomous Maxwell's demon. If we break the left-right symmetry of the circuit by choosing powering biases $V_3 - V_4 > V_1 - V_2$, then we can consider that the right inverter monitors the thermal fluctuations in the output of the left inverter and, by appropriately acting back on its input, can make the electric current I_S to flow against the bias $V_1 - V_2$. For simplicity, all transistors are assumed to have identical parameters.

in v . That expansion reads (again up to a constant):

$$f_{ss}^{\text{fb}}(v) = \frac{1 - \alpha/\alpha_c}{1 + e^{V_S}} v^2 + \mathcal{O}(v^3). \quad (17)$$

Therefore, the steady-state voltage variance is

$$\langle v^2 \rangle \simeq \frac{v_e (e^{\Delta V_S/2} + 1)}{2(1 - \alpha/\alpha_c)}. \quad (18)$$

From Eqs. (18) and (16) and the definition of $I(v)$ in Eq. (15) one can obtain an exact expression for the steady-state current, valid to first order in v_e , which is, however, not very simple. To lower order in ΔV_S it reduces to (neglecting terms proportional to $\Delta V_S v_e$)

$$\langle I_S \rangle \simeq \frac{q_e}{\tau_0} (\Delta V_S/2 - \alpha v_e/n). \quad (19)$$

Thus, the minimum value of the amplification factor needed for the average current to flow against the applied bias is $\alpha_m = n \Delta V_S / 2 v_e$. We note that $\alpha_m \rightarrow \infty$ in the macroscopic limit $v_e \rightarrow 0$. Thus, to reverse the current is increasingly hard as one goes deeper in the macro limit. This is natural, since the rectification strategy exploits the fluctuations in the output voltage, and these fluctuations are negligible in the macro limit [their variance scales as v_e , see Eq. (18)]. As we will see in the next section, the demon or agent implementing the feedback protocol must invest an increasing amount of energy in order to achieve higher amplification factors.

III. AUTONOMOUS CMOS DEMON

A feedback protocol similar to the nonautonomous one can be implemented using an additional inverter (see Fig. 3), which plays the role of a Maxwell's demon monitoring the state of a system (the output of the first inverter) and performing some feedback on it (by changing the input voltage). The second inverter is powered by voltages $V_3 = -V_4 = \Delta V_D/2$. Thus, the bias ΔV_D controls the power available to the demon. We consider for simplicity that the two transistors of the first inverter are equivalent to the two transistors of the second

inverter. Thus, the only asymmetry between system and demon is introduced by the different biasing voltages ΔV_S and ΔV_D . As we explain below, it is possible to compute the steady-state distribution of the two degrees of freedom of the circuit, v and v_{in} , from which average currents can be obtained. In that way we obtain the results of Fig. 3. We see that if the bias ΔV_D applied to the demon is sufficiently high, then its action is able to reverse the direction of the electric current through the system. Before further analyzing the performance of this autonomous implementation of an electronic Maxwell's demon, we explain how to model the dynamics of the circuit in Fig. 3, again both at the deterministic and stochastic levels. It is important to note that this autonomous model automatically takes into account some realistic effects that were not considered in the previous nonautonomous implementation, namely the intrinsic noise affecting the measurement and feedback processes and the delay between them.

A. Deterministic dynamics

The deterministic equations of motion for the voltages v and v_{in} are

$$\begin{aligned} Cd_t v &= I_p(v, v_{\text{in}}; \Delta V_S) - I_n(v, v_{\text{in}}; \Delta V_S) \\ Cd_t v_{\text{in}} &= I_p(v_{\text{in}}, v; \Delta V_D) - I_n(v_{\text{in}}, v; \Delta V_D). \end{aligned} \quad (20)$$

Close to thermal equilibrium (that is, for low biases ΔV_S and ΔV_D), the previous deterministic dynamics has the unique fixed point $v = v_{\text{in}} = 0$. However, the system becomes bistable when $\alpha_S \alpha_D > 1$, where $\alpha_{S/D} \equiv (1 - e^{\Delta V_{S/D}/2})/n$ is the gain of each inverter, given by Eq. (4). The bistability can be exploited to store the value of a bit, and in fact for symmetric powering $\Delta V_S = \Delta V_D = \Delta V$ the circuit in Fig. 3 is the typical CMOS implementation of SRAM memory cells [2]. In that case, also assuming $n = 1$, the bistability is achieved for $\Delta V > 2 \log(2)$.

B. Stochastic dynamics

The master equation describing the stochastic evolution of the circuit in Fig. 3 is constructed in the same way as for the simpler implementation of the previous section. It reads

$$\begin{aligned} d_t P(\mathbf{v}, t) &= \sum_{\rho} \lambda_{\rho}(\mathbf{v} - \mathbf{\Delta}_{\rho} v_e) P(\mathbf{v} - \mathbf{\Delta}_{\rho} v_e, t) \\ &\quad - \lambda_{\rho}(\mathbf{v}) P(\mathbf{v}, t), \end{aligned} \quad (21)$$

where now $\mathbf{v} = (v, v_{\text{in}})$ is a vector with the two degrees of freedom of the circuit, and the vectors $\mathbf{\Delta}_{\rho}$ encode the change in \mathbf{v} for each jump. For example, for forward transitions through the pMOS and nMOS transistors of the first inverter, we have $\mathbf{\Delta}_{\rho} = (1, 0)$ and $\mathbf{\Delta}_{\rho} = (-1, 0)$, respectively. The transition rates $\lambda(\mathbf{v})$ are obtained in the same way as before [see Eq. (8)] and satisfy the LDB conditions ensuring thermodynamic consistency.

In this case, the determination of the steady state $P_{\text{ss}}(\mathbf{v})$ is not as straightforward as before. Since the voltages v and v_{in} take only discrete values spaced by the elementary voltage v_e , the steady state can be found numerically by truncating the infinite state space to a finite space with $\max(|v/v_e|, |v_{\text{in}}/v_e|) \leq N$, for sufficiently large N . Then the master equation can be

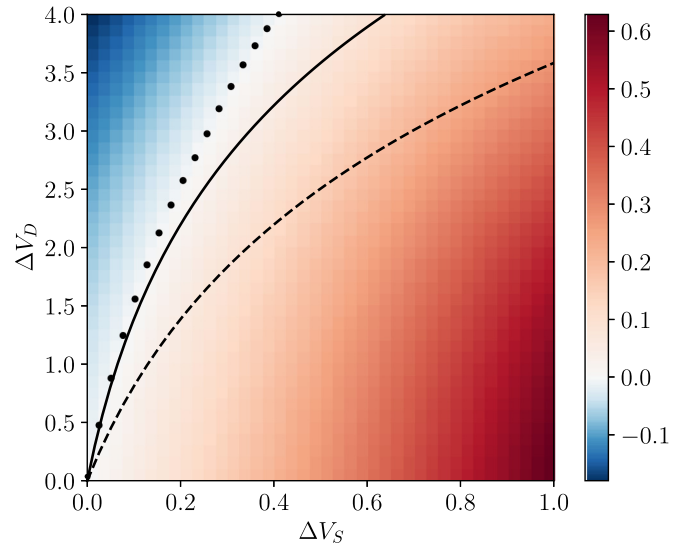


FIG. 4. Average electric current $\langle I_S \rangle$ at steady state for the autonomous circuit of Fig. 3 as a function of the voltage biases ΔV_S and ΔV_D ($v_e = 0.1$, $n = 1$). Negative values (blue) correspond to a current flowing against the voltage bias of the first inverter. The dotted line indicates the actual boundary between positive and negative currents. The solid black line shows the approximation given by Eq. (32), while the dashed black line shows the estimation of Eq. (22) based on the nonautonomous model.

written as $d_t |P(t)\rangle = W |P(t)\rangle$, where the probability vector $|P(t)\rangle$ has $(2N + 1)^2$ components and the generator W is a matrix of dimensions $(2N + 1)^2 \times (2N + 1)^2$. The steady-state vector $|P_{\text{ss}}\rangle$ is given by the right eigenvector of W with zero eigenvalue. Since the matrix W is sparse, that computation can be done efficiently. Typical steady-state distributions for different values of α^2 are shown in Fig. 5.

The results of the above procedure are shown in Fig. 4. We see that if V_D/V_S is sufficiently high, then the current through the first inverter is effectively reversed. The boundary between positive and negative values of that current can be estimated in the following way. In the nonautonomous implementation, a minimum amplification factor $\alpha_m = nV_s/q_e$ was required for the inversion of the current. From Eq. (4), we know that the second inverter can provide that minimum amplification factor if the biasing voltage V_D is above

$$\Delta V_D^m = 2 \log(1 + n^2 \Delta V_S / 2v_e). \quad (22)$$

The previous result is plotted with a dashed line in Fig. 4, where we see that it underestimates the minimum value of V_D required for inversion. The reason is that the previous estimation does not consider the intrinsic noise and delay in the demon side. An improved estimation is obtained below.

C. Macroscopic limit

The solution to Eq. (21) also satisfies an LD principle,

$$P(\mathbf{v}, t) \underset{v_e \rightarrow 0}{\asymp} e^{-[f(\mathbf{v}, t) + o(v_e)]/v_e}, \quad (23)$$

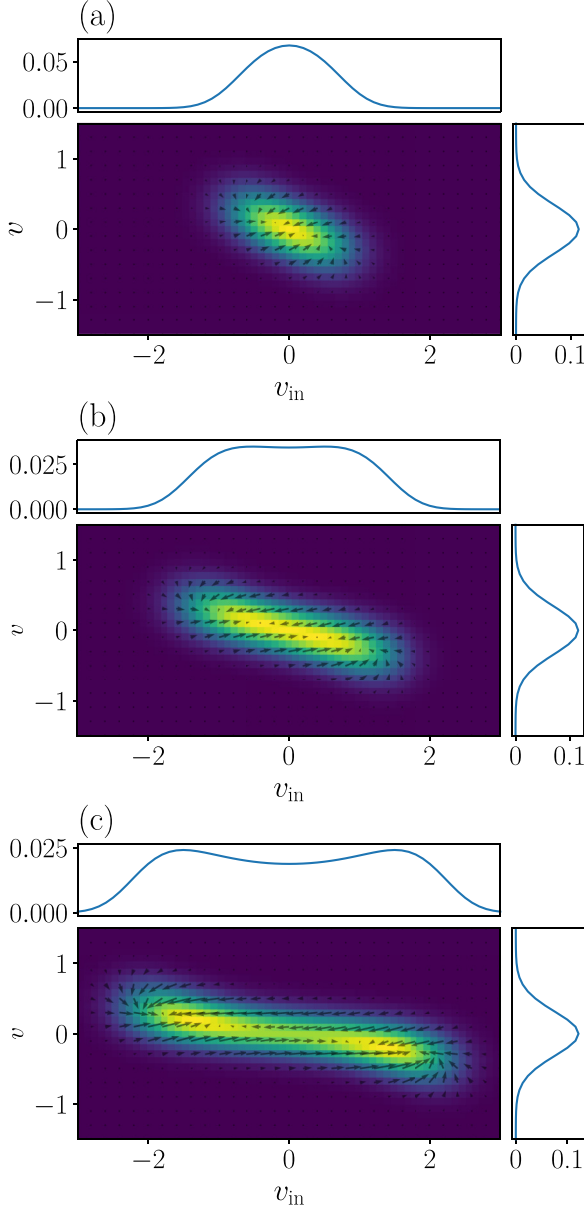


FIG. 5. Histograms of the steady-state distribution $P_{ss}(\mathbf{v})$ for $v_e = 0.1$, $\Delta V_S = 0.4$ and (a) $\Delta V_D = 2$ ($\alpha^2 < 1$, monostable phase), (b) $\Delta V_D \simeq 3.415$ ($\alpha^2 = 1$, critical point), and (c) $\Delta V_D = 5$ ($\alpha^2 > 1$, bistable phase). Lateral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

where the rate function $f(\mathbf{v}, t)$ evolves according to

$$d_t f(\mathbf{v}, t) = \sum_{\rho} \omega_{\rho}(\mathbf{v}) [1 - e^{\Delta_{\rho} \cdot \nabla f(\mathbf{v}, t)}], \quad (24)$$

but in this case the steady-state rate function $f_{ss}(\mathbf{v})$ cannot be computed exactly. One can still perform a Gaussian approximation. Thus, in the monostable phase where the minimum of the rate function is $\mathbf{v} = 0$, we have:

$$f_{ss}(\mathbf{v}) = \mathbf{v}^T \mathcal{C}^{-1} \mathbf{v} / 2 + \mathcal{O}(|\mathbf{v}|^3), \quad (25)$$

where the matrix \mathcal{C} is the solution of

$$0 = \mathcal{C} \mathcal{A} + \mathcal{A}^T \mathcal{C} + \mathcal{B}, \quad (26)$$

and the matrices \mathcal{A} and \mathcal{B} are given by

$$\{\mathcal{A}\}_{ij} = \sum_{\rho} \partial_{v_i} \omega_{\rho}(0) (\Delta_{\rho})_j \quad (27)$$

and

$$\{\mathcal{B}\}_{ij} = \sum_{\rho} \omega_{\rho}(0) (\Delta_{\rho})_i (\Delta_{\rho})_j. \quad (28)$$

Solving Eq. (26) one finds the following correlators:

$$\begin{aligned} \mathcal{C}_{11} &= \frac{1 + e^{\Delta V_D/2} + e^{\Delta V_S} - e^{(\Delta V_D + \Delta V_S)/2}}{2[e^{\Delta V_D/2} + e^{\Delta V_S/2} - e^{(\Delta V_D + \Delta V_S)/2}]} \\ \mathcal{C}_{22} &= \frac{1 + e^{\Delta V_D} + e^{\Delta V_S/2} - e^{(\Delta V_D + \Delta V_S)/2}}{2[e^{\Delta V_D/2} + e^{\Delta V_S/2} - e^{(\Delta V_D + \Delta V_S)/2}]} \\ \mathcal{C}_{12} = \mathcal{C}_{21} &= \frac{e^{(\Delta V_D + \Delta V_S)/2} - 1}{2[e^{(\Delta V_D + \Delta V_S)/2} - e^{\Delta V_S/2} - e^{\Delta V_D/2}]} \end{aligned} \quad (29)$$

As in the previous section, the average steady-state current through the first inverter can be computed as

$$\begin{aligned} \langle I_S \rangle &\simeq \iint d^2 \mathbf{v} \frac{e^{-\mathbf{v}_e^{-1} f_{ss}^{\text{fb}}(\mathbf{v})}}{Z_{ss}} \underbrace{q_e [\lambda_+^p(\mathbf{v}) - \lambda_-^p(\mathbf{v})]}_{I(\mathbf{v})} \\ &\simeq I(0) + \text{Tr}[\mathbf{H}_I(0) \langle \mathbf{v} \mathbf{v}^T \rangle] / 2, \end{aligned} \quad (30)$$

where $\mathbf{H}_I(\mathbf{v})$ is the Hessian of $I(\mathbf{v})$ and $\langle \mathbf{v} \mathbf{v}^T \rangle = v_e \mathcal{C}$ is the covariance matrix of the Gaussian fluctuations around the fixed point (in the second line we have assumed that the distribution of \mathbf{v} is symmetric on the point reflection $\mathbf{v} \rightarrow -\mathbf{v}$, and therefore all odd moments vanish). We can then obtain an analytical expression for the average current $\langle I_S \rangle$, valid to lower nontrivial order in $v_e \rightarrow 0$, which is, however, too complicated to show here (see the Supplementary Material in Ref. [18]). To lower order in V_S it reduces to (neglecting terms proportional to $\Delta V_S v_e$)

$$\langle I_S \rangle = \frac{q_e}{2\tau_0} [\Delta V_S - v_e (e^{\Delta V_D/2} - 1)]. \quad (31)$$

From the previous result it follows that the minimum value of ΔV_D required for inversion is

$$\Delta V_D^m = 2 \log(1 + \Delta V_S / v_e), \quad (32)$$

which differs from Eq. (22) by a factor 2 accompanying ΔV_S . This is shown with a solid line in Fig. 4. Alternatively, from Eq. (31) we can say that inversion is impossible above a maximum scale,

$$v_e^{*-1} = |\alpha_D| / \Delta V_S. \quad (33)$$

Whenever the action of the demon manages to make the current I_S to flow against the voltage bias ΔV_S , heat is extracted from the environment of the first inverter and work is realized on the sources fixing the voltages V_1 and V_2 . In that case the entropy production rate at the system side is $\dot{\Sigma}_S = \langle I_S \rangle \Delta V_S < 0$. As we have seen, for that to happen energy needs to be consumed by the second inverter, and therefore the entropy production rate at the demon side is

$\dot{\Sigma}_D = \langle I_D \rangle \Delta V_D > 0$. We can then define the efficiency of the demon as the ratio:

$$\eta = -\frac{\dot{\Sigma}_S}{\dot{\Sigma}_D} = -\frac{\Delta V_S \langle I_S \rangle}{\Delta V_D \langle I_D \rangle} \leq 1. \quad (34)$$

The last inequality follows from the fact that the total entropy production rate $\dot{\Sigma} = \dot{\Sigma}_S + \dot{\Sigma}_D$ is always positive. The thermodynamic efficiency defined in Eq. (34) allows to characterize the performance of the demon at a global level and can be evaluated for different parameters using the above results. However, it is also possible to introduce detailed thermodynamic efficiencies characterizing the measurement and feedback processes independently, providing additional insight. Then, the net efficiency in Eq. (34) is recovered as the product of the detailed efficiencies. To see this, we review in the following the basic theory of information flows as developed in Refs. [19,20] and adapt it to Gaussian states.

IV. INFORMATION THERMODYNAMICS

The thermodynamics of measurement and feedback protocols in bipartite systems can be understood in terms of information flows which, as explained in Refs. [19,20], enter as additional terms in the entropy balance of each subsystem. In this section we compute those information flows for general bipartite systems displaying a macroscopic limit and then apply it to our autonomous CMOS demon. For this, in the sake of clarity, we will adopt a slightly different notation. We assume that we are dealing with a system with two degrees of freedom x and y (v and v_{in} in our example). The Poisson rates $\lambda_\rho(x|y)$ associated to jumps $x \rightarrow x + \delta x_\rho$ depend on the variable y , and in the same way the Poisson rates $\lambda_\rho(y|x)$ associated to jumps $y \rightarrow y + \delta y_\rho$ depend on the variable x . In our circuit, this corresponds to the fact that the Poisson rates associated to the transistors in the first inverter (which modify the voltage v) depend on the voltage v_{in} (this is the feedback process in which the demon controls the system), while the rates of the transistors in the second inverter (which modify v_{in}) depend on the voltage v (this is the measurement process in which the demon acquires information about the state of the system). The jump sizes δx_ρ and δy_ρ are $\pm v_e$ in our case, with the sign depending on the transition.

The mutual information between x and y is defined as

$$\mathcal{I} = \sum_{x,y} P_t(x,y) \log \frac{P_t(x,y)}{P_t(x)P_t(y)}, \quad (35)$$

where $P_t(x,y)$ is the time-dependent global probability distribution, while $P_t(x)$ and $P_t(y)$ are the corresponding reduced distributions for x and y . The time derivative of the mutual information accepts the following decomposition [19]:

$$d_t \mathcal{I} = \dot{\mathcal{I}}^x + \dot{\mathcal{I}}^y, \quad (36)$$

where

$$\begin{aligned} \dot{\mathcal{I}}^x &= \sum_{x,y} \sum_{\rho>0} j_\rho(x|y) \log \left[\frac{P_t(y|x + \delta x_\rho)}{P_t(y|x)} \right] \\ \dot{\mathcal{I}}^y &= \sum_{x,y} \sum_{\rho>0} j_\rho(y|x) \log \left[\frac{P_t(x|y + \delta y_\rho)}{P_t(x|y)} \right] \end{aligned} \quad (37)$$

are the contributions due to jumps $x \rightarrow x + \delta x_\rho$ and $y \rightarrow y + \delta y_\rho$, respectively. In the previous expressions, $j_\rho(x|y) \equiv \lambda_\rho(x|y)P_t(x,y) - \lambda_{-\rho}(x + \delta x_\rho|y)P_t(x + \delta x_\rho,y)$ is the net current along transition $x, y \rightarrow x + \delta x_\rho, y$, and $j_\rho(y|x)$ is defined similarly. Note that at steady state we have $d_t \mathcal{I} = 0$ and therefore $\dot{\mathcal{I}}^x = -\dot{\mathcal{I}}^y$.

We will now evaluate the steady-state information flows $\dot{\mathcal{I}}^{x/y}$ in the macroscopic limit. Under a Gaussian approximation like the one used previously, the log-ratio of conditional probabilities appearing in the equations above can be written as (for simplicity we assume $\langle x \rangle = \langle y \rangle = 0$):

$$\begin{aligned} \log \left[\frac{P_t(y|x')}{P_t(y|x)} \right] &= \frac{\langle xy \rangle}{\langle xx \rangle \langle yy \rangle - \langle xy \rangle^2} \\ &\times \left[y(x' - x) - \frac{\langle xy \rangle}{\langle xx \rangle} (x'^2 - x^2) / 2 \right], \end{aligned} \quad (38)$$

where the averages are computed for $P_t(x,y)$. It follows that the information flow $\dot{\mathcal{I}}^x$ can be rewritten as:

$$\dot{\mathcal{I}}^x = \frac{\langle xy \rangle}{\langle xx \rangle \langle yy \rangle - \langle xy \rangle^2} \left[\langle y \bar{d}_t x \rangle - \frac{\langle xy \rangle}{\langle xx \rangle} d_t \langle x^2 \rangle / 2 \right], \quad (39)$$

where

$$\begin{aligned} \langle y \bar{d}_t x \rangle &= \sum_{x,y} \sum_{\rho>0} j_\rho(x|y) y \delta x_\rho \\ d_t \langle x^2 \rangle &= \sum_{x,y} \sum_{\rho>0} j_\rho(x|y) (2x \delta x_\rho + \delta x_\rho^2). \end{aligned} \quad (40)$$

The bar in \bar{d}_t denotes that it is not a total derivative. The last term in Eq. (39) vanishes at steady state. Then, at the Gaussian level the steady-state information flows read:

$$\dot{\mathcal{I}}_{ss}^x = \frac{\langle xy \rangle \langle y \bar{d}_t x \rangle}{\langle xx \rangle \langle yy \rangle - \langle xy \rangle^2} \quad \text{and} \quad \dot{\mathcal{I}}_{ss}^y = \frac{\langle xy \rangle \langle x \bar{d}_t y \rangle}{\langle xx \rangle \langle yy \rangle - \langle xy \rangle^2}. \quad (41)$$

Note that $\dot{\mathcal{I}}_{ss}^x + \dot{\mathcal{I}}_{ss}^y \propto \langle x \bar{d}_t y \rangle + \langle y \bar{d}_t x \rangle = d_t \langle xy \rangle = 0$ in steady-state conditions, as discussed above. A calculation similar to the one in the Sec. III C shows that to first nontrivial order in the macroscopic limit:

$$\begin{aligned} \langle y \bar{d}_t x \rangle &= \langle xy \rangle \sum_{\rho>0} [\partial_x \lambda_\rho(x|y)|_{0,0} - \partial_x \lambda_{-\rho}(x|y)|_{0,0}] \delta x_\rho \\ &+ \langle yy \rangle \sum_{\rho>0} [\partial_y \lambda_\rho(x|y)|_{0,0} - \partial_y \lambda_{-\rho}(x|y)|_{0,0}] \delta x_\rho. \end{aligned} \quad (42)$$

Recalling that in the macroscopic limit we have the scalings $\lambda_\rho \propto \Omega$ and $\delta x_\rho, \langle xx \rangle, \langle yy \rangle, \langle xy \rangle \propto 1/\Omega$ with respect to a scale parameter Ω , we see that $\langle y \bar{d}_t x \rangle \propto 1/\Omega$. Therefore, the information flows $\dot{\mathcal{I}}_{ss}^{x/y}$ are scale independent.

For the kind of bipartite systems we are considering, it was shown in Ref. [19] that when the information flows are taken into account in the entropy balance of each subsystem, then the following extended local second laws hold (we assume isothermal settings at temperature T):

$$\begin{aligned} \dot{\Sigma}_x^i &= d_t S_x + \dot{\Sigma}_x - k_b \dot{\mathcal{I}}^x \geq 0 \\ \dot{\Sigma}_y^i &= d_t S_y + \dot{\Sigma}_y - k_b \dot{\mathcal{I}}^y \geq 0, \end{aligned} \quad (43)$$

where $\dot{\Sigma}_{x/y}^i$, $S_{x/y}$, $\dot{\Sigma}_{x/y}$ are, respectively, the total irreversible entropy production, the internal entropy, and the entropy flow into the environment of each subsystem. Since $S_{x/y}$ are the entropies associated to the reduced distributions $P_t(x)$ and $P_t(y)$, we see that by adding the two previous equations we recover the usual global second law $\dot{\Sigma}^i = \dot{\Sigma}_x^i + \dot{\Sigma}_y^i = d_t S + \dot{\Sigma}$, where $\dot{\Sigma} = \dot{\Sigma}_x + \dot{\Sigma}_y$ and $S = S_x + S_y - \mathcal{I}$ is the entropy of the full distribution $P_t(x, y)$.

A. Efficiency of information creation or consumption

We will now compute the information flows for the electronic Maxwell demon based on the previous result, taking $x = v$ (the output voltage of the first inverter) and $y = v_{in}$ (the output voltage of the second inverter). Thus, x represents the system S and y the demon D. Then, when evaluated at steady state, the inequalities in Eq. (43) reduce to

$$\begin{aligned}\dot{\Sigma}_S^i &= \dot{\Sigma}_S + k_b \dot{\mathcal{I}} \geq 0 \\ \dot{\Sigma}_D^i &= \dot{\Sigma}_D - k_b \dot{\mathcal{I}} \geq 0,\end{aligned}\quad (44)$$

$$\dot{\mathcal{I}} = \frac{2v_e}{\tau_0} \frac{[1 - e^{(\Delta V_S + \Delta V_D)/2}](e^{\Delta V_S/2} - e^{\Delta V_D/2})}{1 + e^{\Delta V_S/2} + e^{\Delta V_D/2} + e^{\Delta V_S} + e^{\Delta V_D} - e^{(\Delta V_S + \Delta V_D)/2}} + \mathcal{O}(v_e^2), \quad (46)$$

which is indeed scale invariant (recall that $\tau_0 \propto v_e \propto 1/\Omega$), as discussed above.

As already mentioned, the information flow $\dot{\mathcal{I}}$ is created by the dynamics of the demon, which also has an associated dissipation $\dot{\Sigma}_D$. Thus, one can define the thermodynamic efficiency for the creation of correlations as

$$\eta_D = \frac{k_b \dot{\mathcal{I}}}{\dot{\Sigma}_D}. \quad (47)$$

According to the second inequality in Eqs. (44), $\eta_D \leq 1$. Also, according to the first inequality in Eqs. (44), a positive information flow might be employed to compensate for a locally negative entropy production $\dot{\Sigma}_S \leq 0$. In those situations, the following thermodynamic efficiency can be assigned to that process:

$$\eta_S = -\frac{\dot{\Sigma}_S}{k_b \dot{\mathcal{I}}}, \quad (48)$$

and we also have $\eta_S \leq 1$. Note that the net thermodynamic efficiency defined in Eq. (34) is just $\eta = \eta_S \eta_D$.

B. Scaling laws of information flows and efficiencies

We now discuss how the different flows and efficiencies scale with respect to the scale parameter $\Omega = v_e^{-1}$. All other parameters fixed, we know that entropy flows $\dot{\Sigma}_{S/D}$ are extensive quantities [since the rates $\lambda_{\pm}^{n/p}(\mathbf{v})$ are extensive and therefore the electric currents are]. Also, we have seen that the information flow $\dot{\mathcal{I}}$ is intensive. This means that the efficiency η_D scales as $1/\Omega$. It also means that $\dot{\Sigma}_S$ must become positive above some value of Ω , since otherwise we would have $\eta_S > 1$, violating the first inequality in Eq. (44). As we have shown, this is indeed what happens in our electronic demon: Rectification is impossible above a certain scale [see Eq. (33)].

where $\dot{\mathcal{I}} = -\dot{\mathcal{I}}_{ss}^S = \dot{\mathcal{I}}_{ss}^D$ is the steady-state information flow. As we will see, $\dot{\mathcal{I}}$ is positive, which means that the dynamics of the demon produces correlations that are then consumed at the system side. This allows the entropy flow $\dot{\Sigma}_S$ to be negative while still respecting the first inequality in Eq. (44).

Evaluating Eq. (42) for our circuit we obtain:

$$\begin{aligned}\langle y d_t x \rangle &= \langle xy \rangle [\partial_x I_S^p|_{0,0} v_e + \partial_x I_S^n|_{0,0} (-v_e)] / q_e \\ &+ \langle yy \rangle [\partial_y I_S^p|_{0,0} v_e + \partial_y I_S^n|_{0,0} (-v_e)] / q_e \\ &= (2v_e/q_e) [\langle xy \rangle \partial_x I_S^p|_{0,0} + 2\langle yy \rangle \partial_y I_S^p|_{0,0}] \\ &= -\frac{2v_e}{\tau_0} [\langle xy \rangle + \langle yy \rangle (e^{\Delta V_S/2} - 1)],\end{aligned}\quad (45)$$

where $I_S^{n/p}(x, y) = q_e [\lambda_{+}^{n/p}(x, y) - \lambda_{-}^{n/p}(x, y)]$ is the electric current through the (p/n)MOS transistor in the first inverter, and in the second line we have used the fact that $\partial_{x/y} I_S^p|_{0,0} = -\partial_{x/y} I_S^n|_{0,0}$. Combining Eqs. (41) and (45) and the expression for the correlators in Eq. (29), we find that the information flow is given by:

It was shown in Ref. [18] that a way around the previous limitation is to make the voltage biases ΔV_S and ΔV_D to be scale dependent. In particular, if we take $\Delta V_S = cv_e \propto 1/\Omega$, and $\Delta V_D = 2 \log(1 + 2\alpha^2/cv_e) \propto \log(\Omega)$, then we obtain that $\dot{\Sigma}_S$ is scale independent and that $\dot{\Sigma}_D \propto \Omega^2 \log(\Omega)$. Thus, the demon continues to work for any scale parameter but with a net efficiency that decreases as $\eta = 1/\Omega^2 \log(\Omega)$. It is interesting to note, from Eq. (46), that even under this scaling strategy the information flow remains intensive, reaching the value $\dot{\mathcal{I}} = 2v_e/\tau_0$ for $v_e \rightarrow 0$ (recall that $\tau_0 \propto v_e$). Then the efficiency η_S is also intensive and attains the limiting value:

$$\lim_{v_e \rightarrow 0} \eta_S = \frac{\alpha^2}{2}, \quad (49)$$

as can be seen from Eqs. (46) and (31). In contrast, the efficiency η_D continuously decreases as $\eta_D \propto 1/\Omega^2 \log(\Omega)$. The behavior of the information flow $\dot{\mathcal{I}}$ under the scaling strategy above is shown in Fig. 6, where the analytical result in Eq. (46) is compared to numerical results obtained by computing the steady-state distribution as in Sec. III B and using Eq. (37) for the information flows. The numerical results are limited to low values of the scale parameter, since otherwise the exact computation of the steady-state distribution becomes too expensive. We see that the analytical result of Eq. (46) is only accurate for low values of α^2 . The reason is that the information flows are highly sensitive to the non-Gaussian nature of the exact steady-state distribution.

In relation to this last observation, it is important to discuss the role that the LD principle plays in our calculation and how it is affected by the previous scaling strategy. The LD principle and the associated rate function was only employed here as a computational tool to extract the Gaussian covariance matrix. It provides the correct Gaussian moments (related to the curvature of the rate function around its minimum) for

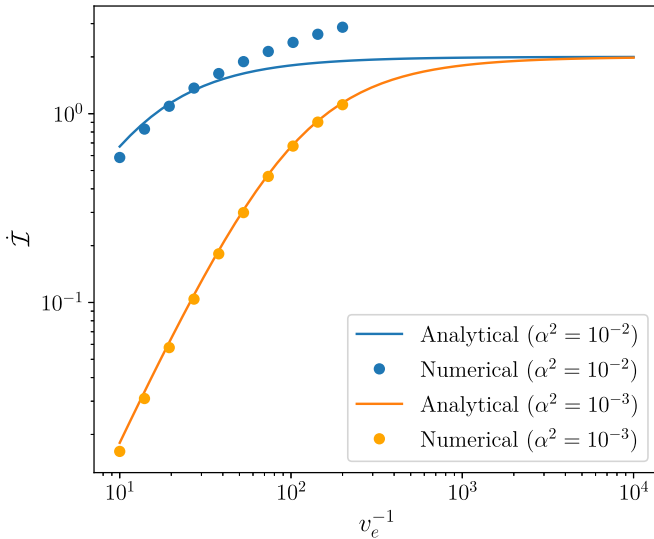


FIG. 6. Information flow as a function of the scale parameter $\Omega = v_e^{-1}$ for the scaling strategy $\Delta V_S = cv_e$ and $\Delta V_D = 2 \log(1 + 2\alpha^2/cv_e)$, for different values of α^2 . The solid lines show the analytical result of Eq. (46), while the points show numerical results obtained from the steady-state distribution and Eq. (37).

arbitrary parameters ΔV_S and ΔV_D (as long as $\alpha^2 < 1$ and the scale parameter is large enough). However, one should note that the LD principle ceases to be valid for the scaling strategy $\Delta V_S \propto 1/\Omega$ and $\Delta V_D = \log(\Omega)$, since the transition rates are no longer extensive. One way to see this is to note that the variance of the voltages does not scale as $1/\Omega$, which is the central limit theorem scaling expected from the LD principle. A related consequence is that a Gaussian approximation (that is, the truncation of $-\log[P_{ss}(v)]$ to second order around its minimum) is not justified anymore in the limit $\Omega \rightarrow \infty$, except if other conditions ($\alpha^2 \ll 1$ in our case) are met. This explains the discrepancies in Fig. 6.

Finally, we note that we have only considered the situation in which system and demon have the same scale. This is natural in this setting, since the output capacitance of the system (demon) inverter is dominated by the gate-body capacitance of the transistors in the demon (system) inverter. A similar analysis holds for the case in which only the demon inverter is scaled up. If the system is scaled up, while the demon scale is held constant, then the demon power still needs to be scaled up in order for it to be able to drive the input capacitance of the system inverter. In other words, when the demon is small compared to the system, the feedback process becomes the limiting factor in our circuit.

V. CONCLUSION

We have presented a in-depth study of the electronic implementation of an autonomous Maxwell's demon proposed in Ref. [18]. We have adapted the thermodynamic theory of information flows developed in Ref. [19] to systems displaying a macroscopic limit and applied it to the electronic demon. This allowed us to define detailed thermodynamic efficiencies for the processes creating and consuming correlations between system and demon, in terms of an information flow that was explicitly computed. Also, we have shown that the information flow is an intensive quantity in any autonomous bipartite system with a macroscopic limit, which implies that any such system will stop working as a demon above a finite scale. Implications in chemical and biological systems [27] will be explored elsewhere.

ACKNOWLEDGMENTS

This research was supported by the Project No. INTER/FNRS/20/15074473 funded by F.R.S.-FNRS (Belgium) and FNR (Luxembourg), and by the FQXi foundation Project No. FQXi-IAF19-05.

-
- [1] H. Leff and A. F. Rex, *Maxwell's Demon 2 Entropy, Classical and Quantum Information, Computing* (CRC Press, Boca Raton, FL, 2002).
 - [2] N. Freitas and M. Esposito, Characterizing autonomous Maxwell demons, *Phys. Rev. E* **103**, 032118 (2021).
 - [3] P. Strasberg, G. Schaller, T. Brandes, and M. Esposito, Thermodynamics of a Physical Model Implementing a Maxwell Demon, *Phys. Rev. Lett.* **110**, 040601 (2013).
 - [4] V. Serreli, C.-F. Lee, E. R. Kay, and D. A. Leigh, A molecular information ratchet, *Nature (Lond.)* **445**, 523 (2007).
 - [5] S. Toyabe, T. Sagawa, M. Ueda, E. Muneyuki, and M. Sano, Experimental demonstration of information-to-energy conversion and validation of the generalized Jarzynski equality, *Nat. Phys.* **6**, 988 (2010).
 - [6] T. K. Saha, J. N. E. Lucero, J. Ehrich, D. A. Sivak, and J. Bechhoefer, Maximizing power and velocity of an information engine, *Proc. Natl. Acad. Sci. USA* **118**, e2023356118 (2021).
 - [7] 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).
 - [8] 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. USA* **111**, 13786 (2014).
 - [9] K. Chida, S. Desai, K. Nishiguchi, and A. Fujiwara, Power generator driven by Maxwell's demon, *Nat. Commun.* **8**, 15310 (2017).
 - [10] P. A. Camati, J. P. S. Peterson, T. B. Batalhão, K. Micadei, A. M. Souza, R. S. Sarthour, I. S. Oliveira, and R. M. Serra, Experimental Rectification of Entropy Production by Maxwell's Demon in a Quantum System, *Phys. Rev. Lett.* **117**, 240502 (2016).
 - [11] M. D. Vidrighin, O. Dahlsten, M. Barbieri, M. S. Kim, V. Vedral, and I. A. Walmsley, Photonic Maxwell's Demon, *Phys. Rev. Lett.* **116**, 050401 (2016).
 - [12] 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. USA* **114**, 7561 (2017).
 - [13] 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).
- [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] M. Ribezzi-Crivellari and F. Ritort, Large work extraction and the Landauer limit in a continuous Maxwell demon, *Nat. Phys.* **15**, 660 (2019).
- [16] B.-L. Najera-Santos, P. A. Camati, V. Métilon, M. Brune, J.-M. Raimond, A. Auffèves, and I. Dotsenko, Autonomous Maxwell's demon in a cavity QED system, *Phys. Rev. Res.* **2**, 032025(R) (2020).
- [17] S. Amano, M. Esposito, E. Kreidt, D. A. Leigh, E. Penocchio, and B. M. W. Roberts, Insights from an information thermodynamics analysis of a synthetic molecular motor, *Nat. Chem.* **14**, 530 (2022).
- [18] N. Freitas and M. Esposito, Maxwell Demon that can Work at Macroscopic Scales, *Phys. Rev. Lett.* **129**, 120602 (2022).
- [19] J. M. Horowitz and M. Esposito, Thermodynamics with continuous information flow, *Phys. Rev. X* **4**, 031015 (2014).
- [20] D. Hartich, A. C. Barato, and U. Seifert, Stochastic thermodynamics of bipartite systems: transfer entropy inequalities and a maxwell's demon interpretation, *J. Stat. Mech.: Theory Exp.* (2014) P02016.
- [21] A. Wang, B. H. Calhoun, and A. P. Chandrakasan, *Sub-threshold Design for Ultra Low-power Systems*, Vol. 95 (Springer, Berlin, 2006).
- [22] R. Sarpeshkar, T. Delbruck, and C. A. Mead, White noise in MOS transistors and resistors, *IEEE Circ. Devices Mag.* **9**, 23 (1993).
- [23] N. Freitas, J.-C. Delvenne, and M. Esposito, Stochastic thermodynamics of nonlinear electronic circuits: A realistic framework for computing around kT , *Phys. Rev. X* **11**, 031064 (2021).
- [24] Y. Tsvividis, *Operation and Modeling of the MOS Transistor* (McGraw-Hill, New York, 1987).
- [25] A. Gopal, M. Esposito, and N. Freitas, Large deviations theory for noisy nonlinear electronics: CMOS inverter as a case study, *Phys. Rev. B* **106**, 155303 (2022).
- [26] J. Nahuel Freitas and M. Esposito, Emergent second law for non-equilibrium steady states, *Nat. Commun.* **13**, 5084 (2022).
- [27] J. Ehrich and D. A. Sivak, Energy and information flows in autonomous systems, [arXiv:2209.10644](https://arxiv.org/abs/2209.10644).