New Journal of Physics
The open access journal at the forefront of physics

Deutsche Physikalische Gesellschaft DPG IOP Institute of Physics

PAPER • OPEN ACCESS

Nonequilibrium fluctuations in quantum heat engines: theory, example, and possible solid state experiments

To cite this article: Michele Campisi *et al* 2015 *New J. Phys.* **17** 035012

New Journal of Physics
The open access journal at the forefront of physics

Deutsche Physikalische Gesellschaft DPG
IOP Institute of Physics

Published in partnership with: Deutsche Physikalische Gesellschaft and the Institute of Physics

CrossMark

OPEN ACCESS

RECEIVED
2 December 2014

REVISED
14 February 2015

ACCEPTED FOR PUBLICATION
23 February 2015

PUBLISHED
20 March 2015

PAPER

Nonequilibrium fluctuations in quantum heat engines: theory, example, and possible solid state experiments

Michele Campisi[1], **Jukka Pekola**[2] **and Rosario Fazio**[1]

[1] NEST, Scuola Normale Superiore and Istituto Nanoscienze-CNR, I-56126 Pisa, Italy
[2] Low Temperature Laboratory, Department of Applied Physics, Aalto University School of Science, P.O. Box 13500, 00076 Aalto, Finland

E-mail: michele.campisi@sns.it, jukka.pekola@aalto.fi and rosario.fazio@sns.it

Keywords: quantum fluctuations, thermal machine, superconducting devices

Abstract

We study stochastic energetic exchanges in quantum heat engines. Due to microreversibility, these obey a fluctuation relation, called the heat engine fluctuation relation, which implies the Carnot bound: no machine can have an efficiency greater than Carnot's efficiency. The stochastic thermodynamics of a quantum heat engine (including the joint statistics of heat and work and the statistics of efficiency) are illustrated by means of an optimal two-qubit heat engine, where each qubit is coupled to a thermal bath and a two-qubit gate determines energy exchanges between the two qubits. We discuss possible solid-state implementations with Cooper-pair boxes and flux qubits, quantum gate operations, and fast calorimetric on-chip measurements of single stochastic events.

1. Introduction

The field of nonequilibrium quantum thermodynamics has gathered a great deal of momentum over the last two decades due to the discovery of a number of exact relations which characterize the response of physical (possibly small) systems to external perturbations, namely, applied mechanical forces or thermodynamic forces (e.g., temperature gradients and chemical potential gradients) [1–3].

Unlike traditional thermodynamics [4], which focuses on macroscopic quantities, fluctuation relations focus on their microscopic, fluctuating, counterparts. To exemplify this, consider the two fundamental objects of thermodynamic investigation, work and heat. A macroscopic thermal engine delivers a certain amount of work while withdrawing a corresponding amount of heat from a hot thermal reservoir. There can be variations in these amounts between different cycles, but typically these fluctuations are negligible. However, as the machine size is scaled down, the work output and heat absorbed will likewise be scaled down. Accordingly, the fluctuations will become more and more relevant. It then becomes useful to investigate the stochastic properties of such fluctuating quantities. Fluctuation relations pose stringent constraints on the statistics of such fluctuating quantities as heat and work due to the symmetries (in particular time-reversal symmetry) characterizing the microscopic motions of atoms and molecules from which they originate.

Fluctuation relations have been reported for both classical and quantum systems [1, 2, 5–7]. In fact, identical fluctuation relations hold regardless of whether the same system is regarded as classical or quantum. Despite their formal identity, classical and quantum fluctuation relations are profoundly different in the way they can be accessed experimentally. Concerning work, for example, although typically one can measure the fluctuating work applied to a classical nano system, e.g., a stretched RNA molecule, by continuously monitoring a displacement x and its conjugate force f (e.g., extension and tension in the molecule) and obtaining the work as $W = -\int f dx$ [8–10], this is typically impossible in a quantum system. In the quantum scenario, the situation is greatly complicated by the invasiveness of the measurement apparatus, which can lead to collapse of the wave function. The prescription accordingly is to measure the energy of the system twice (at the beginning and end of the forcing protocol) by means of two projective measurements and obtain the work as their difference [11–14].

This two-measurement scheme has, however, proved challenging from an experimental point of view [15, 16], so much so that it has been carried out only very recently [17]. This occurrence has triggered the

proposal of a number of alternative methods. One such method proposes to replace the two invasive projections with many less invasive measurements (POVM) carried out on a smaller portion of the system [18]. This method is particularly well suited for studies of transport induced by gradients of temperature and chemical potential [18]. Some experiments already exist which can be explained in terms of these multiple measurements [19, 20]. They take into account the complete counting statistics of electrons transported through a double quantum dot due to an applied chemical potential difference. Note, however, that in those experiments all quantum coherences are suppressed.

Another ingenious method, which is particularly well suited for obtaining the work statistics of a driven system, requires special coupling of the driven system to an ancilla, e.g., a qubit, and replaces the two energy measurements with state tomography of the qubit at the sole final time [21, 22]. This method is a form of Ramsey interferometry and gives experimental access to the characteristic function of work, namely, the Fourier transform of the probability density function (pdf) of work. This has led to the first experimental measurement of quantum work statistics ever performed. It was performed in a liquid-NMR setup and reconstructed the work pdf of a driven two-level system [23]. A proposal for implementing the method with solid-state quantum devices was put forward in [24]. The most promising aspect of this method is that it can be used to asses the work statistics not only of closed systems as in the performed experiment but also of systems which stay in contact with a thermal bath [24].

Roncaglia *et al* [25] have proposed to couple the system to a quantum pointer, e.g., a spin chain. The coupling is engineered so that that a single final projective measurement of the state of the pointer will contain information about the work performed on the system. Like the interferometric method, this method is best suited for the measurement of work. Its experimental realization, however, appears extremely challenging.

In this work we focus on yet another method that has been discussed recently in [26, 27] and that is based on the calorimetric measurement of a photon released and absorbed by thermal reservoirs. This quantum calorimeter is currently under development. The method is well suited for simultaneously measuring both heat and work in a driven quantum system which stays in contact with one or more baths. For this reason it is promising for the experimental study of the stochastic energy exchange of quantum thermal machines.

Since the seminal work of [28], showing how a three-level maser may be understood as a thermal machine, quantum thermal machines have been widely studied in the literature [29–31]; and are still under vigorous investigation [32–45]: see also the recent review [46] and references therein. However, whereas so far the focus has been on the average value of heat and work, here we focus on their fluctuations as well. As recently reported [47], a special form of the fluctuation relation holds for quantum thermal machines. This form implies that no quantum thermal engine can over perform the Carnot efficiency. This universal and exact result was anticipated long ago in [29] but only for those quantum mechanical open systems whose dynamics can be well approximated by a Markovian master equation in Lindblad form.

After revisiting the heat engine fluctuation relation, we introduce a model of thermal engine based on two qubits each coupled to its own reservoir and subject to a unitary gate operation. We identify the regimes when the engine works as a heat engine, a refrigerator, or a heater (dud engine), and we study its full stochastic characteristics, including the pdf of its efficiency. The most intriguing features of the presented machine are (a) that at maximum power it can reach efficiency above the Curzon–Albhorn efficiency and (b) that increasing the speed of its operation increases its power output without affecting its efficiency.

It is important to stress that the engine presented here can be implemented in a real solid-state device and that its stochastic energetic exchanges can be measured using current and soon-to-be-available technology. Hereafter we discuss possible solid-state implementations based on the calorimetric measurement scheme.

2. The heat engine fluctuation relation (HEFR)

Consider a driven bi-partite system:

$$H_S(t) = H_1 + H_2 + V(t) \tag{1}$$

with a factorized initial condition

$$\rho = \frac{e^{-\beta_1 H_1}}{Z_1} \otimes \frac{e^{-\beta_2 H_2}}{Z_2} \tag{2}$$

Without loss of generality we assume throughout this work $\beta_1 \leqslant \beta_2$, i.e., the first subsystem is assumed to be not colder than the second at the initial time. Also we assume that at all times the Hamiltonian is time reversal symmetric [48]. We further assume that the compound system is thermally isolated and that the driving is

turned on at time $t = 0$ and turned off at time $t = \tau$. At these two times simultaneous projective measurements of the energies of both subsystems are performed, giving the results $E^1_{n_1}$, $E^2_{n_2}$ and $E^1_{m_1}$, $E^2_{m_2}$, where $i = 1, 2$ and E^i_k is the kth eigenvalue of subsystem i. According to the quantum exchange fluctuation theorem [18, 49, 50], we have

$$\frac{P\left(\Delta E_1, \Delta E_2\right)}{\widetilde{P}\left(-\Delta E_1, -\Delta E_2\right)} = e^{\beta 1 \Delta E_1 + \beta_2 \Delta E_2} \tag{3}$$

where $\Delta E_i = E^i_{m_i} - E^i_{n_i}$ is the observed energy change in subsystems i, $P(\Delta E_1, \Delta E_2)$ is the joint probability of observing ΔE_1 and ΔE_2, and $\widetilde{P}(-\Delta E_1, -\Delta E_2)$ is the joint probability of observing $-\Delta E_1$ and $-\Delta E_2$ when the reversed driving $V(\tau - t)$ is applied. The driving $V(t)$ injects some amount of energy into the compound system:

$$W = \Delta E_1 + \Delta E_2 \tag{4}$$

which is in fact the work performed by the external driving source to drive the system. Part of this energy, ΔE_1, goes into subsystems 1, and part of it, ΔE_2, goes into subsystem 2. Using the preceding equation to make the change of variable $\Delta E_2 \to W$, we obtain a fluctuation relation for the joint probability of work W and ΔE_1:

$$\frac{P\left(\Delta E_1, W\right)}{\widetilde{P}\left(-\Delta E_1, -W\right)} = e^{\left(\beta_1 - \beta_2\right)\Delta E_1 + \beta_2 W} \tag{5}$$

Multiplying by $\widetilde{P}(-\Delta E_1, -W)$ and integrating over $dW d\Delta E_1$, one obtains the integral form of the fluctuation relation

$$\left\langle e^{\left(\beta_2 - \beta_1\right)\Delta E_1 - \beta_2 W} \right\rangle = 1 \tag{6}$$

Using the Jensen's inequality as usual, one obtains from this

$$\langle W \rangle \geqslant \eta_C \langle \Delta E_1 \rangle \tag{7}$$

where $\eta_C = 1 - \beta_1/\beta_2$ is Carnot's efficiency. The preceding equations hold regardless of the size of the two systems as long as the assumptions introduced are satisfied. In modelling a quantum thermal machine we will consider each subsystem as composed of two parts, namely, a heat reservoir and a small quantum system which constitutes part of the working substance (see figure 1). We will call the small quantum systems the working parts. The driving is applied on the working substance. The received work W is shared between subsystems 1 and 2 as ΔE_1 and ΔE_2. We allow for the possibility of a time-dependence of the couplings between the reservoirs and the working parts, in which case we consider them as part of the time-dependent part of $V(t)$ of the Hamiltonian. This encompasses continuous-mode thermal machines, where the couplings between the working parts and their respective reservoirs are constant in time and non-vanishing, and machines operate in discrete mode (via distinct strokes), where those couplings can be switched on and off during operation. The three-level maser is an example of a continuous-mode engine, whereas Carnot, Otto, diesel engines etc. operate in discrete mode.

The average quantities $\langle \Delta E_1 \rangle$, $\langle \Delta E_2 \rangle$, $\langle W \rangle$ define the operation regime of the machine:

- HEAT ENGINE: $\langle \Delta E_1 \rangle \leqslant 0, \langle \Delta E_2 \rangle \geqslant 0, \langle W \rangle \leqslant 0$
- REFRIGERATOR: $\langle \Delta E_1 \rangle \geqslant 0, \langle \Delta E_2 \rangle \leqslant 0, \langle W \rangle \geqslant 0$
- HEATER: $\langle \Delta E_1 \rangle \leqslant 0, \langle \Delta E_2 \rangle \geqslant 0, \langle W \rangle \geqslant 0$

When the thermal machine works as a heat engine, equation (7) gives

$$\frac{\langle W \rangle}{\langle \Delta E_1 \rangle} \leqslant \eta_C \tag{8}$$

This is the second law of thermodynamics as expressed for a heat engine. Our derivation proves its universality on the basis of the time-reversal symmetric unitary dynamics of the entire system and the initial bi-Gibbsian preparation. In a similar way, when the machine operates as a refrigerator, one finds

$$\frac{-\langle \Delta E_2 \rangle}{\langle W \rangle} \leqslant \frac{1}{\beta_2/\beta_1 - 1} = \eta_C^R \tag{9}$$

Before proceeding it is worth remarking that there is freedom in arranging the position of the border between the two subsystems i.e., to arrange the initial bi-Gibbsian equilibrium. In fact, fluctuation relations for heat engines have been derived previously assuming the working substance is fully included in one of the two subsystems only, say, subsystem 2 [47, 51]. In that case, work W is delivered to subsystem 2, which retains a part ΔE_2 (shared between reservoir 2, $-Q_2$, and working substance ΔU_2) and dumps the other other part $\Delta E_1 = -Q_1$

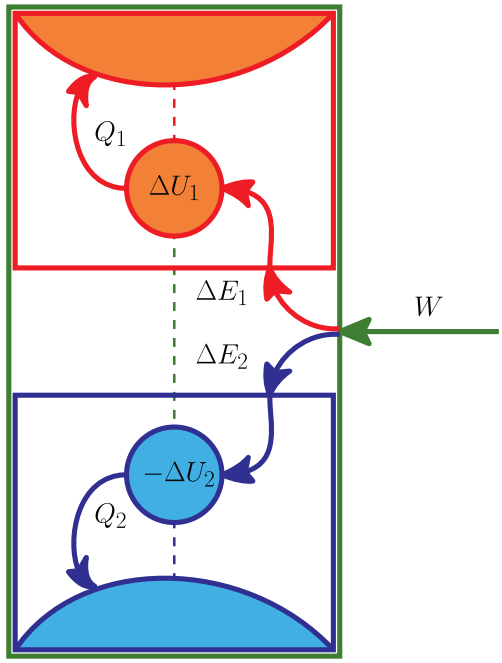

Figure 1. Scheme of a quantum thermal machine. An isolated system (large green rectangle) is driven by an external time-dependent field. The system is composed of two subsystems (red and blue rectangles). Each subsystem is composed of a small quantum system (small circle) and a large system, namely, a thermal reservoir (large circular section). The two small circles form the working substance; we call them working parts. The drive acts on the working substance, thus injecting work W into the entire system. Part of the work, ΔE_1, is delivered to subsystem 1 via working part 1. The rest, $\Delta E_2 = W - \Delta E_1$, is delivered to subsystem 2, via the working part 2. Each working part retains part of the delivered energy ΔU_i and dumps the rest $-Q_i = \Delta E_i - \Delta U_i$ into its reservoir. These energetic exchanges are possible due to possibly time-dependent couplings between the two working parts and between each working part and its reservoir (dashed lines). At the beginning of the driving, each subsystem is at thermal equilibrium with a given temperature T_i.

directly into reservoir 1 as heat. That arrangement is particularly useful for a machine working as a heat engine because equation (7) would read $-\langle Q_1\rangle/\langle W\rangle \leqslant \eta_C$, as in standard thermodynamics books.

Here we adopt instead the scheme in figure 1 because we have in mind an implementation where the coupling of the working parts to the reservoirs is fixed and cannot be manipulated, whereas one can turn the interaction between the working parts on and off. By keeping this coupling off, it is then straightforward to prepare each working part in thermal equilibrium with its own bath. This corresponds to the scenario depicted in figure 1. With our arrangement the average energy $\langle \Delta E_1\rangle$ can be identified with the heat $-\langle Q_1\rangle$ only when the energy $\langle \Delta U_1\rangle$ stored in the working part is null or negligible as compared with $\langle \Delta E_1\rangle$ and $-\langle Q_1\rangle$. This happens when the number of cycles is long and the working substance has a finite energy spectrum. Then $\langle \Delta U_1\rangle$ remains bounded , whereas $\langle \Delta E_1\rangle$ and $-\langle Q_1\rangle$ grow linearly in time. If the condition is met, then $-\langle W\rangle/\langle Q_1\rangle \simeq \langle W\rangle/\langle \Delta E_1\rangle \leqslant \eta_C$. Otherwise, if the condition is not met one can well have the ratio $-\langle Q_1\rangle/\langle W\rangle$ be larger than η_C. This, however, does not have an impact on the second law of thermodynamics stating that a machine *working in a cycle* (implying $\langle \Delta U_1\rangle = \langle \Delta U_2\rangle = 0$) cannot have an efficiency larger than Carnot's efficiency.

It is important to stress that the choice of borders and appropriate associated thermodynamic quantities is the key to obtaining exact transient fluctuation relations, such as in equation (5), that is, fluctuation relations that hold regardless of the time duration of the process under investigation [2, 18, 52, 53]. In the long time limit, steady-state fluctuation relations hold which are independent of the border choice.

3. Optimal two-qubit engine

Our aim is to propose a minimal model of thermodynamic quantum engine which can be implemented and tested experimentally as a solid-state quantum device. The simplest model one can think of is that of a single qubit coupled to two reservoirs at different temperatures. The qubit is driven by an external drive which changes its Hamiltonian in time, for example, by changing its energy spacing $H_{\rm qubit}(t) = \omega(t)\sigma_z/2$. If one has the further ability to couple and decouple the qubit from the two reservoirs one can implement a four-stroke engine, e.g., an Otto cycle. This can be realized, e.g., by interfacing the qubit with the thermal reservoirs by means of band-pass filters, as proposed in [54]. Here we focus instead on the case where the coupling to the reservoirs is fixed in time.

To have a heat engine/refrigerator in continuous mode, a more complex working substance than a mere two-level system is necessary. One needs a working substance that would be able to re-route the energy towards the wanted direction (from the hot bath to the work source and the cold bath for a heat engine; from the cold bath and the work source to the hot bath for a refrigerator). For this reason we introduce a second qubit. Qubit 1 is in contact with the first bath, and qubit 2 is contact with bath 2, as in figure 1. A time-dependent coupling $V(t)$ couples the two qubits for a time-duration $[0,\ \tau]$. The full Hamiltonian is

$$H(t) = H_{q,1} + H_{B,1} + H_{\mathrm{int},1} + H_{q,2} + H_{B,2} + H_{\mathrm{int},2} + V(t) \tag{10}$$

where $H_{B,i}$, $H_{\mathrm{int},i}$, $i = 1, 2$, are the ith bath Hamiltonian and its interaction with qubit i, respectively, and

$$H_{q,i} = \frac{\omega_i}{2}\sigma_i^z \tag{11}$$

is the the ith-qubit Hamiltonian. Here σ_i^z denotes the z Pauli sigma matrix of the ith qubit.

To keep the discussion as simple and intuitive as possible we introduce a useful assumption, namely, that the coupling $V(t)$ is turned on for a time period $[0,\ \tau]$ that is very short compared with the relaxation time of each qubit in its own bath. The effect of the coupling $V(t)$ can accordingly be modelled by a unitary operator U acting in the Hilbert space of the working substance, namely, the two qubits. We will call U the gate operation. Initially, each of the two qubits is in thermal equilibrium with its own bath, i.e., its state is characterized by the density matrix

$$\rho = \frac{e^{-\beta_1 H_{q,1}}}{Z_1} \otimes \frac{e^{-\beta_2 H_{q,2}}}{Z_2} \tag{12}$$

with $Z_i = \mathrm{Tr}\, e^{-\beta_i H_{q,i}} = 2\cosh(\beta_i\omega_i/2)$. The average work injected into the working substance by applying the unitary U is

$$\langle W\rangle = \mathrm{Tr}\left(H_{q,1} + H_{q,2}\right)\left(U\rho U^\dagger - \rho\right) \tag{13}$$

and the energy taken by each subsystem is:

$$\langle \Delta E_i\rangle = \mathrm{Tr}\, H_{q,i}\left(U\rho U^\dagger - \rho\right) \tag{14}$$

After the application of the gate U each qubit is allowed to interact with its respective reservoir for a sufficiently long time so as to reach the state of thermal equilibrium. During this thermalization step they will give the heats $-\langle Q_i\rangle = \langle \Delta E_i\rangle$ to the baths.

We are interested in the unitary that outputs the most work per cycle. Therefore we have searched for the unitary that maximizes $\langle W\rangle$. We have pursued this task by parametrizing a 4×4 unitary by means of 15 angles as discussed in [55] and performing a maximization over the corresponding 15-dimensional space. Numerics clearly indicate that maximum work output is achieved by means of the complex swap unitaries, reading in the $\{|+,+\rangle, |+,-\rangle, |-,+\rangle, |-,-\rangle\}$ basis:

$$U = \begin{pmatrix} e^{i\phi_1} & 0 & 0 & 0 \\ 0 & 0 & e^{i\phi_2} & 0 \\ 0 & e^{i\phi_3} & 0 & 0 \\ 0 & 0 & 0 & e^{i\phi_4} \end{pmatrix}. \tag{15}$$

With these U's we find

$$\langle \Delta E_1\rangle = -\left(\frac{1}{1+e^{\beta_1\omega_1}} - \frac{1}{1+e^{\beta_2\omega_2}}\right)\omega_1 \tag{16}$$

$$\langle \Delta E_2\rangle = \left(\frac{1}{1+e^{\beta_1\omega_1}} - \frac{1}{1+e^{\beta_2\omega_2}}\right)\omega_2 \tag{17}$$

$$\langle W\rangle = \left(\frac{1}{1+e^{\beta_1\omega_1}} - \frac{1}{1+e^{\beta_2\omega_2}}\right)(\omega_2 - \omega_1) \tag{18}$$

In the following we fix the gate to be any complex swap gate in equation (15). Quite remarkably, the same unitaries also maximize the heat engine efficiency. The thermodynamics of one such device, namely, the iSWAP engine (see equation (44)), are discussed in [33]. Its stochastic thermodynamics have not been studied so far.

3.1. Operation

The operation of the swap machine is dictated by the relative signs of $\langle \Delta E_1\rangle$, $\langle \Delta E_2\rangle$, $\langle W\rangle$. With $\beta_1 \leqslant \beta_2$ the conditions for each mode of operation are

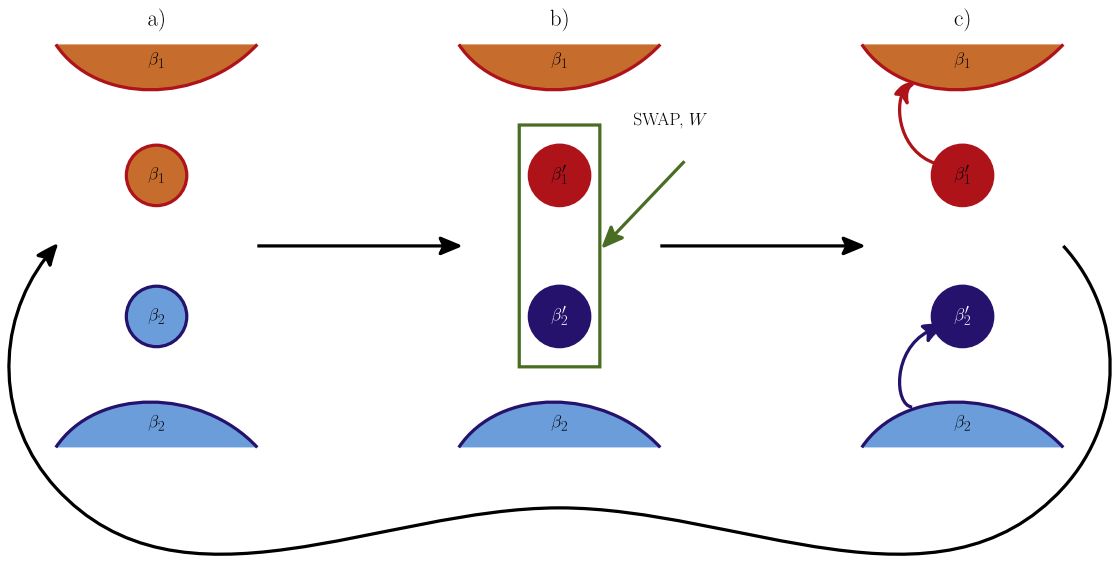

Figure 2. Illustration of the functioning of a swap engine in refrigerator mode. (a) Each qubit is in thermal equilibrium with its respective bath. (b) The instantaneous swap gate is applied, resulting in the injected work W. The hot qubit becomes hotter, whereas the cold qubit becomes colder. (c) Qubit 1 cedes heat to the hot bath. Qubit 2 withdraws heat from the cold bath. The initial equilibrium (a) is reestablished. In heat engine mode the swap cools the hot qubit and heats the cold qubit while outputting work. The sign of the heat and work arrows becomes inverted accordingly.

- HEAT ENGINE: $\frac{\beta_1}{\beta_2} < \frac{\omega_2}{\omega_1} < 1$
- REFRIGERATOR: $0 < \frac{\omega_2}{\omega_1} < \frac{\beta_1}{\beta_2}$
- HEATER: $1 < \frac{\omega_2}{\omega_1}$

The explanation of these conditions is as follows. After the swap-gate operation is performed the two qubits are in the states

$$\rho_1' \propto e^{-\beta_2\omega_2\sigma_z^1/2} = e^{-\beta_1' H_{q,1}} \tag{19}$$

$$\rho_2' \propto e^{-\beta_1\omega_1\sigma_z^2/2} = e^{-\beta_2' H_{q,2}} \tag{20}$$

where

$$\beta_1' = \beta_2\omega_2/\omega_1 \tag{21}$$

$$\beta_2' = \beta_1\omega_1/\omega_2 \tag{22}$$

If $\omega_2/\omega_1 < \beta_1/\beta_2$, then $\beta_1' < \beta_1$ and $\beta_2' > \beta_2$; hence the cold qubit cools down and the hot qubit heats up $\langle\Delta E_1\rangle > 0$, $\langle\Delta E_2\rangle < 0$. Also, since $\omega_2/\omega_1 < \beta_1/\beta_2 < 1$, then $\langle W\rangle > 0$. Hence we have the refrigerator operation. If $\omega_2/\omega_1 > \beta_1/\beta_2$, then $\beta_1' > \beta_1$ and $\beta_2' < \beta_2$; hence the hot qubit cools down and the cold qubit heats up $\langle\Delta E_1\rangle < 0$, $\langle\Delta E_2\rangle > 0$. In this case, depending on the relative size of ω_1 and ω_2, we will have either a heat engine or a heater. Let $u(x) = \mathrm{Tr}\,(\sigma_z/2)e^{-x\sigma_z/2}/\mathrm{Tr}\,e^{-x\sigma_z/2}$. Then $\langle\Delta E_1\rangle = \omega_1[u(\beta_2\omega_2) - u(\beta_1\omega_1)]$ and $\langle\Delta E_2\rangle = \omega_2[u(\beta_1\omega_1) - u(\beta_2\omega_2)] = -(\omega_2/\omega_1)\langle\Delta E_1\rangle$. Accordingly $\langle W\rangle = (1 - \omega_2/\omega_1)\langle\Delta E_1\rangle$. If $\omega_2/\omega_1 < 1$, then $\langle W\rangle < 0$ and we have the heat engine; otherwise, the dud engine. Figure 2 shows a cartoon of the operation of machine in the refrigerator mode.

3.2. Efficiency

For the heat engine operation, it is $0 < \frac{\omega_2}{\omega_1} < \frac{\beta_1}{\beta_2}$; hence:

$$\eta = \frac{-\langle W\rangle}{-\langle\Delta E_1\rangle} = 1 - \frac{\omega_2}{\omega_1} \leqslant 1 - \frac{\beta_1}{\beta_2} = \eta_C \tag{23}$$

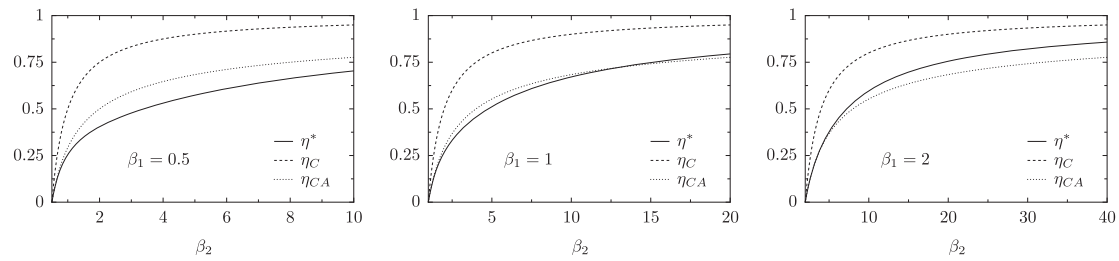

Figure 3. Efficiency at maximum power as a function of β_2 for various values of β_1.

For a refrigerator:

$$\eta^F = \frac{-\langle \Delta E_2 \rangle}{\langle W \rangle} = \frac{\omega_2}{\omega_1 - \omega_2} = \frac{1}{\omega_1/\omega_2 - 1} \leqslant \frac{1}{\beta_1/\beta_2 - 1} = \eta_C^F \tag{24}$$

because for the refrigerator $\frac{\beta_1}{\beta_2} < \frac{\omega_2}{\omega_1} < 1$.

Note that for the engine to function the two qubits must have different energy spacings ω_i; otherwise, the work intake (output) will be exactly null. Note also that the efficiency depends only on the ratio ω_2/ω_1 and not on the temperatures β_1, β_2. This is a peculiar feature of the swap unitary. In the following we will focus on heat engine operation.

3.3. Efficiency at maximum power

Given the two temperatures T_1, T_2 the maximal efficiency, i.e., Carnot's efficiency, is reached when $\omega_2/\omega_1 \to \beta_1/\beta_2$. In this regime, however, the work tends toward zero (see equation (18)). It is interesting that here the power at Carnot efficiency is zero, as with standard stroke engines, but not because of slow operation.

On the other hand, given the two temperatures T_1, T_2 one can find the value of ω_1 and ω_2 for which the power output, $-\langle W \rangle$, is maximum. This can be achieved by maximizing the work output in equation (18). The maximum depends indeed only on the ratio $\Omega = \omega_2/\omega_1$. This can be best seen by setting ω_1 as the unit of energy so that $\omega_1 = 1$ and all energies are measured as multiples of ω_1. With these units:

$$\langle W \rangle = \left(\frac{1}{1 + e^{\beta_1}} - \frac{1}{1 + e^{\beta_2 \Omega}} \right) (\Omega - 1) \tag{25}$$

We denote the value of Ω for which $-\langle W \rangle$ is maximum at given β_1, β_2 as $\Omega^\star(\beta_1, \beta_2)$. The corresponding efficiency, namely, the efficiency at maximum power is

$$\eta^\star\left(\beta_1, \beta_2\right) = 1 - \Omega^\star\left(\beta_1, \beta_2\right) \tag{26}$$

For example, the value of $\Omega^\star$ is $\Omega^\star = 0.83$ for $k_B T_1 = 3/2$, $k_B T_2 = 1$ (in units of ω_1, as previously explained). The corresponding efficiency at maximum power is $\eta^\star \simeq 0.17$.

Figure 3 shows the maximum power efficiency $\eta^\star(\beta_1, \beta_2)$ as a function of β_2 for various fixed values of β_1. The figure also reports the corresponding Carnot efficiency and the Curzon–Albhorn efficiency [56]:

$$\eta_{\mathrm{CA}} = 1 - \sqrt{\frac{\beta_1}{\beta_2}} \tag{27}$$

The figure shows that the the maximum power efficiency can be both larger and smaller than the Curzon–Albhorn efficiency. However, for sufficiently low β_1 (hotter hot reservoir), $\eta^\star < \eta_{\mathrm{CA}}$, whereas, for sufficiently high β_1 (colder hot reservoir), $\eta^\star > \eta_{\mathrm{CA}}$; that is, at very low temperature the efficiency at maximum power is above the Curzon–Albhorn efficiency.

We have performed an analysis of the maximum power efficiency $\eta^\star$ for $\beta_1 \simeq \beta_2$, i.e., in the low η_C limit. In accordance with linear response theory we expect the linear coefficient of the expansion to match the value 1/2 [57]. Since our engine is not a thermoelectric engine (work is provided by time-dependent pulses, rather than by a DC electric potential difference) and does not have left–right symmetry (for it to output some work the energy spacings of the two qubits, ω_1 and ω_2, should be different), we do not expect the value 1/8 for the quadratic coefficient, predicted in the cases [58], to be obeyed. The results of the low η_C analysis, reported in figure 4, corroborate these expectations. The figure presents plots of $\eta^\star/\eta_C$ for various values of β_2 as a function of η_C. The plots clearly show that

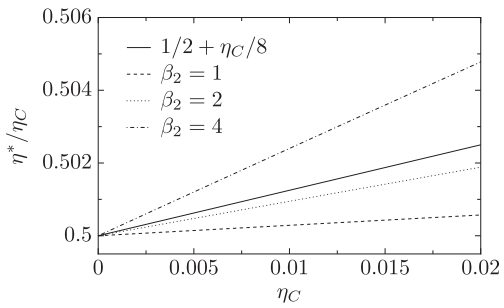

Figure 4. Low η_C behaviour of $\eta^\star$. The plot shows that $\eta^\star$ behaves as in equation (28).

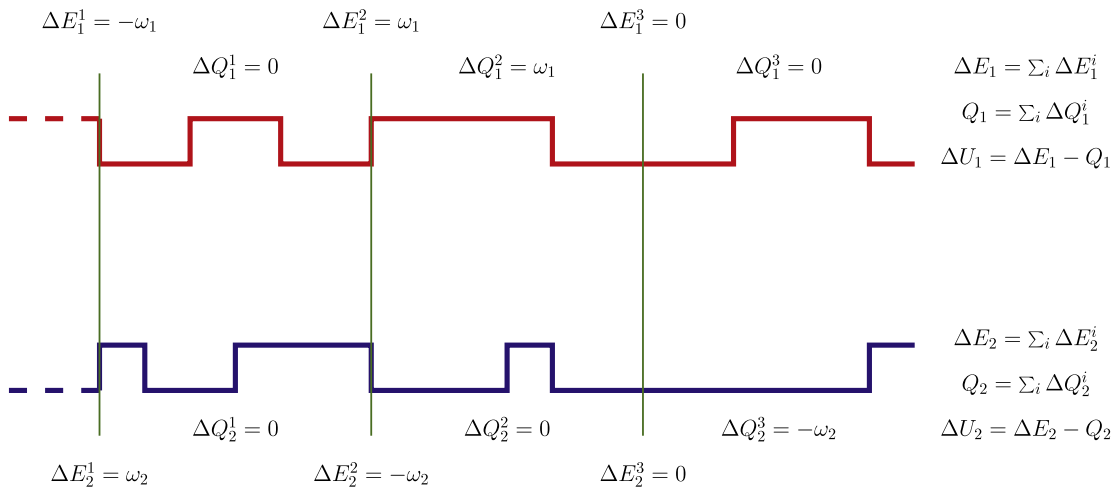

Figure 5. Typical quantum trajectory of the working substance of the swap engine. The top trajectory is for qubit 1. The bottom trajectory is for qubit 2. The green vertical lines indicate times when the instantaneous swap gate is applied. Transitions occurring at these times are due to the work done by the external work source. Transitions occurring between the swap pulses signal heat exchanges with the heat reservoirs.

$$\eta^\star \simeq \frac{\eta_c}{2} + f\left(\beta_2\right)\eta_c^2 + O\left(\eta_C^3\right); \tag{28}$$

that is, the linear coefficient 1/2 is obeyed , whereas the quadratic coefficient is a function $f\left(\beta_2\right)$ whose value may differ from 1/8.

4. Modelling: quantum jumps

The foregoing analysis based on the simplified assumption of unitary gate followed by thermalization allowed us to make predictions about the average work and heats that go into the two reservoirs. It does not suffice, however, for the full stochastic characterization of the engine. To achieve that we need to model the dynamics of the thermalization. We assume then that the effect of the thermal environment on each qubit can be modelled by means of a master equation of Lindblad form:

$$\dot{\rho}_i = -i\left[H_{q,i}, \rho\right] + \mathcal{L}_i\rho \tag{29}$$

where

$$\mathcal{L}_i\rho = \gamma\left(n_i + 1\right)D\left[\sigma_i\right]\rho + \gamma n_i D\left[\sigma_i^\dagger\right]\rho_i, \qquad i = 1, 2 \tag{30}$$

$$n_i = \frac{1}{e^{\beta_i\omega_i} - 1} \tag{31}$$

$$D\left[c\right]\rho = c\rho c^\dagger - \frac{1}{2}c^\dagger c\rho - \frac{1}{2}\rho c^\dagger c \tag{32}$$

and $\sigma_i = \sigma_i^x + i\sigma_i^y$ is the annihilation operator for the qubit i; and $\sigma_i^\dagger$, its adjoint, is the creation operator.

To obtain the statistics of energy exchanges with the bath during the thermalization step, we proceed to unravel the master equation [59], as proposed in [60] and [27]. This results in a stochastic differential equation in the Hilbert space of each qubit:

$$d\left|\psi_i\right\rangle = -iG_i\left(\left|\psi_i\right\rangle\right)dt + \left(\frac{\sigma_i\left|\psi_i\right\rangle}{\|\sigma_i\left|\psi_i\right\rangle\|} - \left|\psi_i\right\rangle\right)dN_i^+ + \left(\frac{\sigma_i^\dagger\left|\psi_i\right\rangle}{\|\sigma_i^\dagger\left|\psi_i\right\rangle\|} - \left|\psi_i\right\rangle\right)dN_i^- \tag{33}$$

The deterministic part is given by

$$G_i\left(\left|\psi_i\right\rangle\right) = H_{q,i}^{\mathrm{eff}}\left|\psi_i\right\rangle + \frac{i}{2}\gamma\left(n_i+1\right)\|\sigma_i\left|\psi_i\right\rangle\|^2\left|\psi_i\right\rangle + \frac{i}{2}\gamma\left(n_i\right)\|\sigma_i^\dagger\left|\psi_i\right\rangle\|^2\left|\psi_i\right\rangle \tag{34}$$

$$H_{q,i}^{\mathrm{eff}} = H_{q,i} - \frac{i}{2}\gamma\left(n_i+1\right)\sigma_i^\dagger\sigma_i - \frac{i}{2}\gamma n_i\sigma_i\sigma_i^\dagger = H_{q,i} - \frac{i}{2}\gamma n_i - i\gamma\sigma_i^\dagger\sigma_i \tag{35}$$

whereas the stochastic Poisson increments have the ensemble expectations

$$E\left(dN_i^+\right) = \gamma\left(n_i+1\right)\|\sigma_i\left|\psi_i\right\rangle\|^2\,dt \tag{36}$$

$$E\left(dN_i^-\right) = \gamma\left(n_i\right)\|\sigma_i^\dagger\left|\psi_i\right\rangle\|^2\,dt \tag{37}$$

The stochastic equations can be solved by means of the Monte Carlo wave function (MCWF) method [61]. In the present case of an undriven single qubit they result in a dichotomic Poisson process governed by the following two rates:

$$\Gamma_i^- = \gamma n_i \tag{38}$$

$$\Gamma_i^+ = \gamma\left(n_i+1\right) \tag{39}$$

depending on whether the qubit is in the down state $|-\rangle$ or up state $|+\rangle$. Note that these rates are detailed balanced:

$$\frac{\Gamma_i^-}{\Gamma_i^+} = e^{-\beta_i\omega_i} \tag{40}$$

Reference [60] has studied the fluctuation relations for such quantum trajectories but for systems being in contact with a single thermal reservoir. Here our working substance, the two qubits, is in contact with two distinct reservoirs. The analysis performed in [60] can, however, be extended to multiple reservoirs. This results in the following fluctuation relation for the probability of a given quantum trajectory γ:

$$\frac{P[\gamma]}{P\left[\tilde{\gamma}\right]} = \exp\left(\beta_1 Q_1[\gamma] + \beta_2 Q_2[\gamma] + \ln\frac{p_a}{p_b}\right) \tag{41}$$

where $\tilde{\gamma}$ is the time reverse of γ. We remark that in the case of multiple reservoirs γ is not only specified by the temporal evolution of the state of the central system (the working substance in our case)—call it χ_t—but also by the succession i_n, indicating which bath (labelled by i) was responsible for each of the N jumps (labelled by n), $\gamma = (\{\chi_t\}, \{i_n\})$. Accordingly the time-reversed trajectory results by requiring that the temporal evolution of the central system state be inverted, and if the nth jump of the forward trajectory γ was caused by the ith bath, so was the last nth jump of the backward trajectory $\tilde{\gamma}$. That is, $\tilde{\gamma}_t = (\{\chi_{T-t}\}, \{i_{N-n}\})$. In our case the trajectory γ has two components $\gamma_t = (\gamma_{1,t}, \gamma_{2,t})$, each specifying the temporal evolution of the state of each qubit. No extra indexes are necessary because all jumps in $\gamma_{1,t}$ are caused by reservoir 1 and all jumps in $\gamma_{2,t}$ are caused by reservoir 2. Accordingly $\tilde{\gamma}_t = (\gamma_{1,\mathcal{T}-t}, \gamma_{2,\mathcal{T}-t})$. The symbol $Q_i[\gamma]$ means the heat ceded to the ith reservoir during the realization of γ. Specifically $Q_i[\gamma] = \int_0^{\mathcal{T}}\omega_i(dN_{i,s}^- - dN_{i,s}^+)ds$. Obviously in our case Q_i is a functional of γ_i only. In equation (41) a, b denote the initial and final state of the trajectory γ, i.e., $\gamma_0 = a$, $\gamma_{\mathcal{T}} = b$, and $p_{a,b}$ are the respective probability that these states are observed. With our choice (12) it is $p_x = \exp\left(-\beta_1 U_1^x - \beta_2 U_2^x\right)/(Z_1 Z_2)$, $x = a, b$. Writing $\Delta U_i = U_i^b - U_i^a$ and using $\Delta U_i = \Delta E_i[\gamma] - Q_i[\gamma]$, we obtain

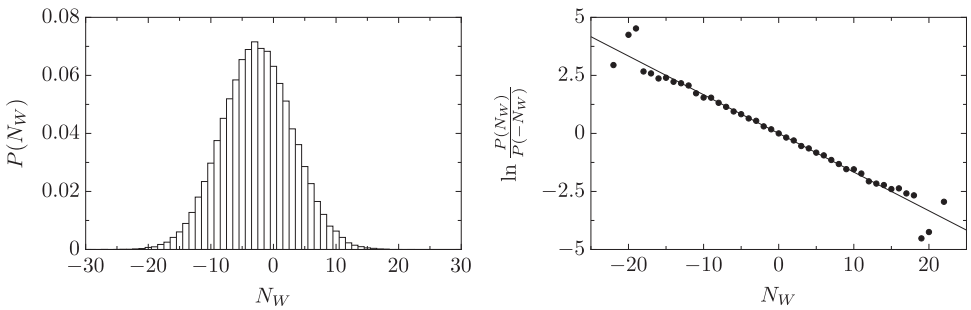

Figure 6. Left panel: probability $P(N_W)$ of work quanta N_W given off by the work source. Right panel: corresponding logarithmic fluctuation ratio $\ln(P(N_W)/P(-N_W))$. Dots: numerics. Solid line: theoretical line $(\beta_1\omega_1 - \beta_2\omega_2)N_W$. The discerepancy at large N_W is ascribed to bad corresponding statistics. The histogram $P(N_W)$ was constructed from a sample of 10^6 trajectories. Here $k_B T_1 = 1.5$, $k_B T_2 = 1$, $\omega_1 = 1$, $\omega_2 = 5/6$, corresponding to heat engine operation. The time between swaps $\tau_2 \simeq 0.65$ was about half the relaxation time τ_{relax} and $N = 100$ swap gates were applied.

$$\frac{P[\gamma]}{P\left[\tilde{\gamma}\right]} = \exp\left(\beta_1 \Delta E_1[\gamma] + \beta_2 \Delta E_2[\gamma]\right) \tag{42}$$

Multiplying by $P[\tilde{\gamma}]\delta(\Delta E_1 - \Delta E_1[\gamma])\delta(\Delta E_2 - \Delta E_2[\gamma])$ and performing a path integral over all trajectories γ, one recovers equation (3). Accordingly all subsequent relations in section 2 are obeyed within our quantum jump modelling.

5. Stochastic thermodynamics of the swap engine

We operate the machine in the following manner. At time $t = 0$ we pick up a state randomly from the initial bi-Gibbsian distribution, equation (12). We apply the complex swap gate, equation (15), and generate the stochastic dynamics of each qubit using the MCWF method until time τ_2, when we apply the complex swap again and let it evolve stochastically until time $2\tau_2$, and so on for a total duration $\mathcal{T} = N\tau_2$. Our assumption is that the swap gate is much faster than the stochastic evolution time: $\tau \ll \tau_2$. Figure 5 shows a sketch of the resulting quantum trajectories of the two qubits, along with the energetic exchanges the various jumps signal.

The first important observation from figure 5 is that any time the energy ΔE_1 is given to subsystem 1, accordingly the energy $\Delta E_2 = -(\omega_2/\omega_1)\Delta E_1$ is taken from subsystem 2. This implies that all trajectories have the same efficiency $\eta = W/\Delta E_1 = (\Delta E_1 + \Delta E_2)/\Delta E_1 = 1 - \omega_2/\omega_1 = \eta$. In other words there are no efficiency fluctuations. This is because the gate swaps the eigenstates of the double qubit without creating superpositions thereof. Hence each swap pulse k deterministically and univocally results in well-defined values of $\Delta E_{1,2}^k$ depending on the state of each qubit before its application. A generic unitary will typically create a superposition of the eigenstates, which can collapse in either the up state or the down state of each qubit with corresponding probability. The value of $\Delta E_{1,2}^k$ would be accordingly not uniquely defined by the state before a generic gate.

The constraint $\Delta E_2 = -(\omega_2/\omega_1)\Delta E_1$ makes it possible to express the heat engine fluctuation relation (5) as a relation for a single variable, say, W. Since W is an integer multiple N_W of $\omega = \omega_1 - \omega_2$, the fluctuation relation can be conveniently expressed as a relation for the probability $P(N_W)$ that N_W of work quanta are given off by the work source. We obtain then

$$\frac{P\left(N_W\right)}{P\left(-N_W\right)} = e^{\left(\beta_1\omega_1 - \beta_2\omega_2\right)N_W} \tag{43}$$

Figure 6 shows $P(N_W)$ and the corresponding logarithmic ratio $\log P(N_W)/P(-N_W)$ for one simulation of our engine. In an experimental realization the probability $P(N_W)$, can be constructed by recording the number and sign of the swaps occurred during each of many realizations in just one of the two qubits.

In figure 7, left panel, we report a plot of the joint probability distribution of heat and work $P(Q_1, W)$.

Note how $W/(\omega_1 - \omega_2)$ differs from Q_1/ω_1 at most by one unit. This is because $\Delta E_1/\omega_1$ differs from Q_1/ω_1 at most by one unit, i.e., one quantum of energy stored in qubit 1 as ΔU_1. Because Q_1 is not exactly equal to ΔE_1, heat efficiency $\eta_Q = -W/Q_1$ has some fluctuations, in contrast to the ΔE-efficiency $\eta = -W/\Delta E_1$. The statistics of η_Q corresponding to the plot in figure 7 are reported in the right panel of figure 7.

Note the very pronounced peak at η. Note also a second peak η^C. We observe that there is a finite probability that η_Q is infinite. Because of the peak at infinity the quantity $\langle\eta_Q\rangle$ is not well defined. As the number of cycles increases the spot in figure 7 left panel drifts and diffuses in the diagonal direction but not in the transverse

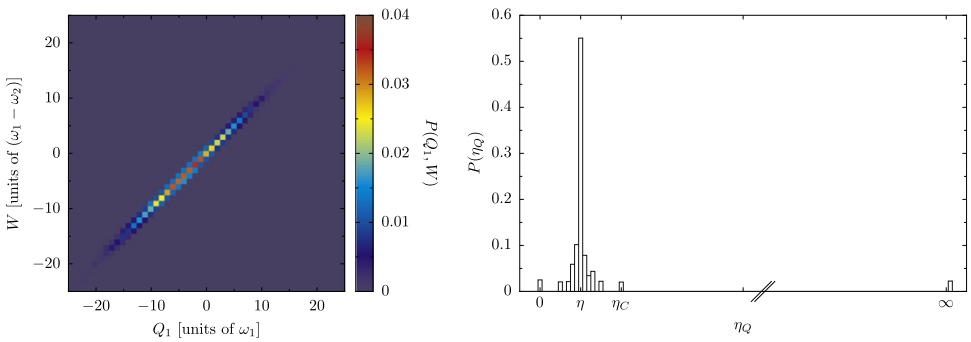

Figure 7. Left panel: joint distribution $p(Q_1, W)$. Right panel: the corresponding Q-efficiency distribution $P(\eta_Q)$. The plots are for $k_B T_1 = 3/2$, $k_B T_2 = 1$, $\omega_1 = 1$, $\omega_2 = 5/6$, corresponding to the heat engine regime. The number of applied pulses is $N = 100$. The sample consists of 10^6 trajectories. The Carnot efficiency is $\eta_C = 1 - 2/3 = 1/3$. The ΔE-efficiency is $\eta = 1 - 5/6 = 1/6 = \eta_C/2$.

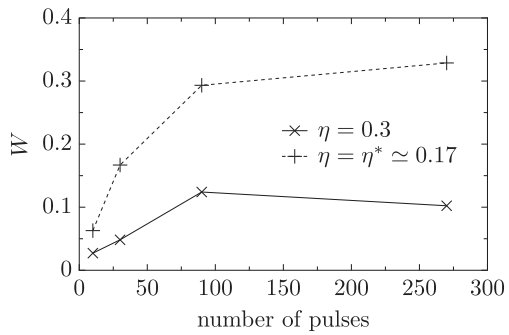

Figure 8. Efficiency and work output as a function of number of applied pulses for a fixed operation time. Here $k_B T_1 = 3/2$, $k_B T_2 = 1$ (in units of ω_1). The operation time of the machine is fixed and equal to $T_{\rm op} = 30\tau_{\rm relax}$ ($\tau_{\rm relax}$ is defined here as the longest among the thermal relaxation times, i.e., $\tau_{\rm relax} = \gamma^{-1} \max((e^{\beta_1\omega_1} - 1, e^{\beta_2\omega_2} - 1))$. At N=10 cycles, the engine has plenty of time to relax to equilibrium, because each pulse is followed by a rest time of $3\tau_{\rm relax}$. By increasing the pulse frequency one can greatly enhance the work output. Solid line: $\omega_1 = 1$, $\omega_2 = 0,7$, corresponding to efficiency $\eta = 0.3$. Dashed line: $\omega_1 = 1$, $\omega_2 \simeq 0.83$, corresponding to efficiency at maximum power $\eta^* = \simeq 0.17$.

direction. A consequence of this is that the peak at η in the efficiency probability increases while all other peaks decay. That is, for long operation time the probability of η_Q coincides with the probability of η as expected[4].

5.1. Increasing the power

As discussed earlier the swap heat engine works at the efficiency $\eta = 1 - \omega_2/\omega_1$ regardless of the power output. This is a great advantage over traditional engines because increasing the power has no cost in terms of reducing the efficiency for our engine. The power of our engine can be increased simply by increasing the frequency of the swap pulses. This is illustrated in figure 8. Figure 8 suggests that the power output saturates at a regime value as the pulse frequency increases. The saturation value gives the maximum power for the given T_1, T_2, ω_1, ω_2. It is important to stress that for too-frequent swap pulses, namely, when their temporal separation τ_2 is on the same order as the temporal duration τ of the swap pulse, our simplifying assumption (namely, that dynamics can be modelled separately by a unitary followed by stochastic relaxation) does not hold any longer. In a real experiment the power output is expected to decay in the range of highly frequent pulses.

6. Solid-state implementation and measurement scheme

The swap engine may be realized by means of solid-state devices, with the working substance consisting of two superconducting qubits [63, 64]. Superconducting qubits are mesoscopic devices (they can be of micrometer size) comprising Josephson junctions, which, when operated below the critical temperature, display a quantum behaviour that is well characterized by just two quantum states. We recall that a Josephson junction is

[4] According to large-deviation theory, all peaks but the most likely fall with an exponential rate, the largest of which is for the Carnot efficiency [62].

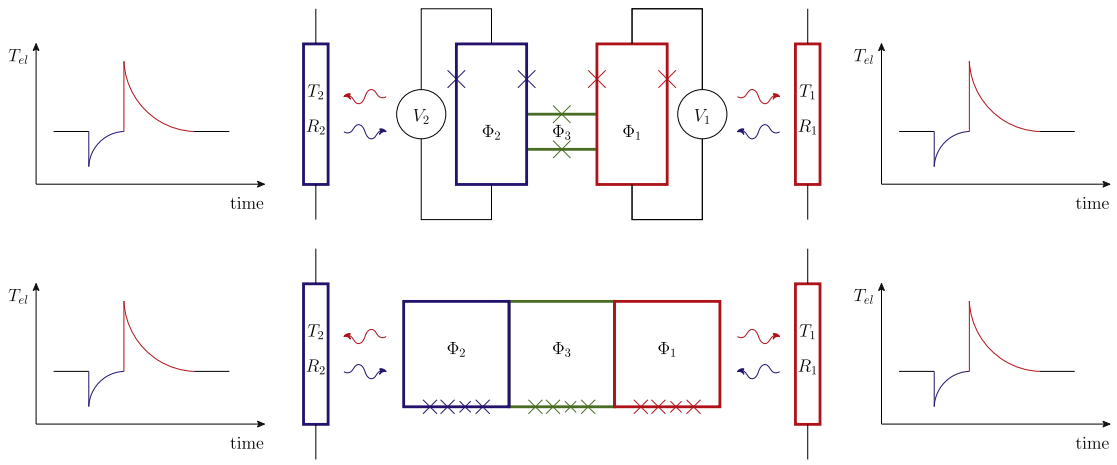

Figure 9. Scheme of two possible experimental setups. Top panel: two Cooper-pair boxes are coupled by means of two Josephson junctions in parallel as in [66]. Bottom panel: two flux qubits are coupled via a third flux qubit as in [69]. Both setups can be used to produce the *i*SWAP gate (equation (44)). Each of the two qubits exchanges photons with a different resistor kept at a given temperature. Two on-chip fast calorimeters detect single-photon emission/absorption in each resistor.

characterized by a capacitance and a nonlinear inductance. Depending on whether the nonlinear inductance or the capacitance dominates, either the phase difference δ of the Cooper-pair condensate wave function across the junction, or its conjugate variable, i.e., the accumulated charge Q, is a good quantum number. In the first case we talk about flux qubits, whereas in the second we talk about charge qubits. Flux qubits are typically realised by means of a superconducting loop interrupted by several Josephson junctions. Accordingly, the two relevant quantum states correspond to clockwise and counterclockwise supercurrents. Charge qubits are realized by means of a so-called Cooper-pair box, namely, a small superconducting island that collects the charge. The island is connected to ground by a gate capacitor in series with a controllable gate potential and by a Josephson junction. In this case the two relevant quantum states correspond to zero or one extra Cooper pair in the island as compared with a reference charge [63, 64]. Many variations of these basic types of Josephson qubits have been developed.

For our implementation proposal we follow the scheme presented in [26] with the necessary modifications and extensions. Reference [26] presents an experimental scheme where a single Cooper-pair box (CPB), namely, a superconducting qubit, is coupled to a resistor at some temperature T. The resistor comprises an electronic system coupled to a phononic one. When a photon is emitted (absorbed) into the resistor, the fast electronic system responds by heating (cooling) abruptly and then relaxing to the thermal equilibrium set by the phononic substrate. A nano calorimeter can then be used to monitor the temperature of the electronic system to detect absorbed/emitted photons. As reported in [65], sufficiently fast and sensitive calorimeters for this purpose are currently under development and should soon be available.

To produce the swap engine, two such CPB + resistor systems should be implemented on the same microchip, which does not seem to pose any particular difficulty. Each resistor is then monitored by an on-chip calorimeter of the type in [65]. At variance with the setup proposed in [26], here the two CPBs have fixed energy gaps, so they exchange photons of well-defined energy $\hbar\omega_i$. This simplifies measurement because each calorimeter does not need to measure the energy of the absorbed/emitted photon but should just detect that a photon has been absorbed/emitted. The gate operation can be implemented by coupling the two CPBs using two tunnel junctions connected in parallel as described in [66]. This allows for the implementation of the *i*SWAP gate, namely the complex swap gate:

$$U = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & i & 0 \\ 0 & i & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{44}$$

Figure 9, top panel, shows the scheme of the implementation.

An alternative implementation uses flux qubits operating at the optimal point, which have much longer coherence and relaxation times compared with CPBs [67, 68]. The switchable coupling is achieved by means of a third qubit sandwiched between the two qubits as demonstrated in [69]. An *i*SWAP gate can be effected by means of microwave driving for a targeted time duration [69]. As demonstrated in [69], the minimum time for a universal gate is, in that set up, about 22 ns, whereas the decoherence and relaxation times are at least 0.2 μs. This

is in agreement with our assumption of fast gate operation, as compared with thermal relaxation. Figure 9, bottom panel, shows this alternative implementation.

By means of calorimetric measurement one can experimentally access the quantum trajectories of the type shown in figure 4. This is achieved in the following way. The calorimeters can detect only the heat quanta ΔQ_k^i ceded to the resistors $i = 1, 2$. If two consecutive emissions (absorptions) are observed to occur in the same bath i, it means that meanwhile a quantum of energy $\Delta E_k^i = +(-)\omega_i$ has been given to (taken from) the ith qubit by the work source. Summing up all the ΔE_k^i one obtains the total energy given to each subsystem, ΔE_1, ΔE_2, and the work $W = \Delta E_1 + \Delta E_2$. Having Q_1, Q_2, ΔE_1, ΔE_2, W one can address the full statistics of energetic exchanges of the engine, and accordingly can check the validity of the fluctuation relations (3), (5). Note that the measurement apparatus can also be employed to check the coincidence of swap-induced jumps in the two qubits, thus quantifying the goodness of the swap operation.

The employment of a flux qubit for implementation of nano coolers was also discussed in [54]. In the work of [54] the working substance is a single flux qubit which is alternatively coupled and decoupled from the two baths. This is attained by embedding each of the two bath-resistors in an LCR circuit, acting as band-pass filters centred at different frequencies ω_1, ω_2. As the qubit- level spacing is switched between these two values, the qubit interacts primarily with one resistor or the other so as to realize an Otto cycle, where interactions with the cold and hot bath occur in alternation and are separated by slow, adiabatic drives. This realize the same average heat and work exchanges as the present engine—hence the same efficiency—with the difference that the present engine works in continuous mode. Heat exchanges with the two baths occur here simultaneously, and no adiabatic drive is employed. This makes this approach more promising in regard to the delivered power.

7. Conclusions

Based on a previous work [47], we have presented a detailed discussion of fluctuation relations for heat and work in quantum heat engines. These fluctuations are illustrated by means of an optimal two-qubit engine working in continuous mode. We studied its full stochastic energetic exchanges, including the statistics of its efficiency. At the average level, this engine achieves the same efficiency as the single-qubit Otto engine of [69] but is expected to deliver higher power due to its continuous mode of operation (no adiabatic sweeps needed; no full thermal relaxation needed). We have presented possible implementations which employ Cooper-pair boxes and flux qubits as working substances, two-qubit quantum gates, and on-chip fast calorimetry for the detection of single exchanged energy quanta. The proposed experiment would allow for the first fully stochastic characterization of a quantum heat engine.

Acknowledgments

This research was supported by the 7th European Community Framework Programme under Grant Agreement Nos. 623085 (MC-IEF-NeQuFlux) (M C), 600645 (IP-SIQS) (R F), 618074 (STREP-TERMIQ) (R F), and 308850 (INFERNOS) (J P); by the Italian Ministry of Education University and Research under Grant Agreement No. MIUR-PRIN-2010LLKJBX (R F); by Academy of Finland (projects 250280 and 272218) (J P); and by COST action MP1209: 'Thermodynamics in the quantum regime'.

References

[1] Esposito M, Harbola U and Mukamel S 2009 *Rev. Mod. Phys.* **81** 1665–702
[2] Campisi M, Hänggi P and Talkner P 2011 *Rev. Mod. Phys.* **83** 771–91
[3] Hänggi P and Talkner P 2015 *Nat. Phys.* **11** 108–10
[4] Fermi E 1956 *Thermodynamics* (New York: Dover)
[5] Seifert U 2012 *Rep. Prog. Phys.* **75** 126001
[6] Jarzynski C 2011 *Annu. Rev. Condes. Matter Phys.* **2** 329–51
[7] Marconi U M B, Puglisi A, Rondoni L and Vulpiani A 2008 *Phys. Rep.* **461** 111–95
[8] Liphardt J, Dumont S, Smith S B, Tinoco I and Bustamante C 2002 *Science* **296** 1832–6
[9] Collin D, Ritort F, Jarzynski C, Smith S B, Tinoco I and Bustamante C 2005 *Nature* **437** 231–4
[10] Douarche F, Ciliberto S, Petrosyan A and Rabbiosi I 2005 *Europhys. Lett.* **70** 593–9
[11] Tasaki H 2000 arXiv:cond-mat/0009244
[12] Kurchan J 2000 arXiv:cond-mat/0007360
[13] Talkner P and Hänggi P 2007 *J. Phys.* A **40** 569–71
[14] Talkner P, Lutz E and Hänggi P 2007 *Phys. Rev.* E **75** 050102
[15] Huber G, Schmidt-Kaler F, Deffner S and Lutz E 2008 *Phys. Rev. Lett.* **101** 070403
[16] Brito F, Rouxinol F, LaHaye M D and Caldeira A O 2014 arXiv:1406.7182
[17] An S, Zhang J N, Um M, Lv D, Lu Y, Zhang J, Yin Z Q, Quan H T and Kim K 2015 *Nat. Phys.* **11** 193–9
[18] Campisi M, Talkner P and Hänggi P 2010 *Phys. Rev. Lett.* **105** 140601

[19] Utsumi Y, Golubev D S, Marthaler M, Saito K, Fujisawa T and Schön G 2010 *Phys. Rev.* B **81** 125331
[20] Küng B, Rössler C, Beck M, Marthaler M, Golubev D S, Utsumi Y, Ihn T and Ensslin K 2012 *Phys. Rev.* X **2** 011001
[21] Dorner R, Clark S R, Heaney L, Fazio R, Goold J and Vedral V 2013 *Phys. Rev. Lett.* **110** 230601
[22] Mazzola L, de Chiara G and Paternostro M 2013 *Phys. Rev. Lett.* **110** 230602
[23] Batalhão T B, Souza A M, Mazzola L, Auccaise R, Sarthour R S, Oliveira I S, Goold J, de Chiara G, Paternostro M and Serra R M 2014 *Phys. Rev. Lett.* **113** 140601
[24] Campisi M, Blattmann R, Kohler S, Zueco D and Hänggi P 2013 *New J. Phys.* **15** 105028
[25] Roncaglia A J, Cerisola F and Paz J P 2014 *Phys. Rev. Lett.* **113** 250601
[26] Pekola J P, Solinas P, Shnirman A and Averin D V 2013 *New J. Phys.* **15** 115006
[27] Hekking F W J and Pekola J P 2013 *Phys. Rev. Lett.* **111** 093602
[28] Scovil H E D and Schulz-DuBois E O 1959 *Phys. Rev. Lett.* **2** 262–3
[29] Alicki R 1979 *J. Phys. A: Math. Gen.* **12** L103
[30] Kosloff R 1984 *J. Chem. Phys.* **80** 1625–31
[31] Geva E and Kosloff R 1992 *J. Chem. Phys.* **96** 3054–67
[32] Quan H, Wang Y, Liu Y X, Sun C and Nori F 2006 *Phys. Rev. Lett.* **97** 180402
[33] Quan H, Liu Y X, Sun C and Nori F 2007 *Phys. Rev.* E **76** 031105
[34] Linden N, Popescu S and Skrzypczyk P 2010 *Phys. Rev. Lett.* **105** 130401
[35] Abah O, Roßnagel J, Jacob G, Deffner S, Schmidt-Kaler F, Singer K and Lutz E 2012 *Phys. Rev. Lett.* **109** 203006
[36] Gelbwaser-Klimovsky D, Alicki R and Kurizki G 2013 *Phys. Rev.* E **87** 012140
[37] Jiang J H, Entin-Wohlman O and Imry Y 2013 *Phys. Rev.* B **87** 205420
[38] Balachandran V, Benenti G and Casati G 2013 *Phys. Rev.* B **87** 165419
[39] Brandner K, Saito K and Seifert U 2013 *Phys. Rev. Lett.* **110** 070603
[40] Uzdin R and Kosloff R 2014 *New J. Phys.* **16** 095003
[41] Zhang K, Bariani F and Meystre P 2014 *Phys. Rev. Lett.* **112** 150602
[42] Correa L A, Palao J P, Alonso D and Adesso G 2014 *Sci. Rep.* **4** 3949
[43] Mazza F, Bosisio R, Benenti G, Giovannetti V, Fazio R and Taddei F 2014 *New J. Phys.* **16** 085001
[44] del Campo A, Goold J and Paternostro M 2014 *Sci. Rep.* **4** 6208
[45] Allahverdyan A E, Hovhannisyan K V, Melkikh A V and Gevorkian S G 2013 *Phys. Rev. Lett.* **111** 050601
[46] Benenti G, Casati G, Prosen T and Saito K 2013 arXiv:1311.4430
[47] Campisi M 2014 *J. Phys. A: Math. Theo.* **47** 245001
[48] Messiah A 1962 *Quantum Mechanics* (Amsterdam: North Holland)
[49] Jarzynski C and Wójcik D K 2004 *Phys. Rev. Lett.* **92** 230602
[50] Andrieux D, Gaspard P, Monnai T and Tasaki S 2009 *New J. Phys.* **11** 043014
[51] Sinitsyn N A 2011 *J. Phys. A: Math. Theo.* **44** 405001
[52] Gaspard P 2013 *New J. Phys.* **15** 115014
[53] Bulnes Cuetara G, Esposito M and Imparato A 2014 *Phys. Rev.* E **89** 052119
[54] Niskanen A O, Nakamura Y and Pekola J P 2007 *Phys. Rev.* B **76** 174523
[55] Hedemann S R 2013 ArXiv e-prints (*Preprint* 1303.5904)
[56] Curzon F L and Ahlborn B 1975 *Am. J. Phys.* **43** 22
[57] van den Broeck C 2005 *Phys. Rev. Lett.* **95** 190602
[58] Esposito M, Lindenberg K and van den Broeck C 2009 *Phys. Rev. Lett.* **102** 130602
[59] Breuer H P and Petruccione F 2002 *The theory of open quantum systems* (Great Clarendon Street: Oxford University Press)
[60] Horowitz J M 2012 *Phys. Rev.* E **85** 031110
[61] Mølmer K, Castin Y and Dalibard J 1993 *J. Opt. Soc. Am.* B **10** 524–38
[62] Verley G, Esposito M, Willaert T and van den Broeck C 2014 *Nat. Commun.* **5** 4721
[63] Makhlin Y, Schön G and Shnirman A 2001 *Rev. Mod. Phys.* **73** 357
[64] Clarke J and Wilhelm F K 2008 *Nature* **453** 1031–42
[65] Gasparinetti S, Viisanen K L, Saira O P, Faivre T, Arzeo M, Meschke M and Pekola J P 2015 *Phys. Rev. Applied* **3** 014007
[66] Echternach P, Williams C P, Dultz S C, Delsing P, Braunstein S and Dowling J P 2001 *Quantum Info. Comput.* **1** 143–50
[67] Chiorescu I, Nakamura Y, Harmans C J P M and Mooij J E 2003 *Science* **299** 1869–71
[68] Mooij J E, Orlando T P, Levitov L, Tian L, van der Wal C H and Lloyd S 1999 *Science* **285** 1036–9
[69] Niskanen A O, Harrabi K, Yoshihara F, Nakamura Y, Lloyd S and Tsai J S 2007 *Science* **316** 723–6