Carnot efficiency at divergent power output

MATTEO POLETTINI[(a)] and MASSIMILIANO ESPOSITO

Physics and Materials Science Research Unit, University of Luxembourg, Campus Limpertsberg
162a avenue de la Faïencerie, L-1511 Luxembourg, G. D. Luxembourg

Abstract – The widely debated feasibility of thermodynamic machines achieving Carnot efficiency at finite power has been convincingly dismissed. Yet, the common wisdom that efficiency can only be optimal in the limit of infinitely slow processes overlooks the dual scenario of infinitely fast processes. We corroborate that efficient engines at divergent power output are not theoretically impossible, framing our claims within the theory of Stochastic Thermodynamics. We inspect the case of an electronic quantum dot coupled to three particle reservoirs to illustrate the physical rationale.

Introduction. – It is common wisdom that the efficiency $\eta = P_1/P_2$ of a thermodynamic machine can achieve optimal (or "Carnot") efficiency only via quasistatic processes that deliver a vanishing fraction of power output P_1 per power input P_2 —thus making the machine useless for all practical purposes. However, this fact does not appear to be an immediate consequence of the laws of thermodynamics. In fact, the feasibility of useful machines operating at optimal efficiency is an active issue of debate, usually framed either in the linear regime (LR), a theory of thermodynamics that describes systems close to equilibrium, or using Stochastic Thermodynamics (ST). The latter assumes an underlying Markovian dynamics of the microscopic degrees of freedom, and encompasses LR by including fluctuations and response far from equilibrium.

Humphrey *et al.* [1] have been among the first to investigate the possibility of Carnot efficiency with finite forces. In the LR, it has first been argued [2], and then debated [3], that asymmetric Onsager coefficients might enhance efficiency at finite power. Further studies on the feasibility of powerful Carnot efficiency [4,5] pivot on special assumptions, for example currents growing less-than linearly (or even discontinuously) in the corresponding forces [6]. On the other hand, Carnot efficiency might be impossible to achieve even in the reversible limit when there are leakages at the interface between system and environment [7]. More relevant to our own analysis, in the context of ST, Shiraishi *et al.* [8] argued that there is a tradeoff between power and efficiency preventing optimal

efficiency at finite power; similar tradeoffs are observed in the study of the maximum efficiency attainable at finite power output [9–16].

The latter no-go results appear to be the death knell of all efforts towards optimal and useful machines. Is this search doomed then? We argue that the door is still open to an extreme, yet tantalising, possibility: that the efficiency can be optimised in a regime that is dual and opposite to the quasistatic limit, that of infinitely fast processes that provide divergent power, yet delivering entropy to the environment at a slower rate. We frame our arguments both in the LR and in the ST of steady-state machines. The case of the simplest nano-battery charger, viz. a quantum dot (QD) weakly coupled to three electron reservoirs, allows us to illustrate how this limit can be attained by the interplay of strong thermodynamic forces (depths of energy wells) and of fast and slow kinetic parameters (heights of activation barriers).

That achieving powerful optimal efficiency would be a difficult task is already clear from the fact that, being a rational function, the efficiency has singular behaviour. For example, in the quasistatic limit the efficiency goes to 0/0, a fuzzy quantity that can take any value according to how the limit is taken. Hence, it is unreasonable to question what happens "at optimal efficiency", and the investigation should rather focus on how limits are approached in certain scaling scenarios towards singularities. The authors analysed in ref. [17] a Gaussian model of efficiency fluctuations where in a regime dubbed *singular coupling* the efficiency tends to ∞/∞, another fuzzy value. A similar limit for the figure of merit has been considered in

(a)E-mail: matteo.polettini@uni.lu

ref. [18]. Campisi and Fazio [19] noticed that, when the working substance of a quantum Otto cycle consists of N coupled components, in the large-N limit close to a critical point a conspiracy of critical exponents might lead to a super-linear scaling of efficiency *vs.* power. They conclude that "obstacles hindering the realisation of the critical powerful Carnot engines appear to be of technological nature, rather than fundamental". Another argument by Shiraishi [20] against the attainability of optimal efficiency with finite thermodynamic forces holds provided transition rates are not singular. In a model of an information machine, Bauer *et al.* [21] found that Carnot efficiency can be achieved for finite cycle times, at infinite precision. Seifert [22] has shown that, beyond the LR, there cannot be a universal bound for efficiency at maximum power smaller than Carnot, and that for strong driving the efficiency can be optimised. Raz *et al.* [23] ideated engines where maximum power and efficiency are attained at fast driving. Lee and Park observed that Carnot efficiency can be reached at divergent currents in a model of a Feynman ratchet [24]; Johnson devised a Carnot cycle near the critical point in the phase space of a charged black hole, showing that it becomes powerful and efficient at large charge and low pressure and temperature [25]. Hence, singular behaviour needs to be inspected more closely.

Setup and linear regime. – To stage our proposition in a general framework, we hereby scale all currents in entropic units of Boltzmann's constant k_B per time, in such a way that the total power delivered to the environment (dissipation rate) ultimately satisfies the Second Law of thermodynamics:

$$\sigma := P_2 - P_1 \geq 0. \qquad (1)$$

In this setup the Carnot efficiency is scaled to unity. Setting *exactly* $\eta = P_1/P_2 \equiv 1$, the total dissipation rate vanishes $\sigma = 0$, yet the power input and output can be finite. Hence, the Second Law alone bears no consequence on finite power at optimal efficiency. While the relation between the Second Law and limits on efficiency is debated [19,26,27], let us notice here that we have not yet formulated a constitutive theory describing the behaviour of P_1 and P_2. As our constitutive theories we will consider LR and ST, but one cannot exclude scenarios where other effective theories might come into play, *e.g.*, when the working substance is close to a critical point [6].

The First Law enters the scene when power is resolved into conserved units of energy (thermodynamic forces $\vec{F}$), and of velocities of their carriers (thermodynamic currents $\vec{J}$):

$$P_1 = -F_1 J_1, \qquad P_2 = F_2 J_2 \geq 0. \qquad (2)$$

The negative sign for P_1 highlights that J_1 is expected to perform work *against* the corresponding force, which is the ultimate purpose of machines at all scales: to lift weights against gravity, to transduce molecules across membranes against osmotic pressure, etc.

We now need a constitutive theory for the relation between currents and forces. The simplest theory is LR, prescribing the constitutive relations $\vec{J} = \mathbb{L}\vec{F}$ with positive-semidefinite response matrix $\mathbb{L} = (L_{ij})_{i,j}$, on the assumption that $L_{12}L_{21} \geq 0$. The efficiency can be expressed as [28]

$$\eta_{\mathrm{LR}} = -\frac{1 + \vartheta\phi\sqrt{1-\zeta}}{\phi^2 + \phi/\vartheta\sqrt{1-\zeta}} \qquad (3)$$

in terms of only three adimensional parameters, namely

$$\vartheta := \sqrt{\frac{L_{12}}{L_{21}}}, \quad \zeta := \frac{\det L}{L_{11}L_{22}}, \quad \phi := \frac{F_2}{F_1}\sqrt{\frac{L_{22}}{L_{11}}}. \qquad (4)$$

This compact form is convenient for studying the functional behaviour of the efficiency. The first two parameters are "structural", as they characterize the apparatus: the first measures the violation of Onsager's symmetry; the second is related to the so-called figure of merit [29]. The third parameter is "contingent", meaning that it depends explicitly on the applied forces, the "knobs" that an ideal observer handles (the condition $P_2 \geq 0$ imposes $\phi \leq -\vartheta^{-1}\sqrt{1-\zeta} \vee \phi \geq 0$). Optimising the efficiency with respect to the latter parameter ϕ we obtain

$$\eta_{\mathrm{LR}}^{\star} := \sup_{\phi}\eta_{\mathrm{LR}} = \vartheta^2\frac{1-\sqrt{\zeta}}{1+\sqrt{\zeta}}, \qquad (5)$$

reached at $\phi^{\star} = -(1+\sqrt{\zeta})/\vartheta\sqrt{1-\zeta}$ within the above-mentioned interval.

For the sake of generality, so far we allowed for asymmetric response coefficients, that apparently can be exploited to manipulate the efficiency at will. However, the quantification of dissipation in models with odd dynamical variables is debated. In the context of the ST of autonomous machines with two terminals (input/output) symmetry of response coefficients is a structural property. To obtain non-symmetric response coefficients, one needs to employ to a larger multi-terminal model and then require that the extra currents vanish [3,30]. An analysis of the effect of these "stalling currents" [31] indicates that Carnot efficiency cannot be achieved. However, asymmetric Onsager coefficients can be achieved in time-periodic machines [27]. Let us here assume that time reversal symmetry holds at a fundamental level and set $\vartheta = 1$ in the following.

Optimal efficiency then only depends on parameter ζ. From eq. (5) it is clear that Carnot efficiency can only be achieved when $\zeta \to 0$ *after* ϕ has been optimised. In ref. [17] we showed that the limit in ζ and the limit in ϕ do not commute, and therefore one can approach the fuzzy limit of vanishing power input/output with any value of the efficiency, even the completely "dud" machine with $\eta = -1$. This shows how delicate efficiency optimisation is and in what sense 0/0 is a fuzzy value. A vanishing ζ is related to a high figure of merit. It is often assumed that the so-called tight-coupling limit $\zeta \to 0$ only occurs when

the determinant of the linear-response matrix vanishes, $\det \mathbb{L} \to 0$. However, $\zeta \to 0$ is more generally achieved when the determinant is much smaller than the product of the diagonal response coefficients $L_{11}L_{22}$, a limit that was briefly analysed by the Authors in ref. [17].

Far from equilibrium. – We now venture far from equilibrium, turning to steady-state ST. As the simplest possible system we consider two states (l and r) occupied with probability $p_{l/r}$, among which three distinguishable transitions $i = 0, 1, 2$ can occur at rates $w_i^{\pm}$ rates of jumping right or left ($+/-$)

$$l \overset{1}{\underset{2}{\rightleftarrows}} 0 \ r \ . \tag{6}$$

The current from one state to the other $J_i = w_i^+ p_l - w_i^- p_r$ counts the net number of transitions. The two fundamental constraints ruling network thermodynamics [32] are Kirchhoff's Loop and Current Laws. They can be implemented using cycle analysis [33,34], which states that independent contributions to the dissipation rate are described by the set of fundamental cycles

$$l \overset{0}{\underset{2}{\rightleftarrows}} r \ , \qquad \overset{1}{l \underset{0}{\leftarrow}} r \ , \tag{7}$$
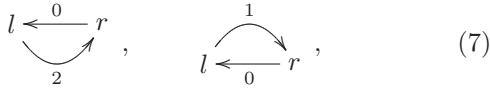
respectively related to the intake of power from reservoir 2 and the outtake of power by reservoir 1, with respect to reservoir 0. Defining $W_{ij} := w_i^+ w_j^-$ and $W_{ij}^{\pm} := W_{ij} \pm W_{ji}$, we obtain for the currents and forces

$$J_1 = \frac{Z_{10}W_{10}^- + Z_{12}W_{12}^-}{Z}, \qquad F_1 = \log\frac{W_{10}^+ + W_{10}^-}{W_{10}^+ - W_{10}^-},$$

$$J_2 = \frac{Z_{20}W_{20}^- - Z_{12}W_{12}^-}{Z}, \qquad F_2 = \log\frac{W_{20}^+ + W_{20}^-}{W_{20}^+ - W_{20}^-}, \tag{8}$$

where $Z := \sum_i (w_i^+ + w_i^-)$ is a normalisation factor associated to the spanning-tree expression of the steady-state probability of being either left or right [34,35]. The coefficients Z_{ij}, to be defined later for the sake of greater generality, are unity for the present system, $Z_{ij} = 1$. Finally the efficiency reads

$$\eta = \frac{(Z_{12}W_{12}^- + Z_{10}W_{10}^-)F_1}{(Z_{12}W_{12}^- - Z_{20}W_{20}^-)F_2}. \tag{9}$$

Notice that (for $Z_{ij} = 1$) the efficiency only depends on five dimensional parameters, which can be made into four independent adimensional parameters, two more than symmetric LR, reflecting the fact that far-from-equilibrium time-symmetric dynamical properties of the system play a crucial role [36]. The LR can eventually be recovered when $W_{ij}^- \ll W_{ij}^+$, yielding the response matrix

$$\mathbb{L} = \frac{1}{2Z}\begin{pmatrix} Z_{10}W_{10}^+ + Z_{12}W_{12}^+ & -Z_{12}W_{12}^+ \\ -Z_{12}W_{12}^+ & Z_{20}W_{20}^+ + Z_{12}W_{12}^+ \end{pmatrix}, \tag{10}$$
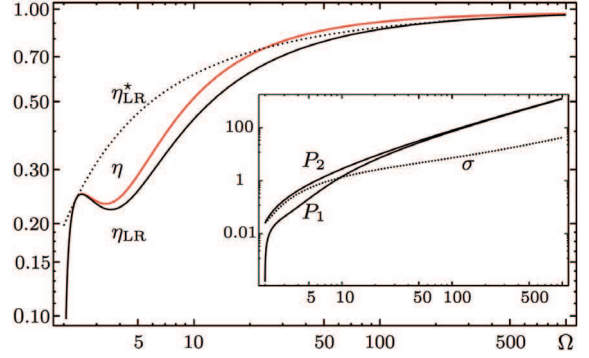

Fig. 1: (Colour online) The efficiency η (upper red curve), its linear-regime approximation η_{LR}, and the maximal linear-regime efficiency $\eta_{\text{LR}}^{\star}$ (dotted) as functions of the scaling parameter, in logarithmic scale, for $\alpha = \delta = 1$, $\beta = 2$, and with choice of rates (at $\Omega = 1$) $w_0^{\pm} = w_2^- = w_1^- = 1$, $w_1^+ = 0.5$, $w_2^+ = 0.6$. Inset: Log-log plot of the power input and output as a function of the scaling parameter. The dotted line is the difference between the two, showing that the dissipation stays order of magnitudes smaller than the power input and output. In fact, by sending $w_1^+/w_2^- \to 1$ (at $\Omega = 1$), one can make the slope of the curve of the dissipation lower at will.

satisfying Onsager's symmetry. The LR efficiency has the simple expression

$$\eta_{\text{LR}} = \frac{(Z_{12}W_{12}^- + Z_{10}W_{10}^-)W_{10}^- W_{20}^+}{(Z_{12}W_{12}^- - Z_{20}W_{20}^-)W_{20}^- W_{10}^+}. \tag{11}$$

But, let us stay far away from equilibrium. We notice that in the limit where $|W_{12}^-|$ is much greater than all other $|W_{ij}^-|$'s, the efficiency goes to

$$\eta \to F_1/F_2. \tag{12}$$

A very large negative $W_{12}^- \to -\infty$ also implies $w_1^+ w_2^- \ll w_1^- w_2^+$, which with a little cycle algebra can be recast as

$$F_2 - F_1 \to \infty. \tag{13}$$

The last two equations are the focal point of our discussion. They highlight that, in order to achieve optimal efficiency, the difference between the affinities should grow large, but at the same time we do not want their ratio to decay to zero. Thus, in our framework we observe that thermodynamic forces cannot be finite for optimal efficiency to be achieved. However, our two focal equations also suggest that optimal efficiency can be achieved in a scaling scenario where both forces grow indefinitely, but proportionally. One scaling scenario that realises this limit is given by $w_0^{\pm} \sim 1$, $w_2^- \sim \Omega^{\alpha}$, $w_2^+ = \Omega^{\alpha+\delta}$, $w_1^- \sim \Omega^{\beta}$, $w_1^+ \sim \Omega^{\beta+\delta}$, with $\beta > \alpha > 1$ and $\delta > 0$. In such scenario, in fig. 1 we plot the efficiency and the power input/output as functions of Ω, clearly showing that the singular limit exists. It can be shown that in this limit the LR parameters $\zeta \to 0$ and $\phi \to 1$. Diagrammatically, it can be interpreted as a prevalence of circulation around cycle
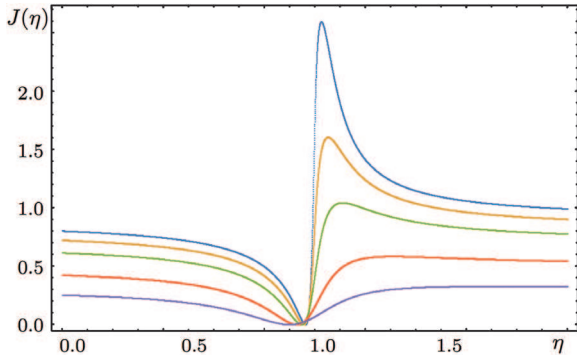
$$\overset{1}{\underset{2}{\rightleftarrows}} \ .$$

Fig. 2: (Colour online) The rate function for the efficiency, for values $\Omega = 10, 20, 50, 100, 200$, from lower to upper curve.

Fluctuations. – So far we have considered the average behaviour of the currents. In the context of ST observables fluctuate. Efficiency as a stochastic variable has first been studied by Verley *et al.* [37], leading to system-specific analysis [38] and to a general characterisation of its universal features [39]. At the steady state, the probability $P_t(\eta)$ of achieving efficiency η at time t is described by the rate function $J(\eta) = -\lim_{t \to \infty} \frac{1}{t} \log P_t(\eta)$. In general, rate functions of stochastic variables are Legendre-dual to time-scaled cumulant generating functions, which are often more tractable. However, the efficiency has no finite moments [17,38], hence its cumulant generating function is nowhere defined. Yet the formula by Verley $J(\eta) = -\min_q \lambda(\eta q, q)$ allows to obtain the efficiency's rate function from the the scaled-cumulant generating function $\lambda(q_1, q_2)$ of the currents, up to a constant. The two-state model is analytically tractable to a great extent. In fig. 2 we plot the rate function of the efficiency for several values of the scaling parameter Ω. The most probable efficiency, given by the infimum of the rate function, slowly creeps to Carnot. The peak, representing the most unlikely value of the efficiency [37], also approaches Carnot. Hence, in the singular limit, Carnot efficiency is at the same time the most and the least probable value. It then follows that in the large-Ω limit small fluctuations might lead to dramatic effects.

Example: quandum dot. – The above two-level system could be physically realised as a single electronic QD acting as an energy filter for electron transport between three reservoirs at the same temperature T, held at different chemical potentials μ_i by the action of external electric fields [1,40–42]. Such device can be seen as a work-to-work conversion nanomachine (a nano-battery charger). The QD can either be (l) occupied at fixed energy E or (r) empty at zero energy. Letting $f(\delta) := (1 + e^\delta)^{-1}$ be the Fermi distribution, the QD is charged by the i-th reservoir at rate $w_i^- = \gamma_i f(\delta_i) =: \gamma_i f_i$ and discharged at rate $w_i^+ = \gamma_i (1 - f_i)$, where γ_i is the tunnelling rate, and $\delta_i := (E - \mu_i)/k_B T$ (see inset of fig. 3 for an illustration and the Appendix for further details). We set $\gamma_0 \equiv 1$, and $\mu_0 \equiv E$, which means that the reference reservoir has
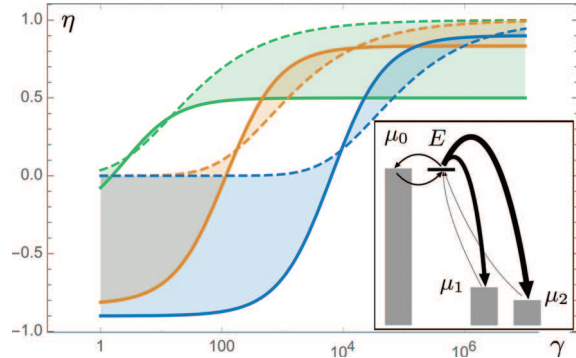

Fig. 3: (Colour online) Inset: illustration of the transitions between the QD and the reservoirs at different chemical potentials. Main frame: thick curves, from left to right, are log-linear plots of the efficiency as a function of the relative tunnelling rate $\gamma = \gamma_1 = \gamma_2$, for $\mu_1 - \mu_2 = k_B T$, and for optimal efficiencies $\eta_{\mathrm{opt}} = .5, .83, .9$ (see eqs. (14), (15)). The corresponding dotted lines (separated by shadings) represent the optimal efficiencies $\eta_{\mathrm{LR}}^\star$ that a machine operating in the LR can achieve for the same values of tunnelling rates.

no bias, $f_0 = 1 - f_0 = 1/2$. In view of eqs. (8), (9) the expressions for the efficiency and the power output are given by

$$\eta = \eta_{\mathrm{opt}} \frac{\gamma_1 \gamma_2 (f_1 - f_2) + \gamma_1 (f_1 - 1/2)}{\gamma_1 \gamma_2 (f_1 - f_2) - \gamma_2 (f_2 - 1/2)}, \quad (14)$$

$$P_1 = \delta_1 \frac{\gamma_1 \gamma_2 (f_1 - f_2) + \gamma_1 (f_1 - 1/2)}{1 + \gamma_1 + \gamma_2}, \quad (15)$$

where $\eta_{\mathrm{opt}} = \delta_1 / \delta_2$. We choose for reservoirs 1 and 2 extremely biased rates in favour of discharging the dot $\delta_2 \gtrsim \delta_1 \gg 0$. Large tunnelling rates then allow to asymptotically reach any given value of the optimal efficiency η_{opt} (see fig. 3), boosting the power output. Comparison with the optimal efficiency of the LR scenario at small chemical potential differences $\epsilon_i = (\mu_i - \mu_0)/k_B T$ indicates that infinitely fast and slow processes might be afflicted by the very same technological issues.

Generalisation and discussion. – Let us now show that the above arguments generalise to more complex models in ST consisting of a large state space, but with only two cycles. The state-space is depicted by a graph

where, *e.g.*, . (16)
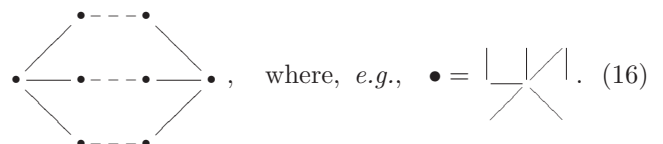
Each bullet in this diagram denotes a so-called *arborescense* of the graph, *i.e.*, subgraphs that do not contain cycles and that do not have sites in common among themselves. Whatever the ramification of the arborescences, indeed the system only has two cycles. Dashed lines denote an arbitrary number of edges connecting the

two bullet sites at their extremes via other bullet sites. The theory of Hill [43] for steady-state currents leads to the expression eq. (9) where

$$
\begin{aligned}
W_{10} &= \\
W_{20} &= \\
W_{12} &=
\end{aligned}
\tag{17}
$$

Here, the diagrammatic expression stands for the product of rates along the oriented links of the diagram, with W_{01}, W_{02}, W_{21} given by the diagrams with inverse orientation. The coefficients Z_{ij} can be obtained as the oriented and rooted spanning tree polynomials in the network where the corresponding cycle is contracted to a unique site [35,44]. This given, the above treatment follows unchanged from eq. (8) to eq. (13).

As regards the bound between the efficiency and power output proven by Shiraishi [8], we notice that the relation is mediated by a dimensional factor, which is needed to compensate for the fact that while the efficiency is adimensional, power has physical units, hence it must be measured against some dynamical property of the system. While a direct application of their result to our setting is not possible, we can retrace their derivation in its fundamental steps[1]. After some work we obtain

$$
\eta(1-\eta) \geq \frac{8}{9}\frac{P_1}{K_2 F_2^2}, \tag{18}
$$

where

$$
K_2 = \frac{1}{Z}(W_{20}^+ + W_{12}^+). \tag{19}
$$

The denominator in eq. (18) can be interpreted as a sort of nonequilibrium LR approximation of the power input, and it is not universal. In particular, in our simple model it is easily seen to be systematically larger than the power output by orders of magnitude. Therefore our treatment is consistent with previous claims of trade-offs between efficiency and finite power. Recently another bound involving power, efficiency and the power's variance has been found by Pietzonka and Seifert [45], which seems to suggest that in an efficient engine with divergent power, the power's variance should also diverge, another fact that is vaguely reminiscent of phase transitions, as already suggested in refs. [17,19].

Our analysis is restrained to two-cycle steady-state (autonomous) models in ST, leaving out more complicated models of multiterminal machines with stalled currents, of multicyclic systems whereby symmetries imply that

---

[1]See appendix A for details. Notice in particular that Shiraishi *et al.* considered weighted currents supported on several edges; this requires several additional passages, including a tilting of the rates, and the application of the Schwarz inequality.

to the same physical process there contribute many configuration cycles [46], and of time-dependent processes. All of the above-mentioned setups rely on the Markov assumption, which in the strong driving limit might fail.

To conclude, our study clearly shows that the quest for optimal efficiency in machines well-described by ST leads in two directions: the reversible and the deeply irreversible limits. While the first is widely known and of limited practical purposes, the second one has been systematically overlooked. We used a simple system to provide a proof-of-concept of the existence of this second limit, and we characterised it physically in terms of the electronic transport across a quantum dot. This research might lead to better design principles for powerful and efficient machines.

$$* * *$$

Appendix A: power/efficiency tradeoff. – We provide a derivation of the power/efficiency tradeoff relation in our context. Let w_{ij} be the rates of a Markov jump process, and p_j the steady-state probability. We follow the derivation of Shiraishi *et al.* [8]. First, as observed by Whitney in ref. [12],

$$
\eta(1-\eta) = \frac{P_1}{P_2}\left(1 - \frac{P_1}{P_2}\right) = \frac{P_1\sigma}{P_2^2}. \tag{A.1}
$$

Then, using Shiraishi's eq. (16) and along the following treatment, the steady-state dissipation rate can be expressed as

$$
\sigma = \sum_{i\neq j} s\left(w_{ij}p_j, w_{ji}p_i\right), \tag{A.2}
$$

where

$$
s(a,b) := a\log\frac{a}{b} + a - b. \tag{A.3}
$$

The following inequality holds:

$$
s(a,b) \geq \frac{8}{9}\frac{(a-b)^2}{a+b}. \tag{A.4}
$$

Hence we obtain

$$
\sigma \geq \frac{8}{9}\sum_{i\neq j}\frac{J_{ij}^2}{K_{ij}}, \tag{A.5}
$$

where

$$
K_{ij} = w_{ij}p_j + w_{ji}p_i. \tag{A.6}
$$

In our example:

$$
\sigma \geq \frac{8}{9}\left(\frac{J_1^2}{K_1} + \frac{J_2^2}{K_2} + \frac{J_0^2}{K_0}\right) \tag{A.7}
$$

and in particular

$$\sigma \geq \frac{8}{9} \frac{P_2^2}{K_2 F_2^2}, \tag{A.8}$$

yielding

$$\eta(1-\eta) \geq \frac{8}{9} \frac{P_1}{K_2 F_2^2}. \tag{A.9}$$

Notice that if instead of throwing all terms away in eq. (A.7) we express $J_0 = J_1 - J_2$, we can obtain a slightly more accurate inequality

$$\eta(1-\eta) \geq P_1 \left(\frac{8}{9} \frac{1}{K_{\text{eff}} F_2^2} - \eta \frac{16}{9} \frac{1}{K_0 F_1 F_2} \right). \tag{A.10}$$

Appendix B: quantum dot. – We provide a more complete treatment of the single Fermionic quantum dot coupled to three reservoirs. Each reservoir charges and discharges the quantum dot at rate

$$w_i^+ = \gamma_i f(\delta_i), \tag{B.1a}$$
$$w_i^+ = \gamma_i [1 - f(\delta_i)], \tag{B.1b}$$

where

$$f(\delta) = \frac{1}{e^\delta + 1} \tag{B.2}$$

is the Fermi distribution and γ_i is the tunnelling rate from the reservoir to the quantum dot. Furthermore,

$$\delta_i = \frac{E - \mu_i}{k_B T}, \tag{B.3}$$

where T is the temperature of all reservoirs, ϵ the energy difference between the charged and uncharged quantum states, and μ_i is the chemical potential of the i-th reservoir. The Fermi distribution satisfies the Kubo-Martin-Schwinger (local detailed balance) condition

$$\frac{f(\delta)}{1 - f(\delta)} = e^{-\delta}. \tag{B.4}$$

For our relevant quantities in the determination of the efficiency we obtain

$$W_{ij}^- = \gamma_i \gamma_j [f(\delta_i) - f(\delta_j)], \tag{B.5a}$$
$$W_{ij}^- = \gamma_i \gamma_j [f(\delta_i) + f(\delta_j) - 2f(\delta_i)f(\delta_j)], \tag{B.5b}$$

and for the thermodynamic forces

$$F_i = \delta_0 - \delta_i = \beta(\mu_i - \mu_0), \tag{B.6}$$

yielding the efficiency

$$\eta = \frac{\gamma_1 \gamma_2 [f(\delta_1) - f(\delta_2)] + \gamma_1 \gamma_0 [f(\delta_1) - f(\delta_0)]}{\gamma_1 \gamma_2 [f(\delta_1) - f(\delta_2)] - \gamma_2 \gamma_0 [f(\delta_2) - f(\delta_0)]} \cdot \frac{\delta_1 - \delta_0}{\delta_2 - \delta_0}. \tag{B.7}$$

Remarkably, the normalisation factor Z simplifies to

$$Z = \gamma_0 + \gamma_1 + \gamma_2. \tag{B.8}$$

The power input and output are then given by

$$P_1 = \gamma_0(\mu_1 - \mu_0) \frac{\gamma_1 \gamma_2 [f(\delta_1) - f(\delta_2)] + \gamma_1 [f(\delta_1) - f(\delta_0)]}{1 + \gamma_1 + \gamma_2}, \tag{B.9a}$$

$$P_2 = \gamma_0(\mu_2 - \mu_0) \frac{\gamma_1 \gamma_2 [f(\delta_1) - f(\delta_2)] - \gamma_2 [f(\delta_2) - f(\delta_0)]}{1 + \gamma_1 + \gamma_2}. \tag{B.9b}$$

In the main text, we further simplified the analysis by setting the timescale $\gamma_0 = 1$, by assuming that the tunnelling rates of the other reservoirs are $\gamma_1 = \gamma_2 = \gamma$, by setting $\delta_0 = 0$ (that is, $\mu_0 = E$) and by setting $\mu_1 - \mu_2 = k_B T$ (that is, $\delta := \delta_1 = \delta_2 - 1$). We then obtain the function plotted in fig. 3:

$$\eta(\gamma, \delta) = \frac{\gamma [f(\delta) - f(\delta + 1)] + [f(\delta) - 1/2]}{\gamma [f(\delta) - f(\delta + 1)] - [f(\delta + 1) - 1/2]} \cdot \frac{\delta}{\delta + 1}. \tag{B.10}$$

The linear regime expression used to plot the maximum theoretical efficiency in fig. 3 is obtained by assuming small thermodynamic forces, *i.e.* by expanding $\delta_i = \delta_0 + \epsilon_i$. Using $f(\delta_i) = f(\delta_0) + f'(\delta_0)\epsilon_i$ one obtains the LR expression for the efficiency

$$\eta = \frac{\gamma_1 \gamma_2 (\epsilon_1 - \epsilon_2) + \gamma_1 \epsilon_1}{\gamma_1 \gamma_2 (\epsilon_1 - \epsilon_2) - \gamma_2 \epsilon_2} \cdot \frac{\epsilon_1}{\epsilon_2}. \tag{B.11}$$

The function $\eta_{\text{LR}}^\star$ can be obtained using the procedure described in the main text.

REFERENCES

[1] HUMPHREY T. E., NEWBURY R., TAYLOR R. P. and LINKE H., *Phys. Rev. Lett.*, **89** (2002) 116801.
[2] BENENTI G., SAITO K. and CASATI G., *Phys. Rev. Lett.*, **106** (2011) 230602.
[3] YAMAMOTO K., ENTIN-WOHLMAN O., AHARONY A. and HATANO N., *Phys. Rev. B*, **94** (2016) 121402.
[4] LEGGIO B. and ANTEZZA M., *Phys. Rev. E*, **93** (2016) 022122.
[5] PONMURUGAN M., arXiv:1604.01912.
[6] KONING J. and INDEKEU J. O., *Eur. Phys. J. B*, **89** (2016) 248.
[7] GONZÁLEZ J. O., ALONSO D. and PALAO J. P., *Entropy*, **18** (2016) 166.
[8] SHIRAISHI N., SAITO K. and TASAKI H., *Phys. Rev. Lett.*, **117** (2016) 190601.
[9] RYABOV A. and HOLUBEC V., *Phys. Rev. E*, **93** (2016) 050101.
[10] HOLUBEC V. and RYABOV A., *J. Stat. Mech.: Theory Exp.* (2016) 073204.
[11] ALLAHVERDYAN A. E., HOVHANNISYAN K. V., MELKIKH A. V. and GEVORKIAN S. G., *Phys. Rev. Lett.*, **111** (2013) 050601.
[12] WHITNEY R. S., *Phys. Rev. Lett.*, **112** (2014) 130601.
[13] BRANDNER K. and SEIFERT U., *Phys. Rev. E*, **91** (2015) 012121.
[14] PROESMANS K., CLEUREN B. and VAN DEN BROECK C., *Phys. Rev. Lett.*, **116** (2016) 220601.

[15] BAUER M., BRANDNER K. and SEIFERT U., *Phys. Rev. E*, **93** (2016) 042112.

[16] JIANG J.-H., *Phys. Rev. E*, **90** (2014) 042126.

[17] POLETTINI M., VERLEY G. and ESPOSITO M., *Phys. Rev. Lett.*, **114** (2015) 050601.

[18] BENENTI GIULIANO, CASATI GIULIO and MEJÍA-MONASTERIO CARLOS, *New J. Phys.*, **16** (2014) 015014.

[19] CAMPISI M. and FAZIO R., *Nat. Commun.*, **7** (2016) 11895.

[20] SHIRAISHI N., *Phys. Rev. E*, **92** (2015) 050101.

[21] BAUER M., ABREU D. and SEIFERT U., *J. Physics A: Math. Theor.*, **45** (2012) 162001.

[22] SEIFERT U., *Phys. Rev. Lett.*, **106** (2011) 020601.

[23] RAZ O., SUBAŞI Y. and PUGATCH R., *Phys. Rev. Lett.*, **116** (2016) 160601.

[24] LEE J. S. and PARK H., arXiv:1611.07665.

[25] JOHNSON C. V., arXiv:1703.06119.

[26] BRANDNER K., SAITO K. and SEIFERT U., *Phys. Rev. X*, **5** (2015) 031019.

[27] PROESMANS K. *et al.*, *Phys. Rev. Lett.*, **115** (2015) 090601.

[28] ENTIN-WOHLMAN O., JIANG J.-H. and IMRY Y., *Phys. Rev. E*, **89** (2014) 012123.

[29] BENENTI G., CASATI G., SAITO K. and WHITNEY R. S., arXiv:1608.05595, to be published in *Phys. Rep.* (2017).

[30] BRANDNER K. and SEIFERT U., *New J. Phys.*, **15** (2013) 105003.

[31] ALTANER B., POLETTINI M. and ESPOSITO M., *Phys. Rev. Lett.*, **117** (2016) 180601.

[32] OSTER G., PERELSON A. and KATCHALSKY A., *Nature*, **234** (1971) 393.

[33] POLETTINI M., *Lett. Math. Phys.*, **105** (2015) 89.

[34] SCHNAKENBERG J., *Rev. Mod. Phys.*, **48** (1976) 571.

[35] ZIA R. K. P. and SCHMITTMANN B., *J. Stat. Mech.: Theory Exp.* (2007) P07012.

[36] BAIESI M., MAES C. and WYNANTS B., *Phys. Rev. Lett.*, **103** (2009) 010602.

[37] VERLEY G., ESPOSITO M., WILLAERT T. and VAN DEN BROECK C., *Nat. Commun.*, **5** (2014) 4721.

[38] PROESMANS K., CLEUREN B. and VAN DEN BROECK C., *EPL*, **109** (2015) 20004.

[39] VERLEY G., WILLAERT T. T., VAN DEN BROECK C. and ESPOSITO M., *Phys. Rev. E*, **90** (2014) 052145.

[40] MAHAN G. D. and SOFO J. O., *Proc. Natl. Acad. Sci. U.S.A.*, **93** (1996) 7436.

[41] SVILANS A., LEIJNSE M. and LINKE H., *C. R. Phys.*, **17** (2016) 1096.

[42] ESPOSITO M., KAWAI R., LINDENBERG K. and VAN DEN BROECK C., *Phys. Rev. E*, **81** (2010) 041106.

[43] HILL T. L., *J. Theor. Biol.*, **10** (1966) 442.

[44] POLETTINI M., *Geometric and Combinatorial Aspects of NonEquilibrium Statistical Mechanics*, PhD dissertation thesis, Alma Mater Studiorum Università di Bologna (2012).

[45] PIETZONKA P. and SEIFERT U., *Universal trade-off between power, efficiency and constancy in steady-state heat engines*, arXiv:1705.05817.

[46] POLETTINI M., BULNES CUETARA G. and ESPOSITO M., *Phys. Rev. E*, **94** (2016) 052117.