


# Emergent second law for non-equilibrium steady states

Received: 4 February 2022

Accepted: 10 August 2022

Published online: 29 August 2022

 Check for updates
José Nahuel Freitas<sup>1,2</sup>  & Massimiliano Esposito<sup>1,2</sup> 

The Gibbs distribution universally characterizes states of thermal equilibrium. In order to extend the Gibbs distribution to non-equilibrium steady states, one must relate the self-information  $\mathcal{I}(x) = -\log(P_{ss}(x))$  of microstate  $x$  to measurable physical quantities. This is a central problem in non-equilibrium statistical physics. By considering open systems described by stochastic dynamics which become deterministic in the macroscopic limit, we show that changes  $\Delta\mathcal{I} = \mathcal{I}(x_t) - \mathcal{I}(x_0)$  in steady state self-information along deterministic trajectories can be bounded by the macroscopic entropy production  $\Sigma$ . This bound takes the form of an emergent second law  $\Sigma + k_b\Delta\mathcal{I} \geq 0$ , which contains the usual second law  $\Sigma \geq 0$  as a corollary, and is saturated in the linear regime close to equilibrium. We thus obtain a tighter version of the second law of thermodynamics that provides a link between the deterministic relaxation of a system and the non-equilibrium fluctuations at steady state. In addition to its fundamental value, our result leads to novel methods for computing non-equilibrium distributions, providing a deterministic alternative to Gillespie simulations or spectral methods.

When a system is at equilibrium with its environment (i.e., when no energy currents are exchanged) the probability of a given microstate  $\mathbf{x}$  is given by the Gibbs distribution<sup>1–3</sup>

$$P_{\text{eq}}(\mathbf{x}) = e^{-\beta\Phi(\mathbf{x})}/Z, \quad (1)$$

where  $\beta = (k_b T)^{-1}$  is the inverse temperature of the environment,  $\Phi(\mathbf{x})$  is the free energy of microstate  $\mathbf{x}$  (for states with no internal entropy,  $\Phi(\mathbf{x})$  is just the energy), and  $Z = \sum_{\mathbf{x}} \exp(-\beta\Phi(\mathbf{x}))$  is the partition function. This central result of equilibrium statistical physics has universal validity and its relevance in most areas of physics cannot be overstated. A natural question is whether or not a similar result also holds for nonequilibrium steady states (NESSs), when the system is maintained out of thermal equilibrium by external drives and subjected to constant flows of energy or matter. In this case, one can always write the steady state distribution over

microstates as

$$P_{ss}(\mathbf{x}) = e^{-\mathcal{I}(\mathbf{x})} \quad (2)$$

in terms of the self-information  $\mathcal{I}(\mathbf{x})$ , also known as fluctuating entropy<sup>4,5</sup>. In order to provide a useful generalization of the Gibbs distribution to NESSs one must relate the self-information  $\mathcal{I}(\mathbf{x})$  to measurable physical quantities. This quest has a long history, starting with the seminal contributions of Lebowitz and MacLennan<sup>6–8</sup> and followed by other works<sup>9–15</sup>. However, it still remains an open problem in non-equilibrium statistical physics, since previous formal results are simply not practical in computations, due to the fact that they involve averages over stochastic trajectories.

In this work we prove, for a very general class of open systems displaying a macroscopic limit where a deterministic dynamics

<sup>1</sup>Complex Systems and Statistical Mechanics, Department of Physics and Materials Science, University of Luxembourg, 162a, avenue de la Faiencerie, Luxembourg L-1511 Luxembourg, Luxembourg. <sup>2</sup>These authors contributed equally: José Nahuel Freitas, Massimiliano Esposito. ✉ e-mail: [nahuel.freitas@uni.lu](mailto:nahuel.freitas@uni.lu); [massimiliano.esposito@uni.lu](mailto:massimiliano.esposito@uni.lu)

emerges, the following fundamental bound on changes of self-information:

$$\Sigma_a \equiv \Sigma + k_b(\mathcal{I}(\mathbf{x}_t) - \mathcal{I}(\mathbf{x}_0)) \geq 0, \quad (3)$$

where  $\Sigma = \int_0^t dt \dot{\Sigma}(\mathbf{x}_t)/T$  is the entropy production along a deterministic trajectory from microstate  $\mathbf{x}_0$  to microstate  $\mathbf{x}_t$ . For example, let us consider the case of chemical reaction networks. The concentrations  $\mathbf{x} = (x_1, x_2, \dots)$  of different chemical species reacting in a solution are stochastic quantities, and their evolution is therefore described by a probability distribution  $P_t(\mathbf{x})$  at time  $t$ . As the volume  $V$  of the solution is increased, the distribution  $P_t(\mathbf{x})$  becomes strongly localised around the most probable values  $\mathbf{x}_t$  for the concentrations at time  $t$ , and these values follow a deterministic dynamics that is in general nonlinear (given in this case by the chemical rate equations). An analogous situation is encountered in electronic circuits, where the state variables  $\mathbf{x}$  are now the voltages at the nodes of a circuit, and the macroscopic limit corresponds to increasing the typical capacitance  $C$  of the nodes (as well as the conductivity of the conduction channels connecting pairs of nodes). The remarkable feature of the result in Eq. (3) is that it provides a link between the deterministic dynamics that emerges in the macroscopic limit and the fluctuations observed at steady state. For example, in an electronic circuit powered by voltage sources and working at temperature  $T$ , the entropy production rate is  $\dot{\Sigma} = -\dot{Q}/T$ , where  $-\dot{Q}$  is the rate of heat dissipation by the conductive elements of the circuit, that can be easily evaluated at the deterministic level. Then, by Eq. (3), the quantity  $-\Sigma/k_b = \int_0^t dt \dot{Q}(\mathbf{x}_t)/(k_b T)$  provides a lower bound to the change of steady state self-information  $\mathcal{I}(\mathbf{x}_t) - \mathcal{I}(\mathbf{x}_0)$  along a trajectory. To arrive at our main result in Eq. (3) we consider stochastic systems with a well defined macroscopic limit in which the self-information  $\mathcal{I}(x)$  can be shown to be extensive, and Eq. (3) is strictly valid in that limit. However, as we also show, our results can be applied to micro or mesoscopic systems whenever sub-extensive contributions to  $\mathcal{I}(x)$  can be neglected. We interpret our result as an emergent second law of thermodynamics, that is stronger than the usual second law  $\Sigma \geq 0$ . This last inequality is recovered from Eq. (3) by considering the fact that  $\Delta\mathcal{I} = \mathcal{I}(\mathbf{x}_t) - \mathcal{I}(\mathbf{x}_0) \leq 0$  (to dominant order in the macroscopic limit, the steady state self-information is a Lyapunov function of the deterministic dynamics<sup>16</sup>). In addition to its conceptual value, our result offers a practical tool to approximate or bound non-equilibrium distributions, that can typically only be accessed via stochastic numerical methods (for example the Gillespie algorithm). In contrast, Eq. (3) only requires to know the deterministic dynamics of the system, which is directly given by the well known network analysis techniques commonly applied in electronic circuits and chemical reaction networks. Furthermore, the inequality in Eq. (3) is saturated close to equilibrium, leading to a powerful linear response theory<sup>17</sup>. As an example, we apply our results to a realistic model of non-equilibrium electronic memory: the normal implementation of SRAM (static random access memory) cells in CMOS (complementary metal-oxide-semiconductor) technology. These memories have a non-equilibrium phase transition from a monostable phase to a bistable phase that allows the storage of a bit of information. As we will see, the transition is well captured by Eq. (3), which also allows to bound the probability of fluctuations around the deterministic fixed points. Finally, we show that a general coarse-graining procedure generates equivalent models with minimal entropy production, and that in this way the bound in Eq. (3) becomes tighter. When applied to the CMOS memory, this improved bound enables the full reconstruction of the steady state distribution arbitrarily away from equilibrium.

## Results

To obtain Eq. (3) we consider stochastic systems described by autonomous Markov jump processes. Thus, let  $\{\mathbf{n} \in \mathbb{N}^k\}$  be the set of possible states of the system, and  $\lambda_\rho(\mathbf{n})$  be the rates at which jumps

$\mathbf{n} \rightarrow \mathbf{n} + \mathbf{\Delta}_\rho$  occur, for  $\rho = \pm 1, \pm 2, \dots$  and  $\mathbf{\Delta}_{-\rho} = -\mathbf{\Delta}_\rho$  ( $\rho$  indexes a possible jump and  $\mathbf{\Delta}_\rho$  is the corresponding change in the state). Each state has energy  $E(\mathbf{n})$  and internal entropy  $S(\mathbf{n})$ . Thermodynamic consistency is introduced by the local detailed balance (LDB) condition<sup>18,19</sup>. It relates the forward and backward jump rates of a given transition with the associated entropy production:

$$\sigma_\rho = \log \frac{\lambda_\rho(\mathbf{n})}{\lambda_{-\rho}(\mathbf{n} + \mathbf{\Delta}_\rho)} = -\beta [\Phi(\mathbf{n} + \mathbf{\Delta}_\rho) - \Phi(\mathbf{n}) - W_\rho(\mathbf{n})]. \quad (4)$$

In the previous equation,  $\Phi(\mathbf{n}) = E(\mathbf{n}) - TS(\mathbf{n})$  is the free energy of state  $\mathbf{n}$ , and  $W_\rho(\mathbf{n})$  is the non-conservative work provided by external sources during the transition. For simplicity, we have considered isothermal conditions at inverse temperature  $\beta = (k_b T)^{-1}$ , and therefore the system is taken away from equilibrium by the external work sources alone. More general situations in which a system interacts with several reservoirs at different temperatures can be treated in the same way, this time in terms of a Massieu potential taking the place of  $\beta\Phi(\mathbf{n})$ <sup>18</sup>. Important classes of systems accepting the previous description are chemical reaction networks and electronic circuits, which are powered by chemical or electrostatic potential differences, respectively. Note that, by energy conservation, the heat provided by the environment during transition  $\mathbf{n} \rightarrow \mathbf{n} + \mathbf{\Delta}_\rho$  is  $Q_\rho(\mathbf{n}) = E(\mathbf{n} + \mathbf{\Delta}_\rho) - E(\mathbf{n}) - W_\rho(\mathbf{n})$ , and therefore  $k_b\sigma_\rho = -Q_\rho(\mathbf{n})/T + S(\mathbf{n} + \mathbf{\Delta}_\rho) - S(\mathbf{n})$ .

The probability distribution  $P_t(\mathbf{n})$  over the states of the system at time  $t$  evolves according to the master equation

$$\partial_t P_t(\mathbf{n}) = \sum_\rho [\lambda_\rho(\mathbf{n} - \mathbf{\Delta}_\rho) P_t(\mathbf{n} - \mathbf{\Delta}_\rho) - \lambda_\rho(\mathbf{n}) P_t(\mathbf{n})]. \quad (5)$$

From the master equation and the LDB conditions one can derive the energy balance

$$d_t \langle E \rangle = \langle \dot{W} \rangle + \langle \dot{Q} \rangle, \quad (6)$$

and the usual version of the second law:

$$\begin{aligned} \dot{\Sigma} &= \dot{\Sigma}_e + d_t \langle S \rangle \\ &= \frac{k_b}{2} \sum_{\rho, \mathbf{n}} (j_\rho(\mathbf{n}) - j_{-\rho}(\mathbf{n} + \mathbf{\Delta}_\rho)) \log \frac{j_\rho(\mathbf{n})}{j_{-\rho}(\mathbf{n} + \mathbf{\Delta}_\rho)} \geq 0, \end{aligned} \quad (7)$$

where  $j_\rho(\mathbf{n}) = \lambda_\rho(\mathbf{n}) P_t(\mathbf{n})$  is the current associated to transition  $\rho$ . In the previous equations,  $\langle S \rangle = \sum_{\mathbf{n}} P_t(\mathbf{n}) (S(\mathbf{n}) - k_b \log(P_t(\mathbf{n})))$  is the entropy of the system,  $\langle E \rangle = \sum_{\mathbf{n}} E(\mathbf{n}) P_t(\mathbf{n})$  is the average energy, and  $\dot{\Sigma}_e$  is the entropy flow rate, given by

$$T \dot{\Sigma}_e = -\langle \dot{Q} \rangle = -\sum_{\rho, \mathbf{n}} Q_\rho(\mathbf{n}) j_\rho(\mathbf{n}) \quad (8)$$

where we also defined the heat rate  $\langle \dot{Q} \rangle$  (the work rate  $\langle \dot{W} \rangle$  is analogously defined as  $\langle \dot{W} \rangle = \sum_{\rho, \mathbf{n}} W_\rho(\mathbf{n}) j_\rho(\mathbf{n})$ ). Finally, Eq. (7) can be also expressed as:

$$T \dot{\Sigma} = -d_t \langle F \rangle + \langle \dot{W} \rangle \geq 0 \quad (9)$$

where  $\langle F \rangle = \langle E \rangle - T \langle S \rangle$  is the non-equilibrium free energy.

### Adiabatic/nonadiabtic decomposition

If the support of  $P_t(\mathbf{n})$  can be restricted to a finite subspace of the state space, the Perron-Frobenius theorem states that the master equation in Eq. (5) has a unique steady state  $P_{ss}(\mathbf{n})$ . Once the steady state is attained, the entropy production rate  $\dot{\Sigma}$  matches the entropy flow rate  $\dot{\Sigma}_e$ . An interesting decomposition of the entropy production rate can be obtained by considering the relative entropy  $D = \sum_{\mathbf{n}} P_t(\mathbf{n}) \log(P_t(\mathbf{n})/P_{ss}(\mathbf{n}))$  between the instantaneous distribution

$P_t(\mathbf{n})$  and the steady state distribution  $P_{ss}(\mathbf{n})$ . Then, it is possible to show that  $\dot{\Sigma} = \dot{\Sigma}_a + \dot{\Sigma}_{na}$ , where

$$\dot{\Sigma}_a = \frac{k_b}{2} \sum_{\rho, \mathbf{n}} (j_\rho(\mathbf{n}) - j_{-\rho}(\mathbf{n} + \mathbf{A}_\rho)) \log \frac{j_\rho^{ss}(\mathbf{n})}{j_{-\rho}^{ss}(\mathbf{n} + \mathbf{A}_\rho)}, \quad (10)$$

and

$$\begin{aligned} \dot{\Sigma}_{na} &= \frac{k_b}{2} \sum_{\rho, \mathbf{n}} (j_\rho(\mathbf{n}) - j_{-\rho}(\mathbf{n} + \mathbf{A}_\rho)) \log \frac{P_t(\mathbf{n})P_{ss}(\mathbf{n} + \mathbf{A}_\rho)}{P_{ss}(\mathbf{n})P_t(\mathbf{n} + \mathbf{A}_\rho)} \\ &= -k_b d_t D \end{aligned} \quad (11)$$

are the adiabatic and non-adiabatic contributions to the entropy production rate  $\dot{\Sigma}$ , respectively. In Eq. (10) we have introduced the steady state probability currents  $j_\rho^{ss}(\mathbf{n}) = \lambda_\rho(\mathbf{n})P_{ss}(\mathbf{n})$ . The non-adiabatic contribution  $\dot{\Sigma}_{na}$  is related to the relaxation of the system towards the steady state, since it vanishes when the steady state is reached. This is further evidenced by the identity in the second line of Eq. (11): a reduction in the relative entropy between  $P_t(\mathbf{n})$  and  $P_{ss}(\mathbf{n})$  leads to a positive non-adiabatic entropy production. The adiabatic contribution  $\dot{\Sigma}_a$  corresponds to the dissipation of ‘house-keeping heat’<sup>20,21</sup>, and at steady state matches the entropy flow rate  $\dot{\Sigma}_e$ . An important property of the previous decomposition is that both contributions are individually positive:  $\dot{\Sigma}_a \geq 0$  and  $\dot{\Sigma}_{na} \geq 0$ <sup>22–25</sup>. Thus, the last inequality and the second line in Eq. (11) imply that the relative entropy  $D$  decreases monotonically, and since  $D$  is positive by definition, it is a Lyapunov function for the stochastic dynamics.

### Macroscopic limit

In the following we will assume the existence of a scale parameter  $\Omega$  controlling the size of the system in question. For example,  $\Omega$  can be taken to be the volume  $V$  of the solution in well-mixed chemical reaction networks, or the typical value  $C$  of capacitance in the case of electronic circuits (see the example below). In addition, we will assume that for large  $\Omega$  i) that the typical values of the density  $\mathbf{x} = \mathbf{n}/\Omega$  are intensive, ii) that the internal energy and entropy functions  $E(\Omega\mathbf{x})$  and  $S(\Omega\mathbf{x})$  are extensive, and iii) that the transition rates  $\lambda_\rho(\Omega\mathbf{x})$  are also extensive. Under those conditions, the probability distribution  $P_t(\mathbf{x})$  satisfies a large deviations (LD) principle<sup>17,26,27</sup>:

$$P_t(\mathbf{x}) \asymp e^{-\Omega I_t(\mathbf{x})}, \quad (12)$$

which just means that the limit  $I_t(\mathbf{x}) \equiv \lim_{\Omega \rightarrow \infty} -\log(P_t(\mathbf{x}))/\Omega$  is well defined. Then,  $I_t(\mathbf{x})$  is a positive, time-dependent ‘rate function’, since it gives the rate at which the probability of fluctuation  $\mathbf{x}$  decays with the scale. Note that, by Eq. (12), the steady state self-information introduced in Eq. (2) satisfies  $\mathcal{I}(\mathbf{x}) = \Omega I_{ss}(\mathbf{x})$  to dominant order in the macroscopic limit. In other words, the large deviations principle states that the instantaneous self-information  $\mathcal{I}_t(\mathbf{x}) \equiv -\log(P_t(\mathbf{x}))$  is an extensive quantity<sup>26</sup>, and we can think of the rate function as the self-information density. Thus, in the following we will consider the ansatz  $P_t(\mathbf{x}) = e^{-\Omega I_t(\mathbf{x})}/Z_t$ , with  $Z_t \equiv \sum_{\mathbf{x}} e^{-\Omega I_t(\mathbf{x})}$ , as an approximation to the actual time-dependent distribution. This amounts to neglecting sub-extensive contributions to the instantaneous self-information. As explained below,  $I_t(\mathbf{x})$  takes its minimum value  $I_t(\mathbf{x}_t) = 0$  at the deterministic trajectory  $\mathbf{x}_t$ , which is equivalent to  $P_t(\mathbf{x}) = \delta(\mathbf{x} - \mathbf{x}_t)$  for  $\Omega \rightarrow \infty$ . Plugging the previous ansatz in the master equation of Eq. (5) we note that  $\lambda_\rho(\mathbf{x} - \mathbf{A}_\rho/\Omega)P_t(\mathbf{x} - \mathbf{A}_\rho/\Omega) \simeq \lambda_\rho(\mathbf{x})P_t(\mathbf{x})e^{\mathbf{A}_\rho \cdot \nabla I_t(\mathbf{x})}$  to dominant order in  $\Omega \rightarrow \infty$ . Noting also that  $\log(Z_t)$  is sub-extensive, it is possible to see that  $I_t(\mathbf{x})$  evolves according to

$$\partial_t I_t(\mathbf{x}) = \sum_{\rho} \omega_\rho(\mathbf{x}) [1 - e^{\mathbf{A}_\rho \cdot \nabla I_t(\mathbf{x})}], \quad (13)$$

where  $\omega_\rho(\mathbf{x}) \equiv \lim_{\Omega \rightarrow \infty} \lambda_\rho(\Omega\mathbf{x})/\Omega$  are the scaled jump rates<sup>17,28</sup>. In a similar way, in the macroscopic limit the LDB conditions in Eq. (4) take the form

$$\log \frac{\omega_\rho(\mathbf{x})}{\omega_{-\rho}(\mathbf{x})} = -\beta [\mathbf{A}_\rho \cdot \nabla \phi(\mathbf{x}) - W_\rho(\mathbf{x})], \quad (14)$$

in terms of the free energy density  $\phi(\mathbf{x}) \equiv \lim_{\Omega \rightarrow \infty} \Phi(\Omega\mathbf{x})/\Omega$  (internal energy and entropy densities  $\epsilon(\mathbf{x})$  and  $s(\mathbf{x})$  satisfying  $\phi(\mathbf{x}) = \epsilon(\mathbf{x}) - Ts(\mathbf{x})$  can be defined in the same way). For the work contributions in Eq. (14) we are abusing notation by writing  $W_\rho(\mathbf{x}) = \lim_{\Omega \rightarrow \infty} W_\rho(\mathbf{n} = \Omega\mathbf{x})$ . Note that we assume that work contributions are intensive. This is justified since they are given by the product of two intensive quantities: a thermodynamic force (for example a potential difference), and the change in a conserved quantity (mass, charge, etc) during a single jump<sup>29</sup>. However, note also that the work rate  $\langle \dot{W} \rangle$  will be extensive in general due to the extensivity of the transition rates.

Many classes of systems satisfy the previous scaling assumptions besides the examples already mentioned. Additional examples include non-equilibrium many-body problems like the driven Potts model<sup>30,31</sup>, reaction-diffusion models<sup>32,33</sup>, and asymmetric exclusion processes<sup>32,34</sup>.

From Eq. (12) we see that as  $\Omega$  is increased,  $P_t(\mathbf{x})$  is increasingly localised around the minimum of the rate function  $I_t(\mathbf{x})$ , which is the most probable value. Also, deviations from that typical state are exponentially suppressed in  $\Omega$ . Thus, the limit  $\Omega \rightarrow \infty$  is a macroscopic low-noise limit where a deterministic dynamic emerges. In fact, from Eq. (13) one can show that the evolution of the minima  $\mathbf{x}_t$  of  $I_t(\mathbf{x})$  is ruled by the closed non-linear differential equations

$$d_t \mathbf{x}_t = \mathbf{u}(\mathbf{x}_t), \text{ with } \mathbf{u}(\mathbf{x}) \equiv \sum_{\rho > 0} i_\rho(\mathbf{x}) \mathbf{A}_\rho, \quad (15)$$

where  $i_\rho(\mathbf{x}) \equiv \omega_\rho(\mathbf{x}) - \omega_{-\rho}(\mathbf{x})$  are the scaled deterministic currents<sup>17</sup>. The vector field  $\mathbf{u}(\mathbf{x})$  corresponds to the deterministic drift in state space. For chemical reaction networks the dynamical equations in Eq. (15) are the chemical rate equations, while for electronic circuits they are provided by regular circuit analysis.

In the following section we obtain bounds for the steady state rate function  $I_{ss}(\mathbf{x})$ , that according to Eq. (13) satisfies:

$$0 = \sum_{\rho} \omega_\rho(\mathbf{x}) [1 - e^{\mathbf{A}_\rho \cdot \nabla I_{ss}(\mathbf{x})}]. \quad (16)$$

### Emergent second law

The positivity of the adiabatic and non-adiabatic contributions to the entropy production,  $\dot{\Sigma}_a \geq 0$  and  $\dot{\Sigma}_{na} \geq 0$ , in addition to the usual second law  $\dot{\Sigma} \geq 0$ , have been called the ‘three faces of the second law’<sup>23</sup>. In<sup>28</sup>, the inequality  $\dot{\Sigma}_{na} = -k_b d_t D \geq 0$  was put forward as an ‘emergent’ second law. There,  $\mathcal{F} = k_b D$  was interpreted as an alternative non-equilibrium free energy, with a balance equation  $d_t \mathcal{F} = \dot{\Sigma}_a - \dot{\Sigma} \leq 0$  (note the analogy with Eq. (9)). Then, the adiabatic contribution  $\dot{\Sigma}_a$  was interpreted as an energy input, which at steady state balances the dissipation  $\dot{\Sigma}$ . Although this point of view is compelling, it is hindered by the fact that there is no clear interpretation of  $\dot{\Sigma}_a$  away from the steady state, that would allow to compute this quantity in terms of actual physical currents. In this work we take the other possible road, and investigate the interpretation and consequences of  $\dot{\Sigma}_a \geq 0$ . We begin by rewriting Eq. (10) using the LDB conditions of Eq. (4) and the definition of  $\mathcal{I}$  in Eq. (2), obtaining:

$$\dot{\Sigma}_a = \dot{\Sigma} + k_b d_t \langle \mathcal{I} \rangle - d_t \langle S_{sh} \rangle \geq 0, \quad (17)$$

where we have defined  $\langle \mathcal{I} \rangle = \sum_{\mathbf{n}} \mathcal{I}(\mathbf{n}) P_t(\mathbf{n})$  as the average of the steady state self-information  $\mathcal{I}(\mathbf{n}) = -\log(P_{ss}(\mathbf{n}))$  and  $\langle S_{sh} \rangle = -k_b \sum_{\mathbf{n}} P_t(\mathbf{n}) \log(P_t(\mathbf{n}))$  as the Shannon contribution to the system entropy, computed over the instantaneous distribution. Eq. (17) has been already obtained in<sup>22–24</sup>, although it was not explicitly written in terms of the self-information  $\mathcal{I}$ . It is important to note that  $\langle S_{sh} \rangle$  is sub-extensive in  $\Omega$  (according to Eq. (12), it grows as  $\log(\Omega)$ ), and therefore can be neglected in the macroscopic limit. Thus, changes in average self-information can be bounded by the entropy production, that can in turn be computed or measured in terms of actual energy and entropy flows (see Eqs. (7) and (8)). However, the result in Eq. (17) is not yet in a useful form, since the average  $\langle \mathcal{I} \rangle$  does not depend only on  $\mathcal{I}(\mathbf{n})$ , the unknown quantity we are interested in, but also on the instantaneous distribution  $P_t(\mathbf{n})$ , that is also typically unknown. This issue is circumvented in the macroscopic limit, since in that case  $P_t(\mathbf{x})$  is strongly localised around the deterministic values  $\mathbf{x}_t$ , and therefore  $\langle \mathcal{I} \rangle \simeq \Omega I_{ss}(\mathbf{x}_t)$  to dominant order in  $\Omega \rightarrow \infty$ . Thus, in the same limit, Eq. (17) for the adiabatic entropy production rate  $\dot{\Sigma}_a$  reduces to

$$\dot{\sigma}_a(\mathbf{x}_t) = \dot{\sigma}(\mathbf{x}_t) + k_b d_t I_{ss}(\mathbf{x}_t) \geq 0, \quad (18)$$

where we have defined  $\dot{\sigma}(\mathbf{x}_t) = \lim_{\Omega \rightarrow \infty} \dot{\Sigma}_a / \Omega$  as the scaled macroscopic limit of the entropy production rate ( $\dot{\sigma}_a(\mathbf{x}_t)$  is defined in a similar way). Eq. (18) is a more rigorous version of our central result in Eq. (3), which is obtained by integrating Eq. (18) along deterministic trajectories (satisfying Eq. (15)) and multiplying by the scale factor. It is also useful to write down the first and second laws in the macroscopic limit. The energy balance in Eq. (6) reduces to

$$d_t \epsilon(\mathbf{x}_t) = \mathbf{u}(\mathbf{x}_t) \cdot \nabla \epsilon(\mathbf{x}_t) = \dot{w}(\mathbf{x}_t) + \dot{q}(\mathbf{x}_t), \quad (19)$$

where the scaled heat and work rates for state  $\mathbf{x}$  are defined as  $\dot{q}(\mathbf{x}) = \sum_{\rho>0} i_{\rho}(\mathbf{x}) Q_{\rho}(\mathbf{x})$  and  $\dot{w}(\mathbf{x}) = \sum_{\rho>0} i_{\rho}(\mathbf{x}) W_{\rho}(\mathbf{x})$ , respectively. Finally, again neglecting the sub-extensive Shannon contribution  $\langle S_{sh} \rangle$ , the second law in Eq. (7) reduces to

$$\begin{aligned} \dot{\sigma}(\mathbf{x}_t) &= -\dot{q}(\mathbf{x}_t)/T + d_t s(\mathbf{x}_t) \\ &= k_b \sum_{\rho>0} (\omega_{\rho}(\mathbf{x}_t) - \omega_{-\rho}(\mathbf{x}_t)) \log \frac{\omega_{\rho}(\mathbf{x}_t)}{\omega_{-\rho}(\mathbf{x}_t)} \geq 0. \end{aligned} \quad (20)$$

### Linear response regime

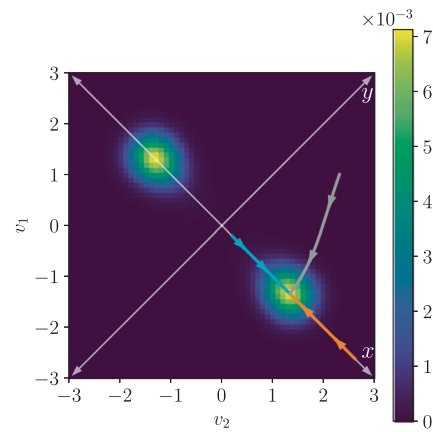
We will now show that to first order in the work contributions  $W_{\rho}(\mathbf{x})$  the inequality in Eq. (18) is saturated. In first place we rewrite Eq. (18) using the macroscopic first and second laws in Eqs. (19) and (20):

$$\dot{\sigma}_a(\mathbf{x}_t)/k_b = \mathbf{u}(\mathbf{x}_t) \cdot \nabla (I_{ss}(\mathbf{x}) - \beta \phi(\mathbf{x}))|_{\mathbf{x}=\mathbf{x}_t} + \beta \dot{w}(\mathbf{x}_t), \quad (21)$$

where we also used  $\phi(\mathbf{x}) = \epsilon(\mathbf{x}) - T s(\mathbf{x})$  and that  $d_t F(\mathbf{x}_t) = \mathbf{u}(\mathbf{x}_t) \cdot \nabla F(\mathbf{x}_t)$  for any function  $F(\mathbf{x})$ . Secondly, we note that in detailed-balanced settings (i.e., if  $W_{\rho}(\mathbf{x}) = 0 \forall \rho, \mathbf{x}$ ) the steady state rate function is just  $I_{ss}(\mathbf{x}) = \beta \phi(\mathbf{x})$  (up to a constant), in accordance to the Gibbs distribution (this follows from Eqs. (14) and (16)). Thus, the difference  $g(\mathbf{x}) \equiv I_{ss}(\mathbf{x}) - \beta \phi(\mathbf{x})$  appearing in Eq. (18) quantifies the deviations from thermal equilibrium. Expanding Eq. (16) to first order in  $W_{\rho}(\mathbf{x})$  and  $g(\mathbf{x})$ , it can be shown that

$$\mathbf{u}^{(0)}(\mathbf{x}) \cdot \nabla g(\mathbf{x}) = -\beta \dot{w}^{(0)}(\mathbf{x}) + \mathcal{O}(W_{\rho}^2), \quad (22)$$

where  $\mathbf{u}^{(0)}(\mathbf{x}) = \sum_{\rho>0} i_{\rho}^{(0)}(\mathbf{x}) \mathbf{\Delta}_{\rho}$  and  $\dot{w}^{(0)}(\mathbf{x}) = \sum_{\rho>0} i_{\rho}^{(0)}(\mathbf{x}) W_{\rho}(\mathbf{x})$  are the lowest-order deterministic drift and work rate, respectively<sup>17</sup>. These are defined in terms of the detailed-balanced deterministic currents  $i_{\rho}^{(0)}(\mathbf{x}) = \omega_{\rho}^{(0)}(\mathbf{x}) - \omega_{-\rho}^{(0)}(\mathbf{x})$  constructed from the scaled transition rates evaluated at  $W_{\rho}(\mathbf{x}) = 0$  that, according to the LDB conditions of Eq. (14),



**Fig. 1 | Steady state and deterministic trajectories for the CMOS memory.** Exact steady state distribution for  $V_{dd} = 1.4$  and  $v_e = 0.1$ . The orange, cyan and grey lines show three deterministic trajectories for the same parameters and different initial conditions, with the arrows indicating the direction of motion.

satisfy  $\log(\omega_{\rho}^{(0)}(\mathbf{x})/\omega_{-\rho}^{(0)}(\mathbf{x})) = -\beta \mathbf{\Delta}_{\rho} \cdot \nabla \phi(\mathbf{x})$ . Comparing the result of Eq. (22) with Eq. (21), we see that  $\dot{\Sigma}_a = 0$  to linear order in  $W_{\rho}(\mathbf{x})$ . Then, in the linear response regime we can write:

$$\begin{aligned} I_{ss}(\mathbf{x}_t) - I_{ss}(\mathbf{x}_0) &\simeq - \int_0^t dt' \dot{\sigma}^{(0)}(\mathbf{x}_{t'}) / k_b \\ &= \beta \left[ \phi(\mathbf{x}_t) - \phi(\mathbf{x}_0) - \int_0^t dt' \dot{w}^{(0)}(\mathbf{x}_{t'}) \right], \end{aligned} \quad (23)$$

where the integration is performed along trajectories solving the detailed-balanced deterministic dynamics  $d_t \mathbf{x}_t = \mathbf{u}^{(0)}(\mathbf{x}_t)$ .

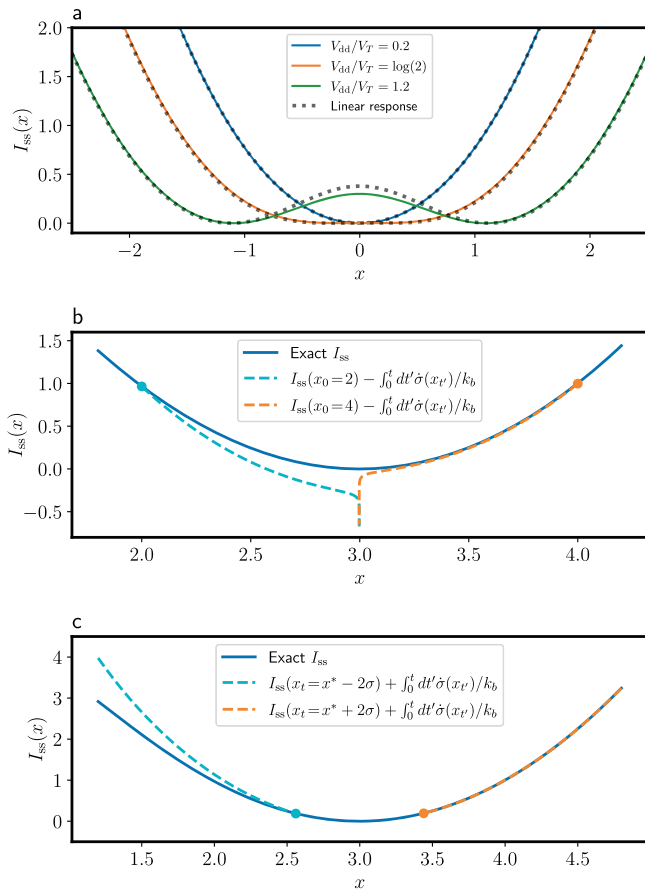
### Example: an electronic memory

In order to illustrate our results we will consider the model of a low-power CMOS memory cell developed in<sup>35,36</sup>, that we review in the Supplementary Note 1. This model involves two CMOS inverters connected in a loop, and each inverter is composed of two MOS transistors. There are two degrees of freedom: voltages  $v_1$  and  $v_2$ , that can take values spaced by the elementary voltage  $v_e = q_e/C$ , where  $q_e$  is the positive electron charge and  $C$  is a value of capacitance that increases with the scale of the MOS transistors. Thus, in this context the scale parameter can be taken to be  $\Omega = V_T/v_e$ , where  $V_T = k_b T/q_e$  is the thermal voltage. In the following all voltages will be expressed in units of  $V_T$ . Figure 1 shows a typical steady state distribution in the bistable phase, and three different deterministic trajectories. The logical state of the memory is codified in the sign of the variable  $x = v_1 - v_2$ . The rate function  $I_{ss}(x)$  along the  $x$  axis in Fig. 1 can be computed exactly<sup>36</sup>:

$$I_{ss}(x) = x^2 + x V_{dd} + \frac{2n}{n+2} [L(x, V_{dd}) - L(x, -V_{dd})], \quad (24)$$

where  $V_{dd}$  is the powering voltage that takes the memory out of thermal equilibrium,  $L(x, V_{dd}) = \text{Li}_2(-\exp(V_{dd} + x(1 + 2/n)))$ , and  $\text{Li}_2(\cdot)$  is the polylogarithm function of second order. Also,  $n \geq 1$  is a parameter that characterizes the transistors (the slope factor), and that will be fixed to  $n = 1$  in the following. In the Supplementary Note 1 we show that the rate function in Eq. (24) provides an essentially exact description of the steady state distribution for scale parameters as low as  $\Omega = 10$  (i.e., states with at most tens of electrons are appreciably populated). Thus, sub-extensive contributions to the self-information can be safely neglected in this case, even away from the strict macroscopic limit.

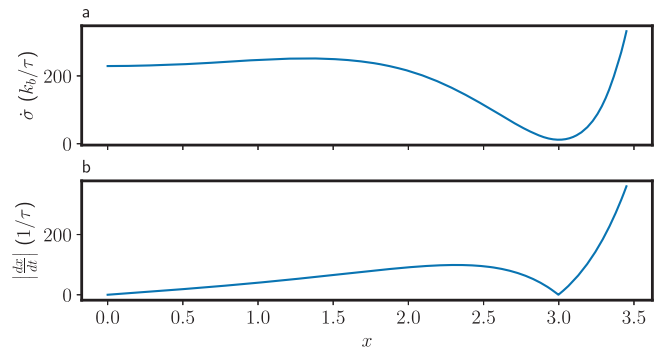
In Fig. 2a we show that there is a non-equilibrium transition from a monostable phase into the bistable phase that allows the storage of a



**Fig. 2 | Rate function of the CMOS memory.** **a** Rate function  $I_{ss}(x)$  for different values of the powering voltage  $V_{dd}$ . We also compare  $I_{ss}(x)$  with the linear response approximation obtained from Eq. (23) for  $V_{dd}=1.2$ . **b** Illustration of the bound to  $I(x_t)$  in Eq. (18) for two different deterministic trajectories starting at  $x_0 = 2$  and  $x_0 = 4$  ( $V_{dd} = 3$ ). **c** Bound to  $I_{ss}(x_t)$  for a fixed  $x_t$  according to Eq. (18), for  $x_t = x^* \pm 2\delta$  ( $\delta = 1/\sqrt{2\Omega}$ ,  $\Omega = 10$ , and  $V_{dd} = 3$ ).

bit of information, occurring at the critical powering voltage  $V_{dd}^* = \ln(2)$ . We also compare the exact rate function in Eq. (24) with the linear response approximation obtained from Eq. (23) for different powering voltages. Remarkably, despite it is only expected to be valid close to equilibrium, the linear response approximation captures the transition to bistability and continues to be a reasonable approximation well into the bistable phase. The reason is that in this case the second order non-equilibrium correction vanishes (as can be checked by expanding Eq. (24) in  $V_{dd}$ ), and therefore the linear response approximation is actually valid up to order  $V_{dd}^3$ .

In Fig. 2b the exact rate function  $I_{ss}(x_t)$  along a deterministic trajectory  $x_t$  is compared with the lower bound  $I_{ss}(x_0) - \int_0^t dt' \dot{\sigma}(x_{t'})/k_b$ , for two different trajectories starting at  $x_0 = 2$  and  $x_0 = 4$ , with  $y_0 = 0$  in both cases (as shown in Fig. 1, deterministic trajectories initialized in the  $x$  axis remain in it). In both cases the trajectory  $x_t$  approaches the fixed point  $x^* \simeq V_{dd} = 3$ . We see that  $I_{ss}(x_0) - \int_0^t dt' \dot{\sigma}(x_{t'})/k_b$  is indeed a lower bound to  $I_{ss}(x_t)$ , in accordance with Eq. (18). Note that this bound diverges when the trajectory approaches the fixed point  $x^*$ . The reason is that once  $x_t \simeq x^*$ , the entropy production  $\int_0^t dt' \dot{\sigma}(x_{t'})$  just continuously integrates the steady state heat dissipation rate  $-\dot{q}(x^*)$  (see Eq. (20)). The linear response approximation avoids this issue since the lowest-order work rate  $\dot{w}^{(0)}(x)$  vanishes at the equilibrium fixed point (see Eq. (23)). The divergence can be also avoided by the coarse-graining procedure discussed in the next section. Alternatively, Eq. (18) can be considered an upper bound to  $I(x_0)$  for a fixed final point  $x_t$ . This



**Fig. 3 | Entropy production rate and speed along deterministic trajectories.** **a** Deterministic entropy production rate  $\dot{\sigma}$ , and **(b)** deterministic speed  $|dx/dt|$  as a function of  $x$  for  $V_{dd} = 3$  (see the Supplementary Note 1 for the definition of the timescale  $\tau$ ).

is shown in Fig. 2c, for final points  $x_t = x^* \pm 2\delta$ , where  $\delta = 1/\sqrt{2\Omega}$  estimates the variance of the fluctuations around the fixed point.

The fact that in Fig. 2b, c the bounds are much tighter to one side of the fixed point than the other can be traced back to the different speeds at which the fixed point is approached. To show this, in Fig. 3 we have plotted the entropy production rate  $\dot{\sigma}$  and the speed  $|dx/dt|$  as a function of  $x$  for  $V_{dd} = 3$ . In Fig. 3a we see that  $\dot{\sigma}$  is minimized close to the deterministic fixed points (this is a design feature of CMOS devices, in order to minimize the static power consumption, which is however not zero). Also, we see that  $\dot{\sigma}$  is actually lower to the left of the fixed point at  $x \simeq 3$  than to the right. However, in Fig. 3b we see that the speed  $dx/dt$  at which the fixed point is approached is also lower to the left side, resulting in a larger total entropy production  $\int_0^t dt' \dot{\sigma}(x_{t'}) = \int_{x_0}^{x_t} dx \dot{\sigma}(x)/(dx/dt)$ , and a looser bound in Fig. 2b, c.

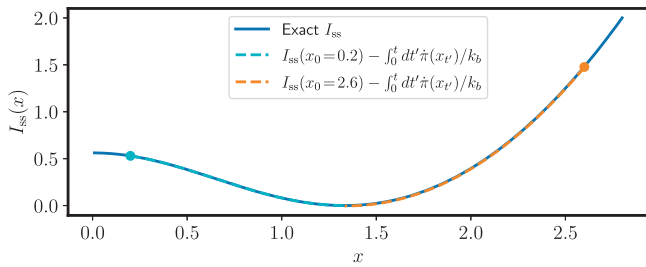
### Tightening the bound

Our main result in Eq. (18) allows to obtain information about the steady state fluctuations by just measuring the physical entropy production along deterministic trajectories. However, if we consider the mathematical problem of bounding the steady state fluctuations given the transition rates  $\lambda_\rho(\mathbf{n})$ , then the full power of our result is achieved by considering a ‘coarse-grained’ entropy production that is in general a lower bound to the actual physical one, as we now explain. In first place we notice that different sets of transition rates  $\{\lambda_\rho(\mathbf{n})\}$  might lead to the same master equation (Eq. (5)) and consequently the same steady state and emerging deterministic dynamics (Eq. (15)), while giving rise to different entropy production rates (Eq. (7)). This is due to the fact that the entropy production depends on how a given stochastic dynamics is split into thermodynamically distinct processes, each of them satisfying a different LDB condition. In particular, the entropy production is not invariant under coarse-graining of the transition rates<sup>23,37</sup>. Specifically, consider that we lump together all transitions going from state  $\mathbf{n}$  to state  $\mathbf{n} + \tilde{\Delta}_\rho$  into a single transition with rate

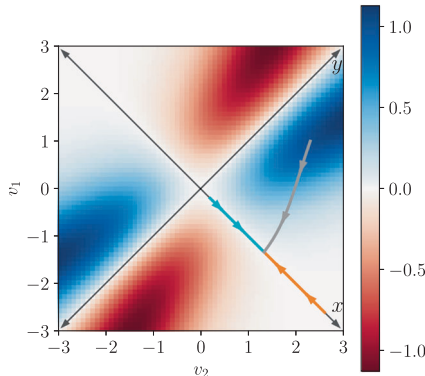
$$\tilde{\lambda}_\rho(\mathbf{n}) \equiv \sum_{\rho/\tilde{\Delta}_\rho = \tilde{\Delta}_\rho} \lambda_\rho(\mathbf{n}). \quad (25)$$

In this way, the jump vectors  $\{\tilde{\Delta}_\rho\}$  associated to the coarse-grained rates  $\tilde{\lambda}_\rho(\mathbf{n})$  are all distinct ( $\tilde{\Delta}_\rho \neq \tilde{\Delta}_{\rho'}$  if  $\rho \neq \rho'$ ). It is easy to check that rates  $\{\tilde{\lambda}_\rho(\mathbf{n})\}$  and  $\{\lambda_\rho(\mathbf{n})\}$  lead to the same master equation. However, the entropy production rate  $\tilde{\Pi}$  corresponding to the rates  $\{\lambda_\rho(\mathbf{n})\}$  is always a lower bound of the original one<sup>23</sup>:

$$\tilde{\Pi} = \frac{k_b}{2} \sum_{\rho, \tilde{\rho}} (\tilde{j}_\rho(\mathbf{n}) - \tilde{j}_{-\rho}(\mathbf{n} + \tilde{\Delta}_\rho)) \log \frac{\tilde{j}_\rho(\mathbf{n})}{\tilde{j}_{-\rho}(\mathbf{n} + \tilde{\Delta}_\rho)} \leq \dot{\Sigma} \quad (26)$$



**Fig. 4 | Coarse-grained bound for the CMOS memory.** Illustration of the bound to  $I_{ss}(x_t)$  from Eq. (18) in terms of the entropy production rate  $\dot{n}$  for the cyan and orange trajectories in Fig. 1, starting respectively at  $x_0 = 0.2$  and  $x_0 = 2.6$  ( $V_{dd} = 1.4$ ).

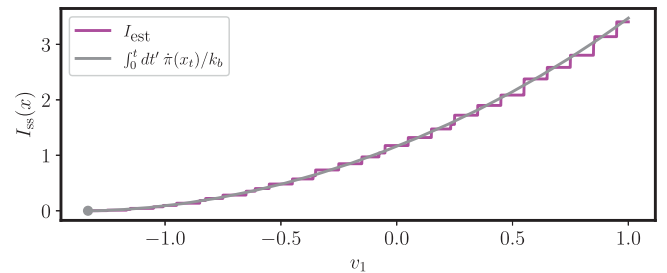


**Fig. 5 | Vorticity and deterministic trajectories.** Vorticity  $f(\mathbf{x})$  of the vector field with components  $\tilde{\sigma}_\rho(\mathbf{x}) \equiv \log(\tilde{\omega}_\rho(\mathbf{x})/\tilde{\omega}_{-\rho}(\mathbf{x}))$ ,  $\rho = 1, 2$ , for the CMOS memory model ( $V_{dd} = 1.4$  and  $v_e = 0.1$ , as in Fig. 1). The orange, cyan and grey lines correspond to the same deterministic trajectories of Fig. 1.

An important property of  $\dot{\Pi}$  is that, whenever the emerging deterministic dynamics has fixed point attractors, its macroscopic version  $\dot{n}(\mathbf{x}_t) \equiv \lim_{\Omega \rightarrow \infty} \dot{\Pi}/\Omega = \sum_{\rho > 0} (\tilde{\omega}_\rho(\mathbf{x}_t) - \tilde{\omega}_{-\rho}(\mathbf{x}_t)) \log(\tilde{\omega}_\rho(\mathbf{x}_t)/\tilde{\omega}_{-\rho}(\mathbf{x}_t))$  vanishes at the fixed points (see the Supplementary Note 2 for a detailed discussion). Thus, the bound obtained using the entropy production rate  $\dot{n}$  is free from the divergence as the fixed point is approached. In the case of the CMOS memory, this is shown in Fig. 4 for the two trajectories along the  $x$  axis of Fig. 1. We see that not only the divergence is avoided, but that the bound matches the exact rate function (see below).

In analogy with Eq. (14), whenever the log-ratio  $\tilde{\sigma}_\rho(\mathbf{x}) \equiv \log(\tilde{\omega}_\rho(\mathbf{x})/\tilde{\omega}_{-\rho}(\mathbf{x}))$  of the scaled coarse-grained rates can be expressed as the gradient of a state function  $-\tilde{\phi}(\mathbf{x})$  along the direction  $\tilde{\Delta}_\rho$ , the steady state rate function will be just  $I_{ss}(\mathbf{x}) = \tilde{\phi}(\mathbf{x})$  (up to a constant). In that case, the system can be considered to be at equilibrium at the coarse-grained level. Under some conditions, one can easily test if  $\tilde{\sigma}_\rho(\mathbf{x})$  derives from a gradient by just checking if the generalized curl  $f_{\rho,\rho'}(\mathbf{x}) \equiv \partial_{x_{\rho'}} \tilde{\sigma}_\rho(\mathbf{x}) - \partial_{x_\rho} \tilde{\sigma}_{\rho'}(\mathbf{x})$  vanishes for all  $\rho, \rho'$ , and  $\mathbf{x}$  (see the Supplementary Note 2). If that is the case, then the bound in Eq. (18) in terms of the entropy production rate  $\dot{n}$  is saturated, and one can fully reconstruct the steady-state rate function in terms of the deterministic dynamics.

In the CMOS memory the state space is two dimensional and therefore there is a single curl component  $f(\mathbf{x})$ , or vorticity. In Fig. 5 we show the vorticity  $f(\mathbf{x})$  for the same parameters as in Fig. 1. We see that the model is genuinely non-equilibrium even at the coarse-grained level. However, we also see that the vorticity vanishes at the  $x$  and  $y$  axes, which explains why the bound in Fig. 4 matches the exact rate function. Finally, we study how the bound in terms of the coarse-grained entropy production rate  $\dot{n}$  performs for the gray trajectory in Figs. 1 and 5, that starts in and goes through areas of non-vanishing



**Fig. 6 | Coarse-grained bound for non-zero vorticity.** Comparison of the upper bound to  $I_{ss}(x_0)$  in terms of the coarse-grained entropy production rate  $\dot{n}$  for the points in the gray trajectory in Figs. 1 and 5 with the function  $I_{est}(\mathbf{x}) = -\log(P_{ss}(\mathbf{x}))/\Omega$  obtained from the exact steady state distribution.

vorticity, and thus genuinely out of equilibrium. Since in this case we do not have the exact rate function to compare with, we compare the bound with the estimation of the rate function  $I_{est}(\mathbf{x}) = -\log(P_{ss}(\mathbf{x}))/\Omega$  obtained from the exact steady state distribution.  $I_{est}(\mathbf{x})$  is only defined for the discrete set of voltages  $v_1$  and  $v_2$  that are a multiple of the elementary voltage  $v_e$ , and in the comparison we choose the closest values of  $v_1$  and  $v_2$  to the continuous deterministic trajectory  $\mathbf{x}_t$ , obtaining a stepwise function. The results are shown in Fig. 6. We see that even in this case our bound provides an accurate approximation of the probability distribution. This shows that the emergent second law in terms of the coarse-grained entropy production has the potential to offer a deterministic alternative to Gillespie simulations and spectral methods. As an example, we provide in the Supplementary Note 3 a full reconstruction of the steady state distribution of the CMOS memory.

It is important to note that the bound given by the emergent second law in terms of the mathematical entropy production in Eq. (26) can be applied to any Markov jump process displaying a macroscopic limit, irrespective of whether it represents a thermodynamically consistent stochastic dynamics or not (the only restriction is that for each jump the reverse jump should also be possible). For example, our results can be relevant in stochastic population<sup>38,39</sup> or gene expression<sup>40,41</sup> models.

## Discussion

For systems accepting a description in terms of Markov jump processes, our results unveil a fundamental connection between the deterministic dynamics that emerges in a macroscopic limit and the non-equilibrium fluctuations at steady state. This is given by an inequality that can be interpreted as an emergent second law. In fact, it is a tighter version of the usual second law, that is saturated in the linear response regime. As shown in the Supplementary Note 4, the emergent second law can be alternatively understood as a generalized fluctuation-dissipation relation. The practical value of our result lies in the fact that the probability of non-equilibrium fluctuations is hard to evaluate, while the deterministic dynamics is directly given by standard methods. The corresponding linear response theory, working at the level of the rate function, was shown in<sup>7</sup> to be highly accurate in some model systems, with a regime of validity beyond that of usual linear response theories. Our result can also be employed in combination with numerical or experimental approaches: once normal or moderately rare fluctuations have been sampled and characterized, Eq. (3) can be used to bound the probability of very rare fluctuations, that otherwise would require extremely long simulation times. In addition, the refinement of our result by a coarse-graining procedure leads to novel numerical techniques to compute non-equilibrium distributions, that only rely on the deterministic dynamics and thus offer an alternative to stochastic simulations or spectral methods.

As a final comment, we note that the extensivity assumptions defining the macroscopic limit we have considered are a very natural extension of the usual notion of extensivity in equilibrium thermodynamics. The only additional requirements, besides the extensivity of the free energy associated to each state, are the extensivity of the jump rates between different states, and the intensivity of the work contributions. These natural requirements lead to an extensive nonequilibrium self-information, whose dominant contribution is constrained by our emergent second law, and that at equilibrium reduces to the regular extensive free energy.

## Methods

As explained in the main text, the emergent second law can be obtained from the fact that the adiabatic contribution to the entropy production rate is positive. The proof of this fact can be found in the mentioned references. The details of the model in the example are given in the Supplementary Note 1. The properties of the coarse grained entropy production are discussed in the Supplementary Note 2. In the Supplementary Note 3 we show how the emergent second law can provide an alternative to stochastic simulations or spectral methods. Finally, in the Supplementary Note 4 we make explicit the connection between the emergent second law derived here and fluctuation-dissipation relations.

## Data availability

All the data shown in the figures can be easily reproduced from the given information, and is also available at the Git repository <https://github.com/nfreitas/emergent-second-law.git>.

## Code availability

The scripts used to generate the data shown in the plots are available at the Git repository: <https://github.com/nfreitas/emergent-second-law.git>.

## References

- Goldstein, S., Lebowitz, J. L., Tumulka, R. & Zanghi, N. Canonical typicality. *Phys. Rev. Lett.* **96**, 050403 (2006).
- Popescu, S., Short, A. J. & Winter, A. Entanglement and the foundations of statistical mechanics. *Nat. Phys.* **2**, 754–758 (2006).
- Georgii, H.-O.: Gibbs Measures and Phase Transitions. De Gruyter, Berlin, Germany <https://doi.org/10.1515/9783110250329/html> (2011).
- Einstein, A. Theorie der Opaleszenz von homogenen Flüssigkeiten und Flüssigkeitsgemischen in der Nähe des kritischen Zustandes. *Annalen der Physik* **338**, 1275–1298 (1910).
- Seifert, U. Entropy production along a stochastic trajectory and an integral fluctuation theorem. *Phys. Rev. Lett.* **95**, 040602 (2005).
- Lebowitz, J. L. & Bergmann, P. G. Irreversible gibbsian ensembles. *Ann. Phys.* **1**, 1–23 (1957).
- Lebowitz, J. Stationary nonequilibrium gibbsian ensembles. *Phys. Rev.* **114**, 1192 (1959).
- McLennan Jr, J. A. Statistical mechanics of the steady state. *Phys. Rev.* **115**, 1405 (1959).
- Zubarev, D. N. *Nonequilibrium Statistical Thermodynamics*. (Consultants Bureau New York, 1974)
- Zubarev, D. Nonequilibrium statistical operator as a generalization of gibbs distribution for nonequilibrium case. *Cond. Matt. Phys* **4**, 19 (1994)
- Komatsu, T. S. & Nakagawa, N. Expression for the stationary distribution in nonequilibrium steady states. *Phys. Rev. Lett.* **100**, 030601 (2008).
- Maes, C. & Netočný, K. Rigorous meaning of mclennan ensembles. *J. Mathematical Phys.* **51**, 015219 (2010).
- Colangeli, M., Maes, C. & Wynants, B. A meaningful expansion around detailed balance. *J. Phys. A: Math. Theoretic.* **44**, 095001 (2011).
- Dhar, A., Saito, K. & Hänggi, P. Nonequilibrium density-matrix description of steady-state quantum transport. *Phys. Rev. E* **85**, 011126 (2012).
- Ness, H. Nonequilibrium density matrix for quantum transport: Hershfield approach as a mclennan-zubarev form of the statistical operator. *Phys. Rev. E* **88**, 022121 (2013).
- Gang, H. Lyapounov function and stationary probability distributions. *Zeitschrift für Physik B Condensed Matter* **65**, 103–106 (1986).
- Freitas, N., Falasco, G. & Esposito, M. Linear response in large deviations theory: a method to compute non-equilibrium distributions. *N. J. Phys.* **23**, 093003 (2021).
- Rao, R. & Esposito, M. Conservation laws shape dissipation. *N. J. Phys.* **20**, 023007 (2018).
- Maes, C. Local detailed balance. *SciPost Phys. Lect. Notes* **32**, 1–17 (2021).
- Hatano, T. & Sasa, S.-i. Steady-state thermodynamics of langevin systems. *Phys. Rev. Lett.* **86**, 3463 (2001).
- Speck, T. & Seifert, U. Integral fluctuation theorem for the house-keeping heat. *J. Phys. A: Math. General* **38**, 581 (2005).
- Esposito, M., Harbola, U. & Mukamel, S. Entropy fluctuation theorems in driven open systems: Application to electron counting statistics. *Phys. Rev. E* **76**, 031132 (2007).
- Esposito, M. & Van den Broeck, C. Three faces of the second law. I. Master equation formulation. *Phys. Rev. E* **82**, 011143 (2010).
- Ge, H. & Qian, H. Physical origins of entropy production, free energy dissipation, and their mathematical representations. *Phys. Rev. E* **81**, 051133 (2010).
- Rao, R. & Esposito, M. Detailed fluctuation theorems: A unifying perspective. *Entropy* **20**, 635 (2018).
- Touchette, H. The large deviation approach to statistical mechanics. *Phys. Rep.* **478**, 1–69 (2009).
- Touchette, H., Harris, R. J.: Large deviation approach to nonequilibrium systems. arXiv [1110.5216](https://arxiv.org/abs/1110.5216). <https://doi.org/10.1002/9783527658701.ch11> (2011).
- Ge, H. & Qian, H. Mesoscopic kinetic basis of macroscopic chemical thermodynamics: A mathematical theory. *Phys. Rev. E* **94**, 052150 (2016).
- Rao, R. & Esposito, M. Conservation laws shape dissipation. *N. J. Phys.* **20**, 023007 (2018).
- Herpich, T., Thingna, J. & Esposito, M. Collective Power: Minimal Model for the Thermodynamics of Nonequilibrium Phase Transitions. *Phys. Rev. X* **8**, 031056 (2018).
- Herpich, T., Cossetto, T., Falasco, G. & Esposito, M. Stochastic thermodynamics of all-to-all interacting many-body systems. *N. J. Phys.* **22**, 063005 (2020).
- Bertini, L., De Sole, A., Gabrielli, D., Jona-Lasinio, G. & Landim, C. Macroscopic fluctuation theory. *Rev. Mod. Phys.* **87**, 593–636 (2015).
- Falasco, G., Rao, R. & Esposito, M. Information Thermodynamics of Turing Patterns. *Phys. Rev. Lett.* **121**, 108301 (2018).
- Enaud, C. & Derrida, B. Large Deviation Functional of the Weakly Asymmetric Exclusion Process. *J. Stat. Phys.* **114**, 537–562 (2004).
- Freitas, N., Delvenne, J.-C. & Esposito, M. Stochastic Thermodynamics of Nonlinear Electronic Circuits: A Realistic Framework for Computing Around  $kT$ . *Phys. Rev. X* **11**, 031064 (2021).
- Freitas, N., Proesmans, K. & Esposito, M. Reliability and entropy production in nonequilibrium electronic memories. *Phys. Rev. E* **105**, 034107 (2022).
- Esposito, M. Stochastic thermodynamics under coarse graining. *Phys. Rev. E* **85**, 041125 (2012).
- Assaf, M. & Meerson, B. WKB theory of large deviations in stochastic populations. *J. Phys. A: Math. Theor.* **50**, 263001 (2017).

39. Riccardo, R. & Stanislas, L. Evolutionary dynamics, evolutionary forces, and robustness: A nonequilibrium statistical mechanics perspective. *Proc. Natl. Acad. Sci. USA.* **119**, 2112083119 (2022).
40. Vahid, S. & Swain Peter, S. Analytical distributions for stochastic gene expression. *Proc. Natl. Acad. Sci. USA.* **105**, 17256–17261 (2008).
41. Bokes, P., Borri, A., Palumbo, P. & Singh, A. Mixture distributions in a stochastic gene expression model with delayed feedback: a WKB approximation approach. *J. Math. Biol.* **81**, 343–367 (2020).

## Acknowledgements

We acknowledge funding from the INTER project "TheCirco" (INTER/FNRS/20/15074473) and CORE project "NTEC" (C19/MS/13664907), funded by the Fonds National de la Recherche (FNR, Luxembourg), and from the European Research Council, project NanoThermo (ERC-2015-CoGAgreement No. 681456). The authors applied a Creative Commons Attribution 4.0 International (CC BY 4.0) license to any Author Accepted Manuscript version arising from this submission.

## Author contributions

J.N.F. and M.E. contributed equally to all aspects of this work.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41467-022-32700-7>.

**Correspondence** and requests for materials should be addressed to José Nahuel Freitas or Massimiliano Esposito.

**Peer review information** *Nature Communications* thanks Fabien Pail-lusson, and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022