The efficiency of a heat engine is traditionally defined as the ratio of its average output work over its average input heat. Its highest possible value was discovered by Carnot in 1824 and is a cornerstone concept in thermodynamics. It led to the discovery of the second law and to the definition of the Kelvin temperature scale. Small-scale engines operate in the presence of highly fluctuating input and output energy fluxes. They are therefore much better characterized by fluctuating efficiencies. In this study, using the fluctuation theorem, we identify universal features of efficiency fluctuations. While the standard thermodynamic efficiency is, as expected, the most likely value, we find that the Carnot efficiency is, surprisingly, the least likely in the long time limit. Furthermore, the probability distribution for the efficiency assumes a universal scaling form when operating close-to-equilibrium. We illustrate our results analytically and numerically on two model systems.
The efficiency of a heat engine operating between a hot and cold reservoir is defined by $\eta = - W/Q_h$, where $Q_h$ is the heat extracted from the hot reservoir, $-W$ is the work produced by the engine and $-Q_c$ is the remaining heat dumped into the cold reservoir. Throughout the study, energy fluxes such as work or heat are positive when flowing into the engine. The Carnot efficiency $\eta_C = 1 - T_c/T_h$ is easily found to be the upper bound of the engine efficiency, $\eta \leq \eta_C$, by combining the first and second law of thermodynamics:

$$\Delta E = W + Q_h + Q_c,$$

$$\Delta S_{tot} = \Delta S - Q_h/T_h - Q_c/T_c \geq 0,$$

in which the engine energy change $\Delta E$ and entropy change $\Delta S$ have to be set equal to zero, as the engine returns to its original state after each cycle.

Early on, Maxwell raised questions concerning the validity of the second law at small scales. Szilard discussed similar issues in the context of information processing. The theoretical breakthroughs in fluctuation theorems and stochastic thermodynamics have fully clarified these points and enable a consistent thermodynamic description of small-scale systems operating arbitrarily far-from-equilibrium. These developments are of crucial practical relevance nowadays, as we are able to design machines operating at the nanoscale as well as to study in great detail biological machines transducing energy and processing information at the sub-micron scale. For such engines, fluctuations are ubiquitous and quantities such as work $w$, heat $q$, energy change $\Delta e$, entropy change $\Delta s$ and entropy production $\Delta s_{tot}$ are stochastic and contain a much richer information than their ensemble average values work $W = \langle w \rangle$, heat $Q = \langle q \rangle$, energy change $\Delta E = \langle \Delta e \rangle$, and so on. At the stochastic level, the first law is essentially the same as at the average level, while the second law in the long time limit as compared with the work and heat. Hence, the heat dumped in the cold reservoir is just $-q_c = w + Q_h$ and we can focus on the work and heat variables, $w$ and $q = Q_h$, and their corresponding output and input powers, $-\dot{w} = -w/t$ and $\dot{q} = q/t$. For systems that harbour no long-time correlations, work $\dot{w}$, heat $\dot{q}$ and their ratio $\eta = -\dot{w}/\dot{q} = -w/q$ are expected to converge in the infinite time limit to their most probable values $\langle \dot{w} \rangle / \langle \dot{q} \rangle$ and $\eta \equiv -\langle w/q \rangle$. The latter efficiency is the one predicted by standard thermodynamics, with the Carnot efficiency as upper bound: $\eta \leq \eta_C$. When considering long trajectories, the probability distribution at time $t$ for $w$ and $q$ as well as for $\eta$ are described by the theory of large deviations and assume the following asymptotic form:

$$P_t(w, q) \sim e^{-J(w, q)}/2 \sqrt{2\pi D_t}$$

Intuitively, the decay rate $J$ for a given efficiency is the smallest among all the decays rates $I$ for the input and output powers which reproduce this efficiency. Note that since the minimization over $q$ in equation (6) includes $q = 0$, we find that $J(q) \leq I(0, 0)$. Expressing the entropy production as $\Delta s_{tot} = q(1/T_c - 1/T_h) + w/T_c$, the fluctuation theorem equation (4) can be shown to take the more detailed form:

$$P_{w, q}(w, q) = \exp(-\Delta s_{tot}).$$

Reversible realizations are characterized by $\Delta s_{tot} = 0$ and thus by an efficiency equal to Carnot efficiency $\eta = -w/q = \eta_C$. In this case, $P(w, q) = P(-w, -q)$ implying $I(w, q) = I(-w, -q)$. Hence $I(-\eta_C q, \eta_C)$ is a symmetric function of the input power $q$. Since $I$ is generically a convex function (assuming no phase transitions), the minimum in equation (6) is, for $\eta = \eta_C$, reached in $q = 0$ and thus $J(\eta_C) = I(0, 0)$. Since $J(0) \leq I(0, 0)$, this proves our main result, namely that the Carnot efficiency becomes the least likely when $t \rightarrow \infty$: $J(\eta) \leq J(\eta_C)$ or $P(\eta) \geq P(\eta_C)$.

Summarizing so far, all values of the efficiency are possible in a small-scale engine running for a large but finite time, including those forbidden by the second law at the average level. The probability for an efficiency different from the standard thermodynamic value $\eta$ decreases exponentially with time with the strongest decrease observed for the Carnot efficiency. Therefore, the probability distribution for the efficiency will develop an exponentially pronounced minimum at the Carnot efficiency as one monitors longer operation times. This observation provides a novel way to define the temperature scale. In standard thermodynamics, the Kelvin temperature scale is introduced by the measurement of the Carnot efficiency of a reversible engine, measurement which is in principle unattainable in a finite time. The identification of the least likely efficiency in a system operating away from reversibility provides an alternative measurement of the Kelvin temperature, which does not suffer from this predicament. For a machine designed to operate like an heat engine (‘on average’), the rare events leading to stochastic efficiency for long times as a direct consequence of the fluctuation theorem equation (4). This result holds for engines with finite state space and therefore with bounded energy so that the energy and entropy contributions $\Delta E$ and $\Delta S$ can be neglected in the first and second law in the long time limit as compared with the work and heat. The following inequality

$${\eta_{\text{c}}} = \frac{\Delta S_{\text{tot}}}{\Delta E} \leq \frac{W}{Q_h} = \eta_C$$

was derived in the context of information processing. The theoretical breakthroughs in fluctuation theorems and stochastic thermodynamics have fully clarified these points and enable a consistent thermodynamic description of small-scale systems operating arbitrarily far-from-equilibrium. These developments are of crucial practical relevance nowadays, as we are able to design machines operating at the nanoscale as well as to study in great detail biological machines transducing energy and processing information at the sub-micron scale. For such engines, fluctuations are ubiquitous and quantities such as work $w$, heat $q$, energy change $\Delta e$, entropy change $\Delta s$ and entropy production $\Delta s_{tot}$ are stochastic and contain a much richer information than their ensemble average values work $W = \langle w \rangle$, heat $Q = \langle q \rangle$, energy change $\Delta E = \langle \Delta e \rangle$, and so on. At the stochastic level, the first law is essentially the same as at the average level, while the second law in the long time limit as compared with the work and heat. Hence, the heat dumped in the cold reservoir is just $-q_c = w + Q_h$ and we can focus on the work and heat variables, $w$ and $q = Q_h$, and their corresponding output and input powers, $-\dot{w} = -w/t$ and $\dot{q} = q/t$. For systems that harbour no long-time correlations, work $\dot{w}$, heat $\dot{q}$ and their ratio $\eta = -\dot{w}/\dot{q} = -w/q$ are expected to converge in the infinite time limit to their most probable values $\langle \dot{w} \rangle / \langle \dot{q} \rangle$ and $\eta \equiv -\langle w/q \rangle$. The latter efficiency is the one predicted by standard thermodynamics, with the Carnot efficiency as upper bound: $\eta \leq \eta_C$. When considering long trajectories, the probability distribution at time $t$ for $w$ and $q$ as well as for $\eta$ are described by the theory of large deviations and assume the following asymptotic form:

$$P_t(w, q) \sim e^{-J(w, q)}/2 \sqrt{2\pi D_t}$$

Intuitively, the decay rate $J$ for a given efficiency is the smallest among all the decays rates $I$ for the input and output powers which reproduce this efficiency. Note that since the minimization over $q$ in equation (6) includes $q = 0$, we find that $J(q) \leq I(0, 0)$. Expressing the entropy production as $\Delta s_{tot} = q(1/T_c - 1/T_h) + w/T_c$, the fluctuation theorem equation (4) can be shown to take the more detailed form:

$$P_{w, q}(w, q) = \exp(-\Delta s_{tot}).$$

Reversible realizations are characterized by $\Delta s_{tot} = 0$ and thus by an efficiency equal to Carnot efficiency $\eta = -w/q = \eta_C$. In this case, $P(w, q) = P(-w, -q)$ implying $I(w, q) = I(-w, -q)$. Hence $I(-\eta_C q, \eta_C)$ is a symmetric function of the input power $q$. Since $I$ is generically a convex function (assuming no phase transitions), the minimum in equation (6) is, for $\eta = \eta_C$, reached in $q = 0$ and thus $J(\eta_C) = I(0, 0)$. Since $J(0) \leq I(0, 0)$, this proves our main result, namely that the Carnot efficiency becomes the least likely when $t \rightarrow \infty$: $J(\eta) \leq J(\eta_C)$ or $P(\eta) \geq P(\eta_C)$.

Summarizing so far, all values of the efficiency are possible in a small-scale engine running for a large but finite time, including those forbidden by the second law at the average level. The probability for an efficiency different from the standard thermodynamic value $\eta$ decreases exponentially with time with the strongest decrease observed for the Carnot efficiency. Therefore, the probability distribution for the efficiency will develop an exponentially pronounced minimum at the Carnot efficiency as one monitors longer operation times. This observation provides a novel way to define the temperature scale. In standard thermodynamics, the Kelvin temperature scale is introduced by the measurement of the Carnot efficiency of a reversible engine, measurement which is in principle unattainable in a finite time. The identification of the least likely efficiency in a system operating away from reversibility provides an alternative measurement of the Kelvin temperature, which does not suffer from this predicament. For a machine designed to operate like an heat engine (‘on average’), the rare events leading to stochastic
efficiencies larger than the Carnot efficiency correspond to realizations along which the machine functions as a heat pump, while those with efficiency lower than zero correspond to a dud engine dissipating heat while absorbing work. The above derivation has been illustrated on a heat engine operating in continuous time, but the results remain valid for all other types of machines, such as isothermal energy transducers, heat engines, refrigerators or heat pumps, operating in non-equilibrium steady states or cyclically as long as the driving cycle is invariant under time reversal. The only difference is that the Carnot efficiency has to be replaced by the reversible efficiency. Below, we give an example of an isothermal work-to-work conversion where the reversible efficiency is equal to 1 and thus corresponds to the least likely efficiency.

Beyond these striking general conclusions about the least and most likely efficiency, we proceed to show that the efficiency fluctuations in the close-to-equilibrium regime have a universal scaling form. Our starting point is that close-to-equilibrium, the relevant work $w$ and heat $q$ fluctuations are generically Gaussian. The resulting large deviation function for the efficiency, being the ratio of two correlated Gaussian variables, is found to be:

$$J(\eta) = \frac{1}{2\eta^2 C_{qq} + 2\eta C_{qw} + C_{ww}},$$

where $C_{qw} \equiv \langle qw \rangle - \langle q \rangle \langle w \rangle$, $C_{qw}$ and $C_{qq}$ are the elements of the symmetric covariance matrix. The crucial thermodynamic ingredient is obtained by combining the Gaussian statistics for $w$ and $q$ with the fluctuation theorem.

More precisely, noting that equation (7) has to be valid for all values of $w = tw$ and $q = t\dot{q}$, one finds (cf. methods section):

$$\langle \dot{q} \rangle = \frac{\eta C_{qw} + C_{ww}}{2T_c}, \quad \langle \dot{w} \rangle = \frac{\eta C_{qw} + C_{ww}}{2T_c}.$$

The large deviation function $J$ can therefore be solely expressed in terms of the covariance matrix:

$$J(\eta) = \frac{1}{8T_c} \left[ \eta \left( \frac{\eta C_{qw} + (\eta + \eta_C)C_{ww} + \frac{C_{ww}}{2}}{\eta C_{qw} + 2\eta C_{ww} + C_{ww}} \right) \right]^2.$$

The above explicit expression for $J(\eta)$ is in agreement with the general properties pointed out above, namely its minimum and maximum are reached for $\eta = \bar{\eta}$ and $\eta = \eta_C$, respectively. These are also the two only extrema of the function. Remarkably, the least likely decay rate, that is, the rate at Carnot efficiency, can be rewritten from equations (9,10) solely in terms of the average heat and work: $J(\eta_C) = \left( \frac{\eta_C \langle \dot{q} \rangle + \langle \dot{w} \rangle}{4T_c} \right)^2$. This relation, which ultimately derives from the fluctuation theorem, should be easy to test experimentally. We also note that the two asymptotic values of $J(\eta)$ at $\eta \rightarrow \pm \infty$ coincide, namely $J(\infty) = \left( \frac{\eta C_{qw} + C_{ww}}{8T_c^2 C_{ww}} \right)^2 = \left( \frac{\bar{\eta}}{2C_{ww}} \right)^2$. We used equation (9) for the second equality. Since the covariance matrix is directly related to the Onsager matrix (cf. methods section), this latter can be obtained from measurements of efficiency fluctuations close to equilibrium. In fact, the covariance matrix $C_{ww}, C_{qw}$ and $C_{qq}$ is uniquely specified by the most probable efficiency $\eta_C$, the value of large deviation function at the Carnot efficiency $J(\eta_C)$, and the asymptotic value of the large deviation function $J(\infty)$ (cf. methods section), so that equation (10) can be rewritten as:

$$\frac{J(\eta)}{J(\eta_C)} = \left( \frac{\eta - \eta_C}{\eta - \eta_C} \right)^2.$$

Brownian work-to-work converter. As a first illustration of our main results, we consider the simplest possible model for work-to-work conversion. An overdamped Brownian particle subjected to two constant forces $F_1$ and $F_2$ diffuses on a plane, as illustrated in Fig. 1. $F_1$ is the driving force, allowing the particle to move against an opposing force $F_1$. For a given displacement $x = x(t)$ of the particle (assuming $x(0) = 0$), the work performed by each force is given by $w_1 = F_1 \cdot x$ and $w_2 = F_2 \cdot x$. The corresponding stochastic efficiency is $\eta = -w_1/w_2$. The displacement $x(t)$ is a Gaussian random variable with average $\langle x(t) \rangle = \mu \cdot t$, with $\mu = \frac{F_1 + F_2}{2}$ and $\mu = \frac{F_2}{2}$. The corresponding correlation functions read $C_{11} = 2D||F_1||^2$, $C_{22} = 2D||F_2||^2$, and $C_{12} = 2DF_1 \cdot F_2$. The large deviation of efficiency is given by

$$J(\eta) = \frac{1}{\eta^2 C_{ww} + 2\eta C_{qw} + C_{ww}} \left( \frac{\eta C_{qw} + (\eta + \eta_C)C_{ww} + \frac{C_{ww}}{2}}{\eta C_{qw} + 2\eta C_{ww} + C_{ww}} \right)^2.$$

One immediately verifies that $J(\eta)$ takes its maximum value $J(1) = \mu^2||F_1||^2/4D$ in $\eta = 1$ which is the predicted reversible efficiency for work-to-work conversion. Furthermore, the above mentioned averages and correlations functions obey the relations equation (9) on setting $\eta_C = 1$ and $D = \mu t$ from the Einstein relation. One can thus also rewrite the $J(\eta)$ as in equation (10) or equation (11) (with $\eta_C = 1$).

Photovoltaic device. Our second model is a nano-sized photovoltaic device powered by black-body radiation at temperature $T_b$ (ref. 29). The device is composed of two quantum dots, each with a single-energy level $E_i$ and $E_q (E_q > E_1)$, respectively, cf. Fig. 2. Coulomb repulsion prevents simultaneous occupation by electrons of both quantum dots. Each dot can exchange electrons with its neighbouring electronic lead. Both leads are at the same temperature $T_c$, but at different voltages and therefore at different chemical potentials $\mu_c > \mu_q$. Electron transfers between the two quantum dots are induced either by hot black-body radiation at $T_b$, or by cold thermal phonons at $T_c$. This device operates as an heat engine fuelled by the heat $q = n_{ph} (E_q - E_1)$, where $n_{ph}$ is the number of photons absorbed from the hot blackbody, and producing a positive work output $-w = n_{el} \Delta \mu$, where $n_{el}$ is the number of electrons transferred from left to right lead against the chemical potential gradient $\Delta \mu = \mu_q - \mu_c > 0$. The stochastic efficiency is thus $\eta = -w/q$. The rates describing the Markovian dynamics of the device as well as the large deviation function for the work and heat statistics are discussed in the methods section. The resulting large deviation function for efficiency is plotted in Fig. 3. All the predicted features—the least

---

**Figure 1** | Work-to-work converter for an overdamped Brownian particle diffusing in a plane. The particle is driven by the two forces $F_1$ and $F_2$. The orange line is a specific trajectory ending at position $x$. The stochastic work $w_1$ and $w_2$ are obtained from the scalar products (dashed lines) projecting $x$ on $F_1$ and $F_2$, respectively. The ratio of the lengths obtained from these projections gives the stochastic efficiency for this specific trajectory.
likely value at Carnot efficiency and the universal shape of the large deviation function close-to-equilibrium—are perfectly reproduced.

Discussion

The efficiency of macroscopic thermal machines is the ratio between two averaged quantities, the extracted work and the heat coming from the hot source. One of the momentous discoveries in science, which lead to the formulation of the second law of thermodynamics, is the observation by Carnot that this efficiency has a maximum called Carnot efficiency. Contrary to macroscopic machines, the behaviour of small machines is subjected to strong fluctuations. Their average behaviour thus provides an incomplete description except in the macroscopic limit where fluctuations are typically strongly peaked around the average. In the present study, we introduce the concept of fluctuating efficiency to accurately characterize the performance of small machines and find universal features in its fluctuations. Using the fluctuation theorem, which generalizes the second law at the fluctuating level, we provide an analogue of the Carnot analysis by proving that the Carnot efficiency becomes the least likely efficiency when long measurement times are considered, independently of any details of the machine or of its mode of operation. Furthermore, we show that close-to-equilibrium, the large deviation function of the efficiency fluctuations obeys a universal form parametrized by the Onsager matrix of the engine. Our study suggests a new direct application of the fluctuation theorem, which was previously mostly invoked to measure free energy differences. Since heat and work fluctuations are nowadays measured in a wide variety of systems, we expect that experimental measurements of the fluctuating efficiency will become a valuable tool to characterize the performance of small engines.

Methods

Linear response and fluctuation theorem. For the photoelectric device, the average photon and electron currents, \( N_p \equiv \langle n_p \rangle / t \) and \( N_e \equiv \langle n_e \rangle / t \), read in the linear regime

\[
\tag{13}
\dot{N}_p = L_{ep} \Delta \mu / T_c + L_{pp} \Delta E \Delta \beta,
\]

\[
\tag{14}
\dot{N}_e = L_{ep} \mu / T_c + L_{pp} \Delta E \Delta \beta,
\]

where \( \Delta \beta = 1 / T_c - 1 / T_h > 0 \), \( \Delta E = E_e - E_i > 0 \) and \( L \) is the symmetric Onsager matrix with \( L_{pp} \geq 0 \), \( L_{ee} \geq 0 \) and \( \det L \geq 0 \). The average work and heat per unit time can thus be written as

\[
\tag{15}
W = \Delta \mu N_e = \frac{1}{T_c} \left( L_{ep} \mu^2 + \eta_{ep} L_{pp} \Delta \mu \Delta E \right),
\]

\[
\tag{16}
Q = \Delta E N_p = \frac{1}{T_c} \left( L_{ep} \mu \Delta E + \eta_{ep} L_{pp} \Delta E \Delta \beta \right).
\]

From Green–Kubo relation, the linear response coefficients are related to equilibrium fluctuations by

\[
\tag{17}
L_{op} = \lim_{\Delta t \to \infty} \frac{1}{2 \Delta t} \langle n_p(t) - \langle n_p \rangle \rangle \langle n_q(0) - \langle n_q \rangle \rangle\rangle_{eq}.
\]

This implies that in the long time limit, \( C_{op} = 2 - \Delta \mu \Delta E \) and \( C_{ep} = 2 + \Delta \mu \Delta E \). Proceeding similarly for the other response coefficients, we find \( C_{pp} = 2 - \Delta \mu \Delta E \) and \( C_{pp} = 2 + \Delta \mu \Delta E \). Equations (15–16) thus lead to equation (9) of the results section. These equalities may also be derived using the fluctuation theorem for work and heat in the Gaussian limit. Indeed, using

\[
\tag{18}
I(\dot{w}, \dot{q}) - I(-\dot{w}, -\dot{q}) = -\langle (\dot{w} \dot{q} + \dot{w}) \rangle / T_c,
\]

and the quadratic large deviation function

\[
\tag{19}
I(\dot{w}, \dot{q}) = \frac{\langle \dot{w} \dot{q} \rangle - \langle \dot{w} \dot{q} \rangle_{eq}^2}{2 \det C},
\]

we get

\[
\tag{20}
\det C = 2 T_c \langle \dot{w} \dot{q} \rangle_{eq} = \dot{w} W C_{eq} + \dot{q} Q C_{eq} - C_{eq}(\dot{w} W + \dot{q} Q).
\]

Since this relation must hold true for any values of \( \dot{w} \) and \( \dot{q} \), we obtain

\[
\tag{21}
\det C = \left[ \begin{array}{cc} C_{eq} & -C_{eq} \\ -C_{eq} & C_{eq} \end{array} \right] \left[ \begin{array}{c} W \\ Q \end{array} \right] = \left[ \begin{array}{c} 1/2 T_c \eta_{eq} (2 T_c) \end{array} \right],
\]

which reproduces the expected result when solved for \( W \) and \( Q \).
according to
\[
\begin{bmatrix}
    g_i(0; \gamma; \lambda) \\
    g_i^j(0; \gamma; \lambda) \\
    g_i(\tau; \gamma; \lambda)
\end{bmatrix} =
\begin{bmatrix}
    -k_0 - k_0 & k_0 & k_0 e^{-\Delta \beta \lambda} \\
    k_0 & -k_0 - k_0 & k_0 e^{-\Delta \beta \lambda} \\
    k_0 & k_0 & -k_0 - k_0
\end{bmatrix}
\begin{bmatrix}
    g_i(0; \gamma; \lambda) \\
    g_i^j(0; \gamma; \lambda) \\
    g_i(\tau; \gamma; \lambda)
\end{bmatrix}.
\]
\tag{22}

When \( \lambda = \gamma = 0 \), equation (22) becomes a Markovian master equation for the probability \( P_i = g_i(j, 0, 0) \) to find the device in state \( j \) at time \( t \). The rates \( k_0 \) denote the probability per unit time to jump from state \( j \) to \( i \). Introducing the Fermi–Dirac distribution \( f(x) \equiv (e^{x-1} + 1) \) and the Bose–Einstein distribution \( b(x) \equiv (1/e^{x-1} - 1) \), they are defined by
\[
\begin{align*}
    k_{0i} &= \Gamma f \left( \frac{E_i - \mu_i}{T} \right), \\
    k_{0i} &= \Gamma f \left( \frac{E_i - \mu_i}{T} \right), \\
    k_{0i}^e &= \Gamma b \left( \frac{E_i - \mu_i}{T} \right),
\end{align*}
\tag{23}
\]
and \( k_0 = k_0^b + k_0^f \), where \( c = c_i \) denotes the cold and hot reservoir and \( \Gamma \) is the coupling strength with the various reservoirs as depicted in Fig. 2 of the result section. For long times \( t \), the work and heat-generating function is dominated by the highest eigenvalue \( \phi(i; \lambda) \) of the rate matrix in equation (22)
\[
\begin{equation}
\langle \varepsilon^{\alpha + i \eta} \rangle = \sum_{i,j,l} g_i(\alpha, \gamma; \lambda) \approx e^{\phi(i; \lambda)}.
\tag{24}
\end{equation}
\]

The latter can be calculated analytically. The corresponding large deviation function is obtained by the Legendre transform \( \Gamma(w, q) = \max decided to study efficiency

References
1. Maxwell, J. C. Tait’s “Thermodynamics”. Nature 17, 278–280 (1878).
2. Szilard, L. On the decrease of entropy in a thermodynamic system by the intervention of intelligent beings. Phys. Z. 53, 840–856 (1929).
3. Jarzynski, C. Equations and inequalities: irreversibility and the second law of thermodynamics at the nanoscale. Annu. Rev. Condens. Matter 2, 329–351 (2011).
4. Sevick, E. M., Prabhakar, R., Williams, S. R. & Searles, D. J. Fluctuation theorems. Annu. Rev. Phys. Chem. 59, 603–638 (2008).
5. Campisi, M., Hanggi, P. & Talkner, P. Colloquium: quantum fluctuation relations: foundations and applications. Rev. Mod. Phys. 83, 771–791 (2011).
6. Seifert, U. Stochastic thermodynamics, fluctuation theorems and molecular machines. Rep. Prog. Phys. 75, 126001 (2012).
7. Van den Broeck, C. & Esposito, M. Ensemble and trajectory thermodynamics: a brief introduction. Phys. A 360, 1016–1045 (2006).
8. Esposito, M., Harbola, U. & Mukamel, S. Nonequilibrium fluctuations, fluctuation theorems, and counting statistics in quantum systems. Rev. Mod. Phys. 81, 1665–1702 (2009).
9. Liphardt, J., Dumont, S., Smith, B. T., Tinoco Jr. & Bustamante, C. Equilibrium information from nonequilibrium measurements in an experimental test of Jarzynski’s equality. Science 296, 1832–1835 (2002).
10. Bialek, V., Speck, T., Helden, I., Seifert, U. & Bechinger, C. Thermodynamics of a colloidal particle in a time-dependent nonharmonic potential. Phys. Rev. Lett. 96, 070603 (2006).
11. Küng, B. et al. Test of the fluctuation theorem for single-electron transport. J. Appl. Phys. 113, 136507 (2013).
12. Liphardt, J., Dumont, S., Smith, S. B., Tinoco Jr. & Bustamante, C. Time-dependent nonequilibrium fluctuations. Phys. Rev. Lett. 109, 180601 (2012).
13. Bialek, V. & Bechinger, C. Realization of a micrometre-sized stochastic heat engine. Nat. Phys. 8, 143–146 (2012).
14. Moffitt, J. R. et al. Intersubunit coordination in a homomeric ring ATPase. Nature 457, 446–450 (2009).
15. Yau, S., Hoj, H., Yoshida, M., Kinoshita, K. & Itoh, H. Resolution of distinct rotational substeps by submillisecond kinetic analysis of F1-ATPase. Nature 410, 898–904 (2001).
16. Toyabe, S., Sagawa, T., Ueda, M., Muneyuki, E. & Sano, M. Experimental demonstration of information-to-energy conversion and validation of the generalized Jarzynski equality. Nat. Phys. 6, 988–992 (2010).
17. Bürk, A. et al. Experimental verification of Landauer’s principle linking information and thermodynamics. Nature 483, 187–189 (2012).
18. Allemany, A., Mossa, A., Junier, J. & Ritort, F. Experimental free-energy measurements of kinetic molecular states using fluctuation theorems. Nat. Phys. 8, 688–694 (2012).
19. Collin, D. et al. Verification of the Crooks fluctuation theorem and recovery of RNA folding free energies. Nature 437, 231–234 (2005).
20. Koski, J. V. et al. Distribution of entropy production in a single-electron box. Nat. Phys. 9, 644–648 (2013).
21. Küng, B. et al. Irreversibility on the level of single-electron tunneling. Phys. Rev. X 2, 011001 (2012).
22. Caliberto, S., Imparato, A., Naert, A. & Tanase, M. Heat flux and entropy produced by thermal fluctuations. Phys. Rev. Lett. 110, 180601 (2013).
23. Bustamante, C., Liphardt, J. & Ritort, F. The nonequilibrium thermodynamics of small systems. Phys. Today 58, 43–48 (2005).
24. Matthews, J., Battista, P., Sanchez, D., Samuelsson, P. & Linke, H. Experimental verification of reciprocity relations in quantum thermoelectric transport. Preprint at http://arxiv.org/abs/1306.3694 (2013).
25. Touchette, H. The large deviation approach to statistical mechanics. Phys. Rep. 478, 1–69 (2009).
26. García-García, R., Domínguez, D., Lecomte, V. & Kolton, A. B. Unifying approach for fluctuation theorems from joint probability distributions. Phys. Rev. E 82, 030104 (2010).
27. Pytte, R. A., Cavaglia, F., Caroli, E., Porras, A. M. & Saitta, I. B. Experimental test of the fluctuation theorem for single-electron transport. Phys. Rev. B 80, 235122 (2009).
28. Jarzynski, C. Nonequilibrium equality for free energy differences. Phys. Rev. Lett. 78, 2690–2693 (1997).
29. Crooks, G. E. Path ensemble averages in systems driven far from equilibrium. Phys. Rev. E 61, 2361–2366 (2000).

Acknowledgements
G.V. acknowledges insightful comments from Andreas Engel. This work was supported by the National Research Fund, Luxembourg under Project No. ENR/A11/02 and INTER/FWO/13/09 and also benefited from support by the ESF network 'Exploring the Physics of Small Devices'.

Author contributions
G.V. explicitly derived the central results of the study and suggested to study efficiency fluctuations using large deviation theory. T.W. was involved in preliminary studies of model systems which led to this work. C.V.d.B. supervised the work at every stage and provided key contributions on the connection between the fluctuation theorem and efficiency fluctuations. M.E. supervised the work at every stage, proposed to study efficiency fluctuations and made key suggestions about the close-to-equilibrium limit.

Additional information
Competing financial interests: The authors declare no competing financial interests.

Reprints and permission information is available online at http://npg.nature.com/reprintsandpermissions/

How to cite this article: Verley, G. et al. The unlikely Carnot efficiency. Nat. Commun. 5:4721 doi: 10.1038/ncomms5721 (2014).