Quantum limit to nonequilibrium heat-engine performance imposed by strong system-reservoir coupling

DOI:
10.1103/PhysRevE.101.052129

Document Version
Accepted author manuscript

Link to publication record in Manchester Research Explorer

Citation for published version (APA):
https://doi.org/10.1103/PhysRevE.101.052129

Published in:
Physical Review E

Citing this paper
Please note that where the full-text provided on Manchester Research Explorer is the Author Accepted Manuscript or Proof version this may differ from the final Published version. If citing, it is advised that you check and use the publisher's definitive version.

General rights
Copyright and moral rights for the publications made accessible in the Research Explorer are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

Takedown policy
If you believe that this document breaches copyright please refer to the University of Manchester’s Takedown Procedures [http://man.ac.uk/04Y6Bo] or contact uml.scholarlycommunications@manchester.ac.uk providing relevant details, so we can investigate your claim.
Quantum limit to nonequilibrium heat-engine performance imposed by strong system-reservoir coupling

David Newman,1,2 Florian Mintert,2 and Ahsan Nazir1,*

1Department of Physics and Astronomy, The University of Manchester, Oxford Road, Manchester, M13 9PL, UK
2Department of Physics, Imperial College London, London, SW7 2AZ, UK

(Dated: May 20, 2020)

We show that finite system-reservoir coupling imposes a distinct quantum limit on the performance of a non-equilibrium quantum heat engine. Even in the absence of quantum friction along the isentropic strokes, finite system-reservoir coupling induces correlations that result in the generation of coherence between the energy eigenstates of the working system. This coherence acts to hamper the engine’s power output, as well as the efficiency with which it can convert heat into useful work, and cannot be captured by a standard Born-Markov analysis of the system-reservoir interactions.

In the same way that fundamental limitations such as the Carnot bound of classical engine cycles impose strict constraints on classical devices, the fundamental limitations of quantum mechanical cycles set the stage for what is achievable with quantum devices. We now have a good understanding of quantum heat engines in the idealized situation of adiabatic dynamics and weak coupling to heat baths [1–21]. Recent experimental realisations of nanoscale heat engines [22–26], however, operate well outside this regime, and effects resulting from quantum coherence are known to impact a machine’s efficiency. In particular, operating quantum thermodynamic cycles faster than adiabatically typically results in coherence generation during the work extraction strokes [4, 27] with correspondingly reduced efficiencies. That this effect, termed quantum friction, can be avoided by dephasing (quantum lubrication [28]) or with control protocols such as shortcuts to adiabaticity [29, 30] is now well established. It thus implies a technical challenge, but not necessarily a fundamental limitation.

In contrast, the impact of quantum coherence entering the cycle via the heat exchange strokes has not so far been considered. Primarily, this is because typical treatments of quantum engine cycles assume idealized weak-coupling, in which system-reservoir correlations are not explicit. In the adiabatic regime where no coherence accumulates during the work extraction strokes, this assumption also results in a reduced system state that remains diagonal in its energy eigenbasis during the heat exchange strokes. Hence, the quantum and classical cycles become largely equivalent.

Here we move beyond the limitations of the weak-coupling approach to consider the impact of system-reservoir correlations accrued as a result of finite coupling during the heat exchange strokes of a quantum Otto cycle. By remaining adiabatic along the work extraction strokes, we show that coherences resulting from system-reservoir correlations lead to performance losses even when conventional quantum friction plays no role. This constitutes a distinct finite-coupling quantum limit to non-equilibrium engine cycles that cannot be captured by a standard weak-coupling approach. Neither can it be avoided by straightforward generalisation of the techniques applied along the isentropic strokes. As we show, for finite system-reservoir coupling the appropriate basis for employing quantum lubrication comprises both system and reservoir components, and is thus difficult to identify experimentally. Likewise, control protocols are challenging to analyse even theoretically beyond the weak-coupling limit. Alternatively, dynamical decoupling could be employed to prevent the growth of system-reservoir correlations, but as this effectively decouples the system and reservoir, it would also simply stop the machine from working. Our work therefore reveals a crucial restriction on quantum engine performance, which is not yet clearly avoidable even in principle.

We consider a two-level system (TLS) interacting separately with two heat reservoirs, at temperatures \(T_h\) and \(T_c\) \((T_h > T_c)\). The Otto cycle consists of four strokes labelled by eight points: \(A'B'B'C'C'D'D'A\). Hot isochore: at \(A'\) the TLS is coupled to the hot reservoir with which it interacts for a time \(\tau_i\) to reach \(B\). The interaction is then switched off instantaneously \((B')\). Isentropic expansion: the TLS Hamiltonian, \(H_S\), is tuned over a time \(\tau\) to reduce the gap between the two energy eigenvalues, reaching point \(C\). Now, the TLS-cold reservoir interaction is switched on suddenly \((C')\). Cold isochore: the TLS interacts with the cold reservoir for a time \(\tau_i\), to reach point \(D\). The system is then decoupled from the cold reservoir \((D')\). Isentropic compression: \(H_S\) is tuned for a time \(\tau\) to increase the gap between energy eigenvalues back to the level at \(A'\), reaching point \(A\). The cycle is completed by turning on the interaction with the hot reservoir \((A')\). These strokes are repeated until a limit cycle is reached [31].

The full Hamiltonian reads

\[
H(t) = H_S(t) + \sum_j (H_{R_j} + H_{T_j}),
\]

where \(H_S(t) = \epsilon(t)\sigma_z/2 + \Delta(t)\sigma_z/2\), \(H_{R_j} = \sum_k \omega_k b_k^\dagger b_k\), and \(H_{T_j} = -\sigma_z \sum_k f_k(b_k^\dagger + b_k)\) for \(j = h, c\), denoting the hot \((h)\) and cold \((c)\) reservoir.
Here, $\sigma_z$ and $\sigma_x$ are the usual TLS Pauli operators. The time dependence of the system is addressed by considering the splitting $\mu(t) = \sqrt{\alpha(t)^2 + \Delta(t)^2}$ between the eigenvalues of $H_S(t)$ at time $t$. We label the values of $\mu$ during the hot or cold stage of the cycle as $\mu_h$ and $\mu_c$, respectively. Bosonic reservoir annihilation operators for excitations at frequencies $\omega_k$ are given by $b_{\omega_k}$. The TLS-reservoir coupling is via $H_{I_j}$ with strengths $f_{\omega_k}$, and is present only during the relevant isochore.

Studies of quantum Otto cycles generally invoke the weak coupling assumption, i.e. that the interaction terms are (negligibly) small. This leads to a tractable analysis in terms of a quantum state of the TLS and reservoirs approximated to remain in tensor product form at all times. In the finite coupling regime of interest, where the interaction terms cannot be neglected, it is a more involved task to compute the state around the cycle [19–21]. In order to access this regime, we extend the finite time Otto cycle considered in Ref. [35] to inherently non-equilibrium, finite time cycles. In this approach, the interaction terms in Eq. (1) are unitarily mapped to collective modes (the RCs) such that

$$H_{I_j} = -\sigma_z \sum_{k_j} f_{\omega_k}(b_{\omega_k}^+ + b_{\omega_k}) = -\lambda_j \sigma_z (a_j^+ + a_j),$$

where $a_j$ annihilates an excitation in the RC mode for reservoir $j$ with natural frequency $\Omega_j$, and $\lambda_j = (\sum_k f_{\omega_k}^2)^{1/2}$. The reservoir Hamiltonians become $H_{R_j} = \Omega_j a_j^+ a_j + \sum_k g_{\omega_k} (a_j^+ + a_j)(r_{\omega_k}^+ + r_{\omega_k}) + \sum_k \nu_{\omega_k} r_{\omega_k}^+ r_{\omega_k}$,  
while the system Hamiltonian remains unchanged. Here, $r_{\omega_k}$ annihilates an excitation at frequency $\nu_{\omega_k}$ in a redefined residual reservoir which interacts weakly with the corresponding RC mode via couplings $g_{\omega_k}$. The full Hamiltonian becomes $H(t) = H_{S'}(t) + \sum_j H_{R_j}$, where $H_{S'}(t) = (t)\sigma_z/2 + \Delta(t)\sigma_x/2 - \sum_j \lambda_j \sigma_z (a_j^+ + a_j) + \Omega_j a_j^+ a_j$, and $H_{R_j} = \sum_k g_{\omega_k} (a_j^+ + a_j)(r_{\omega_k}^+ + r_{\omega_k}) + \sum_k \nu_{\omega_k} r_{\omega_k}^+ r_{\omega_k}$. The RC mapping involves an enlarged view of a redefined system $S'$ whose self-energy $H_S'(t)$ now additionally incorporates the self-energies of the reservoir RCs as well as the TLS-RC coupling terms. The remaining terms represent residual environments $R_j$, and their interactions with the corresponding RC, which may be treated as Markovian [33, 34]. Nevertheless, through the RCs, correlations between the TLS and the reservoirs will form and indeed persist even in the limit cycle.

In the original representation we characterise the TLS-reservoir interactions via spectral densities $J_j(\omega) = \sum_k f_{\omega_k}^2 \delta(\omega - \omega_{\omega_k}) = \alpha \omega \omega_c/(\omega^2 + \omega_c^2)$, taken to be the same for each reservoir, with coupling strength $\alpha$ and cutoff frequency $\omega_c$. As the residual reservoirs are traced out when deriving a master equation for the enlarged system $S'$, to determine the RC mapping one simply needs to find the spectral density $J_j(\nu) = \sum_k g_{\omega_k}^2 \delta(\nu - \nu_{\omega_k})$ that characterizes the coupling between $S'$ and the residual reservoirs, as well as $\Omega_j$ and $\lambda_j$, such that the Heisenberg equations of motion for operators in the TLS subspace are equivalent in both pictures. Imposing this results in $\Omega_j = 2\pi \gamma \omega_c$, $\lambda_j = \sqrt{\pi \alpha \Omega_j/2}$, and $J_j(\nu) = \nu \sqrt{\nu^2 + \Delta^2}/2\pi \omega_c$ [34].

The master equation governing the dynamics of $S'$ during the $j = h,c$ isochore is $\rho(t) = L_j[\rho(t)]$, with $L_j[\rho(t)] = -i[H_{S'}(\rho(t)) - [A_j, \{\chi_j, \rho(t)\}]] + [A_j, \{\Xi_j, \rho(t)\}]$ and $\rho(t)$ the state of the TLS plus both RCs. Here, $A_j = a_j + a_j^+$, the self-Hamiltonian $H_{S'}$ only includes interaction terms for the $j$ RC, $\chi_j = \gamma \int_0^\infty dt \int_0^\infty dw \omega \cos(\omega \tau) \coth(\frac{\beta \omega}{2}) A_j(-\tau)$, and $\Xi_j = \gamma \int_0^\infty dt \int_0^\infty dw \cos(\omega \tau) [H_{S'}(A_j(-\tau)]$, for $A_j(\tau) \equiv e^{iH_{S'}(\tau)} A_j(t) e^{-iH_{S'}(\tau)}$ [34].

Thermodynamic treatments of the Otto cycle usually consider thermal reservoir resources. We wish to isolate strong-coupling effects from any due to coupling to heat reservoirs that are out of equilibrium at the start of each isochore. To this end, we include a mechanism in our finite time cycle to ensure that any uncoupled reservoir returns to thermal equilibrium by the time it is coupled to the TLS once more. For the uncoupled reservoir $j$, which has been driven out of equilibrium during the previous isochore, to return to thermal equilibrium at temperature $T_j$ while the other reservoir and TLS interact, we can add terms to the master equation that act only on the uncoupled RC and hence do not depend on the full system plus RC eigenstructure. The uncoupled RC is a harmonic oscillator so we add standard dissipators $D_{\beta_j}[\rho(t)] = \gamma_d(N_j + 1)L_{\beta_j}[\rho(t)] + \gamma_d N_j L_{\beta_j}^*[\rho(t)]$, where $L_{\beta_j}[\rho(t)] = O\rho(t)O^{-1} - \frac{1}{2}\{O^\dagger O, \rho(t)\}$ and $N_j = (e^{\beta_j \Omega_j} - 1)^{-1}$, with $\beta_j = 1/T_j$ [38, 39]. We choose $\gamma_d$ to ensure thermalisation occurs over a timescale such that the TLS re-couples to a thermal reservoir at the start of the subsequent isochore.

The cycle work output is given by the net energy change of the system across each of the isentropic strokes. For strong coupling, this also involves accounting for the energetic costs associated with turning off the interaction term at the end of the isochores; there are no coupling costs due to rethermalisation of the RCs when uncoupled. This leads to work output [35]

$$W = \text{tr}[H_{S'}^{\beta_c} \rho^{C}] - \text{tr}[H_{S'}^{\beta_h} \rho^{B}] + \text{tr}[H_{S'}^{\beta_d} \rho^{A}]$$
$$- \text{tr}[H_{S'}^{\beta_d} \rho^{D}] - \text{tr}[H_{U}^{\beta_c} \rho^{P}] - \text{tr}[H_{U}^{\beta_h} \rho^{P}],$$

where $h$ and $c$ superscripts indicate that the TLS splitting is set to $\mu_h$ and $\mu_c$, respectively. The density operator $\rho$ is labelled with superscripts $A - D$ indicating the various points around the cycle and as before represents the state of the enlarged $S'$. The energy transferred into the system during the hot isochore is given by $Q = \text{tr}[H_{S'}^{\beta_h} \rho^{B}] - \text{tr}[H_{S'}^{\beta_c} \rho^{A}]$, and we shall term this heat. Through $H_{S'}$, this expression contains contributions from
the TLS-hot RC interaction energy as well as from the hot RC being pulled out of equilibrium, both of which are neglected in weak-coupling treatments.

To evaluate the heat and net work output we need to compute the states $\rho^{A-D}$ of the enlarged system $S'$ when the engine is operating in the limit cycle. In the infinite time (equilibrium) version of the cycle, this involves taking the steady state solution of the RC master equation along each isochore, followed by unitary evolution along the isentropic strokes. For the non-equilibrium case considered here, the calculation is more involved and we must numerically solve the dynamical equation of motion for the full state of $S'$ for a particular isochore time $\tau_i$.

During the isentropic strokes we tune $\epsilon(t)$ and $\Delta(t)$ such that $(\epsilon_h, \Delta_h) \leftrightarrow (\epsilon_c, \Delta_c)$, with $[H_S(t), H_S(t')] = 0$. This ensures that the TLS Hamiltonians at the start of the stroke, $H_S(0)$, and at the end, $H_S(\tau)$, share a common energy eigenbasis. The TLS quantum state then adiabatically follows the change in splitting and no coherence develops even when the stroke is carried out in a finite time $\tau$. We thus avoid any quantum friction along the isentropic strokes, and any variations in performance due to the generation of coherence can instead be attributed to the isochores. For the purposes of maximising power output, it is then preferable to complete this stroke quickly and we consider the limit $\tau \to 0$.

In Fig. 1 (a) we plot the engine’s power output and efficiency at finite coupling. The power output initially increases with coupling strength until a turnover is reached as reservoir decoupling costs begin to dominate over the increase in work output. Note that we are considering here a finite isochore time, $\tau_i$, shorter than that necessary for a stationary state to be reached along the isochores. The energy absorbed from the hot reservoir increases with coupling strength too and this leads to an engine efficiency which decreases monotonically with coupling strength. A similar finding is also reflected in the analysis of the strong coupling regime for a heat engine in Ref. [19] but there the isochores are long enough that the engine is considered to have reached arbitrarily close to equilibrium. In the present case, we show that in fact power output can be maximised before this equilibrium has been reached in Fig. 1 (b). Here we plot the work and power outputs as a function of the isochore time $\tau_i$ for an intermediate coupling strength. As the system $S'$ approaches equilibrium, the work output saturates. Power output, however, is maximised before this equilibrium is reached. If the desired engine metric is how much power can be produced, it is thus preferable to operate out of equilibrium, by choosing shorter isochore times and intermediate coupling strengths.

We now wish to explore the quantum nature of the heat engine, and specifically isolate the effects on engine performance of system-reservoir correlations accrued during the (heat exchange) isochoric strokes as a result of finite coupling with the reservoirs. These correlations manifest themselves in finite coherences in the working system state. We shall therefore make a distinction between a fully quantum version of the cycle, where system-reservoir correlations lead to quantum coherence being generated along the isochores, and one where coherence is prevented from accumulating. In this latter version of the cycle, we introduce into the master equation terms that induce pure dephasing [28, 38, 40–42] in the energy eigenbasis of the working system, while taking care that these have no energetic contribution to the overall evaluation of work output or energy exchange with the reservoirs. To meet with these criteria, the pure dephasing terms must commute with the mapped Hamiltonian $H_{S'}$, which we diagonalise and write as $H_{S'} = \sum_n E_n |E_n\rangle \langle E_n|$. We then construct a pure dephasing Liouvillian

$$L_{dep}[\rho(t)] = \gamma_{dep} \sum_n \left[ |E_n\rangle \langle E_n|, [\rho(t), |E_n\rangle \langle E_n|] \right].$$

We are free to choose the value of $\gamma_{dep}$ to ensure dephasing occurs on an appropriate timescale, such that coherence is prevented from developing during the isochoric strokes of the engine. We can then compute the states of the working system at various points around the cycle as previously described, but with the addition of these terms to both isochores. We define this as the incoherent engine. We stress that these terms achieve
pure dephasing in the energy eigenbasis of the enlarged TLS plus RCs (i.e. $H_{S'}$) rather than just the TLS energy eigenbasis. This is the natural choice for a system interacting strongly with an environment. Introducing pure dephasing terms that only act on the TLS would be inappropriate since they do not commute with the unitary part of the master equation, and thus have a non-zero energetic contribution.

We compare these two versions of the Otto cycle and analyse the effect of quantum coherence on engine performance in Fig. 2 as a function of isochore time $\tau_i$. At large $\tau_i$, the incoherent and coherent engines converge. Here, the state of $S'$ approaches thermal equilibrium with the relevant residual reservoir which is maintained at the hot or cold temperature. This state is then diagonal in the energy eigenbasis of $S'$, and so no coherence is present in either type of engine at points B or D if the isochore is long enough. For shorter times, however, this is not the case, and pure dephasing does have an appreciable effect on the engine metrics. The dephased engine absorbs more heat along the hot isochore (not shown) but outputs a net work large enough to compensate, yielding a higher efficiency at short times than the fully coherent engine. Decoupling costs remain comparable. These are dominated by the RC being driven out of equilibrium along the isochores, which is not prevented by dephasing. In Fig. 3 we show this improved engine performance in parametric plots of power versus efficiency for a selection of coupling strengths. Even at weaker but finite coupling, where coherence generation is less severe, an improvement over the coherent engine can be achieved by dephasing.

In a weak coupling treatment of the Otto cycle the TLS equilibrates to a Gibbs state for long isochore times, which is diagonal in the energy eigenbasis. With the isentropic strokes carried out adiabatically, populations in these energy eigenstates are kept constant and no coherence accumulates in the working system state. In fact, if the isentropic strokes remain adiabatic when the cycle is treated at finite time (as we consider), then no coherence is present during the weak-coupling cycle at any time. At non-negligible system-reservoir interaction strengths, however, correlations that are generated between the TLS and reservoirs lead to quantum coherence in the energy eigenbasis at the end of the isochores. This coherence will, in general, persist during the isentropic strokes even if they are performed adiabatically. In this way, the strongly coupled TLS Otto cycle is inherently quantum in nature.

We further illustrate how dephasing in the TLS plus RC basis accelerates the process of equilibration in Fig. 4. Of crucial importance for work calculations are the TLS population differences in the eigenbasis of $H_S$ [see Eq. (3)]. In Fig. 4 we show how these are affected by dephasing at point B, just before the hot adiabatic stroke begins, and at point D, just before the cold adiabatic stroke begins. In a weak-coupling analysis the TLS populations would be completely insensitive to pure dephasing (which for weak coupling would be defined relative to $H_S$ to possess no energetic contribution). In stark contrast, for the finite coupling theory presented here, TLS populations are significantly affected by pure dephasing in the correct strong-coupling basis of $H_{S'}$. Specifically, the dephased engine displays a faster approach to the steady state than the fully coherent engine. At point B there is therefore a smaller population difference (a) at shorter $\tau_i$ and at point D a larger population difference (b). This results in greater work extracted during the hot adiabat (c) and smaller work invested during the cold adiabat (d) for the dephased engine. The reason is that at point B we wish to have as much population as possible in the

![Fig. 3. Parametric plot of power output against efficiency for various coupling strengths $\alpha$. The dashed curves represent the incoherent engine. The dotted, dot-dashed, and solid curves depict the fully coherent engine. The parameter varied along the curves is the isochore time $\tau_i$. Other parameters as in Fig. 1.](image1)

![Fig. 4. TLS eigenstate population difference at point B in the cycle (a), and at point D (b); work extracted during the hot adiabatic stroke (c) and work input during the cold adiabatic stroke (d), plotted against isochore time $\tau_i$. Solid curves represent the fully coherent engine, dashed curves depict the incoherent engine. Parameters as in Fig. 2.](image2)
excited state as this implies a higher (effective) TLS temperature. This entails that more heat has been absorbed from the hot reservoir during the hot isochore and that the subsequent hot adiabatic stroke can extract a larger amount of work. At point D, on the other hand, it is desirable to have as little population in the excited state as possible. This corresponds to a lower temperature, which means that as much heat as possible has been dissipated into the cold reservoir and that on the subsequent adiabatic stroke, the work performed on the system can be kept to a minimum.

To summarise, we have shown that finite system-reservoir coupling imposes a bound on quantum heat engine performance that is not captured by standard weak-coupling approaches. Our results highlight that even when work extraction is adiabatic, quantum coherence enters the cycle through finite system-reservoir interactions. Protocols designed to avoid quantum friction along the work strokes are not straightforwardly extendable to suppress the generation of coherence along the isochores. We thus expect our findings to be of broad relevance to engine performance that is not captured by standard weak-coupling approaches.

We thank Luis Correa and Ronnie Kosloff for discussions. D.N. is supported by the EPSRC and the CDT in Controlled Quantum Dynamics. A.N. is supported by the EPSRC, Grant No. EP/N008154/1.

* ahsan.nazir@manchester.ac.uk

