Non-Markovian thermal operations boosting the performance of quantum heat engines

Krzysztof Ptaszyński [*]

Institute of Molecular Physics, Polish Academy of Sciences, Mariana Smoluchowskiego 17, 60-179 Poznań, Poland

It is investigated whether non-Markovianity, i.e., the memory effects resulting from the coupling of the system to its environment, can be beneficial for the performance of quantum heat engines. Specifically, two physical models are considered. The first one is a well-known single-qubit Otto engine; the non-Markovian behavior is there implemented by replacing standard thermalization strokes with so-called extremal thermal operations which cannot be realized without the memory effects. The second one is a three-stroke engine in which the cycle consists of two extremal thermal operations and a single qubit rotation. It is shown that the non-Markovian Otto engine can generate more work-per-cycle for a given efficiency than its Markovian counterpart, whereas performance of both setups is superior to the three-stroke engine. Furthermore, both the non-Markovian Otto engine and the three-stroke engine can reduce the work fluctuations in comparison with the Markovian Otto engine, with their relative advantage depending on the performance target. This demonstrates the beneficial influence of non-Markovianity on both the average performance and the stability of operation of quantum heat engines.

I. INTRODUCTION

In recent years much research interest has been devoted to nanoscopic quantum heat engines converting heat into work (or conversely, using work to cool the environment) [1–3]. On the fundamental level, such devices enable one to investigate basic aspects of heat-to-work conversion at the nanoscale. Studies in this area focused on topics such as fundamental resource-theoretic limits to the work extraction [4,5], engineering of optimal driving protocols [5–8], geometric approaches to thermodynamics [9–11], and the role of quantum coherence [12–29], relativist effects [30], or cooperative coupling to the baths [31]. Apart from characterizing the mean behavior of such devices, fluctuations of thermodynamic quantities, which are nonnegligible at the nanoscale, recently received much attention [32–44]. In particular, topics such as fluctuation theorems, characterizing the universal properties of fluctuations [45,46], or thermodynamic trade-offs between power, efficiency, and stability of operation on nanoengines [27–29,47,48] have been investigated. Though most studies so far have been theoretical in nature, quantum heat engines have been experimentally realized (or at least emulated) using setups such as trapped ions [49,50], NV centers [26], or spin systems [51]. Furthermore, practical applications of such engines for cooling of qubits employed in the quantum computing have been proposed [52,53].

This paper investigates whether the performance of nanoscopic heat engines can be improved by using non-Markovian effects. Non-Markovianity is a term used to describe a situation when dynamics of the system is not wholly determined by its actual state, but rather is influenced by the memory about its past evolution, which is encoded in the correlations between the system and its environment [54,55]. This is usually (though not strictly [56]) related to the strong system-environment coupling [54]. As already demonstrated, non-Markovianity is not only interesting from the fundamental point of view—being a feature of many realistic systems—but may also affect the performance of practical appliances, especially nanodevices. In particular, non-Markovian effects were found to be advantageous in applications such as quantum communication channels [57,58], quantum metrology [59,60], entanglement generation [61,62], algorithmic cooling of qubits [52], or photochemical switches [63].

In the context of quantum heat engines, the beneficial influence of non-Markovian effects on the performance of the quantum Otto engine based on a single harmonic oscillator has been previously claimed in Ref. [64]. The observed power enhancement was related to generation of states with occupation numbers exceeding the thermal value due to strong coupling to the environment; it was also claimed that such an engine can surpass the Carnot efficiency. However, subsequent works demonstrated this apparent violation of the second law of thermodynamics to be a result of neglect of the work performed during the coupling and the uncoupling of the system to the thermal bath [65–67]. When this work is taken into account, the validity of the Carnot bound is secured. Indeed, after proper inclusion of the coupling/uncoupling work, the strong coupling has been demonstrated to be detrimental rather than beneficial, though its influence can be mitigated

---

[*]krzysztof.ptaszynski@ifmpan.poznan.pl

by applying appropriate driving protocols [66,67]. Two other works [68,69] demonstrated the power enhancement in the non-Markovian regime for the Otto engine with a specific type of system-bath interaction in which the system is coupled to a Markovian bath through an auxiliary qubit. However, since in both papers the Markovian relaxation rate was kept fixed, while coupling to the auxiliary qubit has been tuned, the observed power enhancement was actually a result of a relatively trivial enhancement of the effective system-bath coupling. Finally, some other publications demonstrated the speed-up of thermalization due to the non-Markovian anti-Zeno effect, and the related boosting of power, for a single-qubit Otto engine coupled to structured baths [70–72]. To summarize, most of the discussed works focused on specific microscopic models of non-Markovian dynamics; therefore, the demonstration of the advantage of non-Markovianity on a more fundamental level would be still desirable.

The goal of this paper is to rigorously demonstrate the beneficial influence of non-Markovian effects on the performance of a quantum heat engine based on a single qubit, without referring to any particular model of the system-bath interaction. To do so, I will apply resource theory methods which provide fundamental constraints on transformations between quantum states due to interaction with the thermal environment (independent of details of microscopic dynamics) [5,73–77]. This formalism has been previously used, for example, to demonstrate the enhancement of photoisomerization yield through non-Markovian effects beyond values achievable in the Markovian regime [63]. Within this framework the state transformations due to coupling to the heat bath are described by means of so-called thermal operations, i.e., energy preserving unitary operations U acting on the system and the bath; the condition of energy preservation is given by the equation $[U, H_S + H_B] = 0$, where H_S and H_B are the Hamiltonians of the system and the bath, respectively. The state of the system is transformed as

$$\rho_S \xrightarrow{TO} \mathcal{E}\rho_S = \mathrm{Tr}_B\big[U \rho_S \otimes \rho_B^{\mathrm{th}} U^\dagger\big], \qquad (1)$$

where $\mathcal{E}$ is the dynamical map representing the transformation, Tr_B is the partial trace over the states of the bath and $\rho_B^{\mathrm{th}} = e^{-\beta H_B}/\mathrm{Tr}(e^{-\beta H_B})$ is the Gibbs state of the bath, with $\beta = 1/(k_B T)$ being the inverse temperature. It can be noted that besides its generality the applied formalism has another advantage: the energy preservation provides that the energy change of the system due to a thermal operation can be unambiguously identified with the heat exchanged with the environment (in contrast to Refs. [64–67]), which significantly simplifies the analysis.

The paper is organized as follows. Section II presents basic elements of the resource-theoretic framework, focusing on the distinction between Markovian and non-Markovian operations. Section III analyzes the average performance of two heat engine models based on a single qubit, with heat strokes described by means of thermal operations: a well known quantum Otto engine [78] and a much less studied three-stroke engine [79]. Section IV is concerned with the behavior of work fluctuations. Section V discusses the experimental realizability of quantum heat engines employing non-Markovian thermal operations. Finally, Sec. VI brings the conclusions.

II. THERMAL OPERATIONS

A. Criteria of Markovianity of thermal operations

Let us start our discussion with a rigorous definition of the distinction between Markovian and non-Markovian state transformations. The thermal operation is deemed to be Markovian when it can be expressed as the time ordered exponential [80]

$$\mathcal{E} = \mathcal{T}\big\{e^{\int_0^t \mathcal{L}_\tau d\tau}\big\} \qquad (2)$$

of a (possibly time-dependent) Lindblad generator $\mathcal{L}_t$ defined as

$$\mathcal{L}_t \rho_S(t) = \mathcal{H}\rho_S(t) + \mathcal{D}_t \rho_S(t), \qquad (3)$$

where $\mathcal{H}$ and $\mathcal{D}_t$ are generators of the unitary and the dissipative dynamics, respectively. They are given by the equations

$$\mathcal{H}\rho_S(t) = -i[H_S, \rho_S(t)], \qquad (4)$$

$$\mathcal{D}_t \rho_S(t) = \sum_i r_i(t)\bigg[L_i \rho_S(t)L_i^\dagger - \frac{1}{2}\{L_i^\dagger L_i, \rho_S(t)\}\bigg], \qquad (5)$$

where L_i and $r_i(t) \geqslant 0$ are the Lindblad operators and the transition rates, respectively. Here and from hereon $\hbar = 1$ is taken. The Lindblad generator is required to fulfill two properties [80]:

(1) Stationary thermal state: the Gibbs state of the system,

$$\rho_S^{\mathrm{th}} = \frac{e^{-\beta H_S}}{\mathrm{Tr}(e^{-\beta H_S})}, \qquad (6)$$

is the stationary state of the dynamics at all times, i.e.,

$$\forall t: \quad \mathcal{L}_t \rho_S^{\mathrm{th}} = 0. \qquad (7)$$

(2) Covariance: $\mathcal{H}$ and $\mathcal{D}_t$ commute for an arbitrary density matrix ρ, i.e.,

$$\forall \rho: \quad \mathcal{H}\mathcal{D}_t \rho = \mathcal{D}_t \mathcal{H}\rho. \qquad (8)$$

As shown by Lostaglio and Korzekwa [80], any Markovian thermal operation on an n-level quantum system can be realized by a sequence of full or partial thermalizations acting on pairs of energy levels. Thus, in particular, a Markovian thermal operation acting an a two-level system (qubit) corresponds to its full or partial thermalization.

It should be noticed that Eq. (2) is not an only way to define Markovian operations present in the literature. As a matter of fact, there exists a plethora of definitions of quantum Markovianity, with certain hierarchy relations between them [81,82]. In particular, the condition given by Eq. (2) is stronger than either the CP-divisibility [81] or the operational divisibility defined within the process tensor framework [82], since the dynamical map can be always divided into mutually independent maps: $\mathcal{E}_t = \mathcal{E}_{t-s}\mathcal{E}_s$, with $\mathcal{E}_t$ given by the right-hand side of Eq. (2). Furthermore, Spaventa et al. [63] investigated, within the same resource-theoretic framework, how efficiency of photoisomerization processes is limited by a condition even stronger than Markovianity, namely, the embeddability. This

is defined by a requirement that the dynamical map can be written as $\mathcal{E} = e^{\mathcal{L}t}$ with a time-independent Lindbladian $\mathcal{L}$. However, for qubits both conditions are equivalent, which is demonstrated in the Appendix; this is no longer true for higher dimensions of the Hilbert space, e.g., for three-level systems investigated in Ref. [63].

It needs also to be mentioned that the stationarity of the Gibbs state holds for both Markovian and non-Markovian thermal operations: $\mathcal{E}\rho_S^{\mathrm{th}} = \rho_S^{\mathrm{th}}$ [74]. This is a consequence of the requirement of energy preservation discussed in the Introduction. However, for a generic (energy nonpreserving) evolution induced by the coupling to the bath, the Gibbs state may be no longer a stationary state due to correlations between the system and the environment; in particular, for a time-independent coupling the fixed point is given by a so called mean force Gibbs state [83,84]. As a matter of fact, operations with nonthermal fixed points have been applied in some previous works on non-Markovian heat engines [64,65].

B. Thermal operations on a single qubit

The paper focuses on the case when the heat engine is based on a single qubit, for which an exact analytic form of an arbitrary thermal operation can be found. The qubit is described by the Hamiltonian

$$H_S = \omega |e\rangle\langle e|, \tag{9}$$

where ω is the qubit energy gap and $|e\rangle$ denotes the excited state of the qubit.

In general, a dynamical map describing the evolution of an open quantum system may couple the dynamics of the diagonal elements of the density matrix (populations) and the off-diagonal ones (coherences). However, due to energy conservation, for thermal operations the evolution of populations and coherences is decoupled in the eigenstate basis, and thus may be considered in a separate way [85]. Let us first consider the evolution of the population vector $\mathbf{p} = (p_g, p_e)^T$, where p_g and $p_e = 1 - p_g$ are populations of the ground and the excited state, respectively. It undergoes the transformation [77,86]

$$\mathbf{p} \xrightarrow{TO} \Lambda\mathbf{p}, \tag{10}$$

where

$$\Lambda(\omega, \beta, \lambda) = (1-\lambda)\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \lambda\begin{pmatrix} 1 - e^{-\beta\omega} & 1 \\ e^{-\beta\omega} & 0 \end{pmatrix} \tag{11}$$

is referred to as a Gibbs-stochastic matrix [77,86], with parameter $\lambda \in [0, 1]$ describing the strength of interaction.

The qubit coherences ρ_{ge} and $\rho_{eg} = \rho_{ge}^*$ undergo the transformation [77,85]

$$\rho_{ge} \xrightarrow{TO} \gamma\rho_{ge}, \tag{12}$$

where $\gamma \in [0, \sqrt{1 - \lambda e^{-\beta\omega}(1-\lambda)}] \leqslant 1$. For the sake of simplicity, in this paper the dynamics of coherences will be neglected. This corresponds to the case when $\gamma = 0$, and thus the qubit is fully dephased at the end of each thermal operation.

It appears that the parameter λ determines whether the thermal operation is Markovian or not. It is Markovian
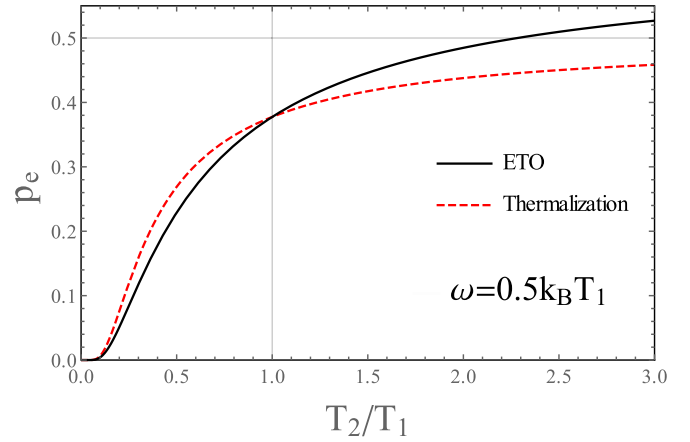

FIG. 1. The population of the excited state p_e after the application of ETO (black solid line) or thermalization (red dashed line) for a qubit initialized in the thermal state with the temperature T_1 and the applied thermal operation corresponding to the reservoir temperature T_2. The qubit gap fixed at $\omega = 0.5k_BT_1$.

when [52]

$$\lambda \leqslant p_g^{\mathrm{th}} = \frac{1}{1 + e^{-\beta\omega}}, \tag{13}$$

where p_g^{th} is the thermal population of the ground state for the inverse bath temperature β. In particular, $\lambda = p_g^{\mathrm{th}}$ and $\lambda < p_g^{\mathrm{th}}$ correspond to the full and the partial thermalization, respectively.

C. Extremal thermal operation (ETO)

Let us now consider the non-Markovian regime of $\lambda > p_g^{\mathrm{th}}$, focusing on the limiting case of $\lambda = 1$. The thermal operation Λ with $\lambda = 1$ is referred to as the *extremal thermal operation* (ETO) [79]. The physical nature of such an operation can be understood qualitatively by analyzing the example of a qubit initialized in the thermal state with some temperature T_1 and than transformed by ETO corresponding to the reservoir temperature T_2. Figure 1 presents the population of the excited state p_e after the application of ETO (black solid line) and after thermalization to the temperature T_2 (red dashed line). As one can observe, for $T_2 < T_1$ ETO "cools the qubit more" than thermalization, i.e., the population of the excited state is reduced below the thermal population corresponding to the temperature T_2 (and, consequently, more heat is extracted from the system than during thermalization). Conversely, for $T_2 > T_1$ ETO generates the population of the excited state exceeding the thermal value. Most notably, for a high enough T_2 the extremal thermal operation may even produce a population inversion $p_e > 1/2 > p_g$; this is impossible to achieve using Markovian thermal operations acting on a noninverted state.

III. AVERAGE WORK

In this section the average work generated during a single cycle of the considered quantum heat engines is analyzed. Specifically, Secs. III A and III B present the models of the Otto engine and the three-stroke engine as well as analytic
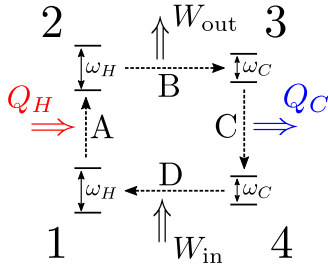
FIG. 2. A schematic representation of the Otto engine cycle. Successive strokes denoted as A, B, C, and D and their starting points as 1, 2, 3, and 4. Q_H and Q_C denote the heat flows whereas W_{out} and W_{in}—the work output and the work input, respectively.

formulas describing work, heat and efficiency, whereas in Sec. III C the performance of both models is compared.

A. Otto engine

Let me first present a principle of operation of the quantum Otto engine. It is based on a single qubit attached to two thermal baths: the hot bath H with the temperature T_H and the cold one C with the temperature T_C. The operation of the Otto engine is based on a periodic coupling (decoupling) of the qubit to (from) the thermal baths, accompanied by a periodic modulation of the qubit gap ω between two values: ω_H and ω_C. A schematic representation of the Otto engine cycle is presented in Fig. 2, with strokes denoted with letters A, B, C, and D and their starting points denoted with numbers 1, 2, 3, and 4. The strokes can be described as follows:

(A) Heating: the qubit gap $\omega = \omega_H$ is taken and the system is placed into contact with the hot bath; as a result, the population of the excited state is increased from $p_{e,1}$ to $p_{e,2}$ and heat $Q_H = \omega_H(p_{e,2} - p_{e,1})$ is delivered from the bath to the system.

(B) Work output: the qubit gap is shifted from ω_H to ω_C without change of qubit state ($p_{e,3} = p_{e,2}$); as a result, work $W_{\text{out}} = (\omega_H - \omega_C)p_{e,3}$ is generated.

(C) Cooling: the system is placed into contact with the cold bath; as a result, the population of the excited state is decreased from $p_{e,3}$ to $p_{e,4}$ and heat $Q_C = \omega_C(p_{e,4} - p_{e,3})$ is delivered from the system to the bath.

(D) Work input: the qubit gap is shifted from ω_C to ω_H without change of the qubit state ($p_{e,1} = p_{e,4}$); as a result, work $W_{\text{in}} = (\omega_H - \omega_C)p_{e,1}$ is delivered to the engine.

The strokes A and C, during which heat is exchanged with the bath, are referred to as the heat strokes, while the strokes B and D, during which work is performed, as the work strokes.

The net work generated during a cycle is equal to

$$W = W_{\text{out}} - W_{\text{in}} = (\omega_H - \omega_C)(p_{e,3} - p_{e,1}), \quad (14)$$

and the efficiency of the Otto engine is given by a simple formula

$$\eta = \frac{W}{Q_H} = 1 - \frac{\omega_C}{\omega_H}. \quad (15)$$

Of course, though the equation above is mathematically well defined for arbitrary ω_C and ω_H, the setup works as a heat engine only when $T_C/T_H \leqslant \omega_C/\omega_H \leqslant 1$, such that $0 \leqslant \eta \leqslant \eta_C$, where $\eta_C = 1 - T_C/T_H$ is the Carnot efficiency; in particular,

for $T_C/T_H \geqslant \omega_C/\omega_H$ both the work output W and the heat extracted from the hot bath Q_H become negative (work is delivered to heat the hot bath).

The populations of the exited states $p_{e,1}$ and $p_{e,3}$, which are required to calculate the work-per-cycle, can be found in the following way. First, the steady state of the system in point 1 is found by solving the equation

$$\Lambda_C \Lambda_H \mathbf{p}_1 = \mathbf{p}_1, \quad (16)$$

where Λ_C and Λ_H are Gibbs-stochastic maps given by Eq. (11) describing the interaction with the cold bath and the hot bath, respectively:

$$\Lambda_\alpha = \Lambda(\omega_\alpha, \beta_\alpha, \lambda_\alpha), \quad (17)$$

with $\alpha \in \{C, H\}$. The populations in point 3 are then calculated as

$$\mathbf{p}_3 = \Lambda_H \mathbf{p}_1. \quad (18)$$

Two versions of the Otto engine are then considered: ones with Markovian and non-Markovian thermal operations. For the Markovian Otto engine I take

$$\lambda_\alpha = \frac{1}{1 + e^{-\beta_\alpha \omega_\alpha}}, \quad (19)$$

which corresponds to thermalization of the system. The stationary excited state populations take then the thermal values

$$p_{e,1}^{\text{M}} = \frac{1}{1 + e^{\beta_C \omega_C}}, \quad (20)$$

$$p_{e,3}^{\text{M}} = \frac{1}{1 + e^{\beta_H \omega_H}}. \quad (21)$$

For the non-Markovian setup $\lambda_\alpha = 1$ is taken, which generates the extremal thermal operation. The excited state populations read then as

$$p_{e,1}^{\text{NM}} = \frac{e^{\beta_H \omega_H} - 1}{e^{\beta_C \omega_C + \beta_H \omega_H} - 1}, \quad (22)$$

$$p_{e,3}^{\text{NM}} = \frac{e^{\beta_C \omega_C} - 1}{e^{\beta_C \omega_C + \beta_H \omega_H} - 1}. \quad (23)$$

B. Three-stroke engine

Let us now turn our attention to a less known model of a quantum heat engine, namely, the three-stroke engine proposed by Łobejko *et al.* [79]. As before, it is based on a single qubit attached to two thermal baths. A schematic representation of its operation is presented in Fig. 3, with strokes denoted as A, B, and C and starting points of the strokes as 1, 2, and 3. The strokes can be described as follows:

(A) Heating: the system is placed into contact with the hot bath; as a result, the population of the excited state is increased from $p_{e,1}$ to $p_{e,2}$ and heat $Q_H = \omega(p_{e,2} - p_{e,1})$ is delivered from the bath to the system.

(B) Work extraction stroke: the qubit is coherently flipped such that populations of the ground and the excited states are interchanged: $p_{e,3} = p_{g,2} = 1 - p_{e,2}$; as a result work

$$W = \omega(p_{e,2} - p_{g,2}) = \omega(2p_{e,2} - 1) \quad (24)$$
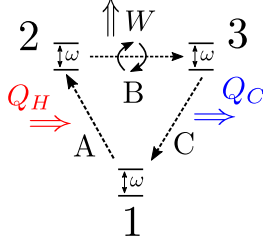
is generated.

FIG. 3. A schematic representation of the three-stroke engine cycle. Successive strokes denoted as A, B, and C and their starting points as 1, 2, and 3. Q_H and Q_C denote the heat flows whereas W—the work output.

(C) Cooling: the system is placed into contact with the cold bath; as a result, the population of the excited state is decreased from $p_{e,3}$ to $p_{e,1}$ and heat $Q_C = \omega(p_{e,1} - p_{e,3})$ is delivered from the system to the bath.

The be precise, in original work of Łobejko *et al.* [79] the work stroke consisted of the energy exchange with the battery; here it is replaced by the qubit flip for the sake of simplicity. As Eq. (24) implies, the engine generates work only in the case when heating produces the population inversion: $p_{e,2} > 1/2 > p_{g,2}$. Therefore, since the Markovian thermal operations cannot generate the population inversion, the engine requires non-Markovian effects to operate.

Analogously to the Otto engine, the stationary populations in point 1 can be found by solving the equation

$$\Lambda_C \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \Lambda_H \mathbf{p}_1 = \mathbf{p}_1, \tag{25}$$

where

$$\Lambda_\alpha = \Lambda(\omega, \beta_\alpha, \lambda_\alpha) \tag{26}$$

is the Gibbs-stochastic matrix given by Eq. (11). Probability vectors in points 2 and 3 read as

$$\mathbf{p}_2 = \Lambda_H \mathbf{p}_1, \tag{27}$$

$$\mathbf{p}_3 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \Lambda_H \mathbf{p}_1. \tag{28}$$

From hereon, the extremal thermal operations will be applied and thus $\lambda_\alpha = 1$ will be taken. The stationary excited state populations read then as

$$p_{e,1} = \frac{1}{1 + e^{(\beta_H + \beta_C)\omega}}, \tag{29}$$

$$p_{e,2} = \frac{1}{e^{\beta_H \omega} + e^{-\beta_C \omega}}, \tag{30}$$

$$p_{e,3} = 1 - \frac{1}{e^{\beta_H \omega} + e^{-\beta_C \omega}}. \tag{31}$$

Using the expressions above, the average work-per-cycle and the efficiency are given by equations [79]

$$W = \omega \left[\frac{2}{e^{\beta_H \omega} + e^{-\beta_C \omega}} - 1 \right], \tag{32}$$

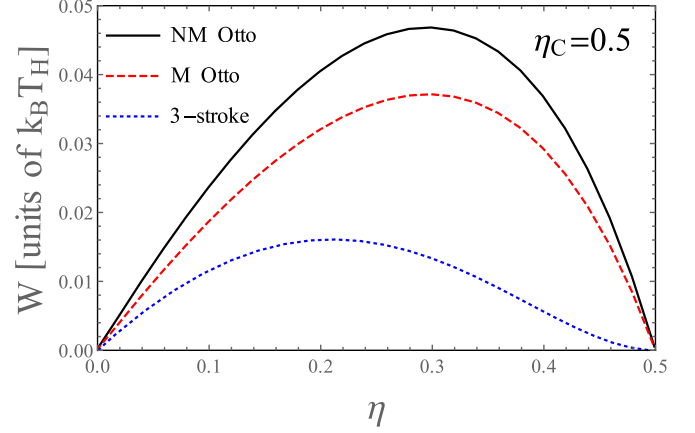$$\eta = \frac{W}{Q_H} = 1 - \frac{e^{\beta_H \omega} - 1}{1 - e^{-\beta_C \omega}}. \tag{33}$$

FIG. 4. The average work-per-cycle as a function of efficiency η for the non-Markovian Otto engine (black solid line), the Markovian Otto engine (red dashed line) and the three-stroke engine (blue dotted line). The Carnot efficiency fixed at $\eta_C = 1 - T_C/T_H = 0.5$.

C. Comparison

I will now compare the average work-per-cycle as a function of efficiency η for a fixed Carnot efficiency η_C. Before that, let me first note that the Otto engine and the three-stroke engine are characterized by different numbers of free parameters. The three-stroke engine is characterized by three free parameters: ω, T_H and T_C. As a result, by fixing η and η_C one fully determines the ratio $W/(k_B T_H)$. In contrast, the Otto engine is characterized by four free parameters: ω_H, ω_C, T_H, and T_C. Therefore, for a chosen efficiency η the ratio ω_C/ω_H becomes fixed [see Eq. (15)], however, a single qubit gap ω_H can be tuned. Therefore, the presented work-per-cycle is here a maximum value, which is optimized over the full range of ω_H. The results are presented in Fig. 4. As one can observe, the non-Markovian Otto engine generates more work-per-cycle than the Markovian one. This is because by heating (cooling) the qubit through the extremal thermal operation one can produce a higher (lower) population of the excited state than using thermalization: $p_{e,3}^{NM} > p_{e,3}^{NM}$ and $p_{e,1}^{NM} < p_{e,1}^{M}$ [see Eqs. (20)–(23)]. Therefore, the work-per-cycle, which is proportional to the difference $p_{e,3} - p_{e,1}$ [see Eq. (14)], is larger for the non-Markovian case. This demonstrates that non-Markovianity can be beneficial for the performance of quantum heat engines. Let me note that a similar advantage of non-Markovian thermal operations for the purpose of cooling the qubit has been previously demonstrated by Alhambra *et al.* [52]

One can also observe that in a full parameter range both the Markovian and the non-Markovian engine generate more work than the three-stroke engine. This might be related to a smaller number of free parameters of the three-stroke engine, which hinders the performance optimization. Note that this conclusion differs from the one provided by Łobejko *et al.* [79], who stated that the three-stroke engine becomes optimal for high efficiencies. However, in the mentioned paper the work was expressed in the units of the "energy gap," which was defined in a different way for the models analyzed: as $\omega_H - \omega_C$ for the Otto engine while as ω for the three-stroke

engine. In this article, in contrast, the same units of $k_B T_H$ are used for both models.

Let me here discuss the nature of the performance boost due to non-Markovianity in the context of previous works on the topic. As in Refs. [64,65], the increase of the work output is here related to the nonthermal nature of the state generated at the end of a heat stroke; in contrast to these publications, however, this is not achieved at the cost of work associated with the coupling and the decoupling of the system from the bath. In contrast, a qualitatively different physical mechanism of the power enhancement was considered in Refs. [68–72], where non-Markovianity was used to speed-up the approach to the thermal state rather then to modify the final state.

Finally, it needs to be noticed that the work-per-cycle may not always be an optimal measure of the average performance. Indeed, for a Markovian heat engine with a fixed maximum value of the system-bath coupling the maximum power is achieved in the limit of the cycle period going to 0, in which the work-per-cycle also approaches 0 [7]. However, the work-per-cycle might be still a valuable measure, for example, when maximum frequency of driving is limited due to technical constraints.

IV. WORK FLUCTUATIONS

So far, the paper focused on the average performance of heat engines. Let us now turn to another topic which received much attention in the recent years, namely, the work fluctuations [27–29,32–48]. This field of investigation is related to the issue of stability of operation of heat engines: a good engine is expected to generate a well defined amount of work during a given time period, and thus, to minimize fluctuations. As a measure of work fluctuations this paper will use the variance of the work generated during N successive cycles,

$$\langle \Delta w_N^2 \rangle = \langle w_N^2 \rangle - \langle w_N \rangle^2, \qquad (34)$$

where $\Delta w_N = w_N - \langle w_N \rangle$. Here lowercase w_N denotes the stochastic (fluctuating) work generated during N cycles to avoid confusion with capital W denoting the average work-per-cycle. From the operational perspective, w_N for the Otto engine can be defined as a sum of stochastic work increments associated with each work stroke:

$$w_N = \sum_{i=1}^{N} [(E_{i,2} - E_{i,3}) + (E_{i,4} - E_{i,1})], \qquad (35)$$

where $E_{i,k}$ is the outcome of the energy measurement (i.e., a projective measurement in the energy eigenbasis) in point k of the ith cycle. This generalizes a standard two-point measurement protocol [87–89] to the case of multiple strokes. Importantly, since the density matrix is diagonal in the energy eigenbasis at the end of each stroke, the state of the system is not disturbed by the energy measurements. Analogously, the stochastic work for the three-stroke engine can be expressed as

$$w_N = \sum_{i=1}^{N} (E_{i,2} - E_{i,3}). \qquad (36)$$

In Sec. IV A the methods used to calculate the work fluctuations will be presented while Sec. IV B presents the results.

A. Methods

The work fluctuations can be determined by means of the full counting statistics formalism [90,91]. Within this framework one defines the counting-field-dependent stochastic map Π_χ describing the evolution of the system during a single cycle. For the Otto engine such a map reads

$$\Pi_{\chi,\text{Otto}} = B_- \Lambda_C B_+ \Lambda_H, \qquad (37)$$

with

$$B_\pm = \begin{pmatrix} 1 & 0 \\ 0 & \exp[\pm \chi(\omega_H - \omega_C)] \end{pmatrix}. \qquad (38)$$

Here the counting field $\pm \chi$ counts the events during which the energy of the excited state is decreased (increased) which corresponds to the positive (negative) work generation. The analogous map for the three-stroke engine reads

$$\Pi_{\chi,\text{3-stroke}} = \Lambda_C \begin{pmatrix} 0 & \exp[\chi \omega] \\ \exp[-\chi \omega] & 0 \end{pmatrix} \Lambda_H. \qquad (39)$$

The average work generated during N cycles, as well as the work variance, can be then calculated using the equations

$$\langle w_N \rangle = \left[\frac{d}{d\chi} G_N(\chi) \right]_{\chi=0}, \qquad (40)$$

$$\langle \Delta w_N^2 \rangle = \left[\frac{d^2}{d\chi^2} G_N(\chi) \right]_{\chi=0}, \qquad (41)$$

where

$$G_N(\chi) = \ln[\mathbf{1}^T (\Pi_\chi)^N \mathbf{p}_1] \qquad (42)$$

is the cumulant generating function. Here $\mathbf{1} = (1, 1)^T$ while $\mathbf{p}_1$, as in Sec. III, is the stationary state of the system in point 1. Since the average value of work generated in each cycle is the same, the average work generated during N cycles grows linearly with the number of cycles: $\langle w_N \rangle = NW$. However, as will be later discussed more thoroughly, the higher work cumulants may not grow linearly due to correlations of work generated in the successive cycles. In particular, the work variance grows sublinearly ($\langle \Delta w_N^2 \rangle < N \langle \Delta w_1^2 \rangle$) for negative intercycle correlations while supralinearly ($\langle \Delta w_N^2 \rangle > N \langle \Delta w_1^2 \rangle$) in the opposite case [43].

To characterize the long-time properties of work fluctuations it is useful to consider the limit of infinite number of cycles ($N \to \infty$). Fluctuations in such a limit are well defined when described by means of scaled cumulants, i.e., work cumulants divided by N. In particular, the scaled mean work and the work variance can be calculated as [90,91]

$$\lim_{N \to \infty} \frac{\langle w_N \rangle}{N} = \left[\frac{d}{d\chi} g(\chi) \right]_{\chi=0},$$

$$\lim_{N \to \infty} \frac{\langle \Delta w_N^2 \rangle}{N} = \left[\frac{d^2}{d\chi^2} g(\chi) \right]_{\chi=0}, \qquad (43)$$

where

$$g(\chi) = \ln \lambda_0(\chi) \qquad (44)$$

is the scaled cumulant generating function, with $\lambda_0(\chi)$ being the dominant eigenvalue of Π_χ.
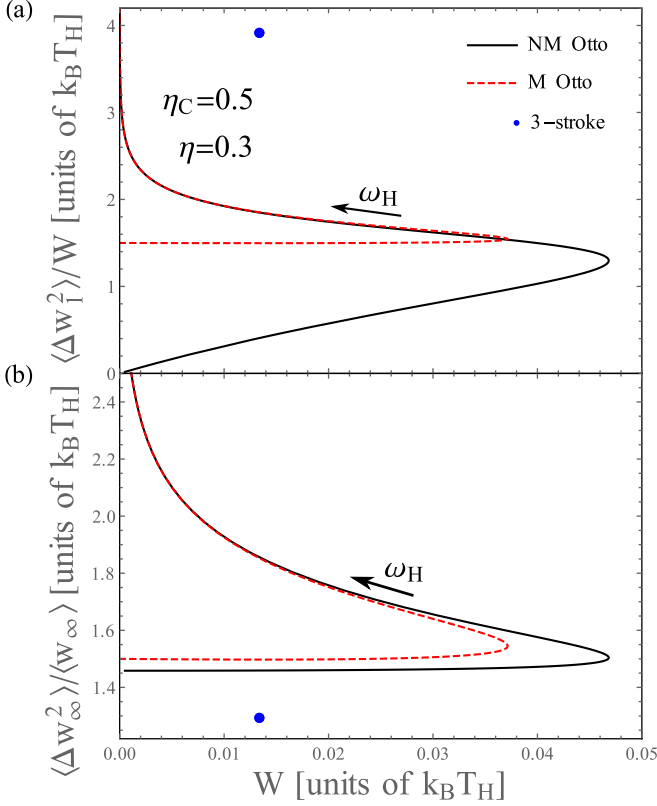
FIG. 5. The ratio of the work variance to the average work for a single-cycle (a) and in the limit of infinite number of cycles (b) for the non-Markovian Otto engine (black solid line), the Markovian Otto engine (red dashed line) and the three-stroke engine (blue point). The arrow near the curves points the direction of increasing ω_H. Results for $\eta = 0.3$ and $\eta_C = 1 - T_C/T_H = 0.5$.

B. Results

Let me now present the results. Figure 5 shows the ratio of the work variance to the mean work as a function of work-per-cycle for a fixed efficiency η and the Carnot efficiency $\eta_C = 1 - T_C/T_H$. The fluctuations are analyzed in two opposite limits: the short-time and the long-time behavior is described by means of fluctuations during a single cycle [Fig. 5(a)] and in the limit of infinite number of cycles [Fig. 5(b)], respectively. As one can note, the work fluctuations are represented by curves for the Otto engines while by a point for the three-stroke engine. This is related to a different number of free parameters of the models considered, which was discussed in Sec. III C: for the three-stroke engine fluctuations and the average work become fixed for a specific choice of η and η_C, whereas for the Otto engine different values can be accessed by tuning ω_H; the direction of increasing ω_H is denoted by the arrow near the curves.

One can observe that for a single cycle [Fig. 5(a)] the non-Markovian Otto engine generates less fluctuations for a given work-per-cycle than the Markovian one. For $\omega_H \to 0$ the ratio $\langle \Delta w_1^2 \rangle/W$ goes to 0, which is, however, also accompanied by the mean value of work going to 0. Additionally, both the non-Markovian and the Markovian Otto engine generate less fluctuations than the three-stroke engine. Interestingly, a qualitatively different behavior is observed in the limit of in-
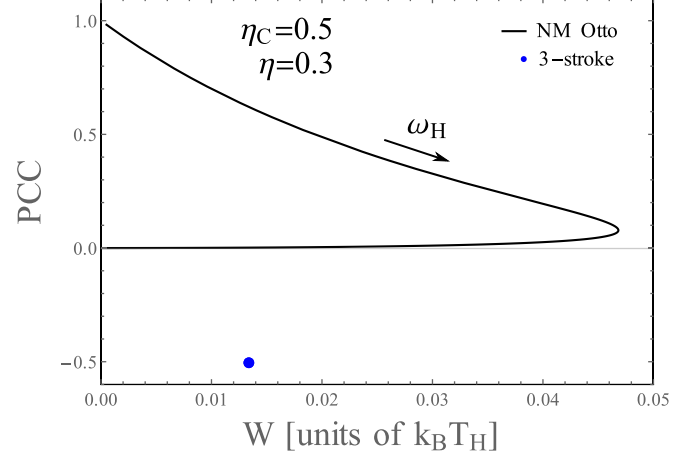

FIG. 6. The Pearson correlation coefficient of work correlations as a function of work-per-cycle for the non-Markovian Otto engine (black solid line) and the three-stroke engine (blue point). The arrow near the black curve points the direction of increasing ω_H. Results for $\eta = 0.3$ and $\eta_C = 1 - T_C/T_H = 0.5$.

finite number of cycles [Fig. 5(b)]. While the non-Markovian Otto engine still generates less fluctuations than the Markovian one, the difference is much less pronounced and the ratio $\langle \Delta w_\infty^2 \rangle/\langle w_\infty \rangle$ no longer goes to 0 for $\omega_H \to 0$. Even more importantly, the three-stroke engine now generates the smallest amount of work fluctuations, which is opposite to the single-cycle case. However—for the parameters considered—the non-Markovian Otto engine can generate about 4 times more work-per-cycle than the three-stroke engine for the fluctuation-to-work ratio $\langle \Delta w_\infty^2 \rangle/\langle w_\infty \rangle$ larger only by about 10%. Therefore, either the three-stroke engine or the non-Markovian Otto engine may be optimal depending on whether one is more interested in maximizing power or minimizing the work fluctuations.

As shown before by Xu and Watanabe [43], a different behavior of single-cycle and long-time fluctuations is a result of correlations of the work generated in the successive cycles. Specifically, the long-time fluctuations are enhanced (reduced) by positive (negative) correlations. This can be demonstrated by analyzing the work covariance

$$\mathrm{Cov}(w) = \langle \Delta w_I \Delta w_{II} \rangle, \qquad (45)$$

where w_I and w_{II} are the amounts of work generated in two successive cycles. As one can easily show, the work covariance is related to work variances through the relation

$$\langle \Delta w_2^2 \rangle = 2\langle \Delta w_1^2 \rangle + 2\mathrm{Cov}(w). \qquad (46)$$

Figure 6 shows the Pearson correlation coefficient

$$\mathrm{PCC} = \frac{\mathrm{Cov}(w)}{\langle \Delta w_1^2 \rangle} = \frac{\langle \Delta w_2^2 \rangle}{2\langle \Delta w_1^2 \rangle} - 1 \qquad (47)$$

for the non-Markovian Otto engine and the three-stroke engine, with parameters as in Fig. 5. It takes values within the range $\mathrm{PCC} \in [-1, 1]$, with a positive (negative) sign corresponding to positive (negative) correlations. The intercycle correlations are not present for the Markovian Otto engine (employing thermalization strokes) because the system is

resetted to the thermal state after each heat stroke, and thus the initial state at the beginning of each cycle does not depend on the previous evolution (of course, this is a consequence of the full thermalization of the system during heat strokes rather than the Markovian dynamics itself – correlations will still appear in the Markovian partial thermalization regime of $\lambda < p_g^{\text{th}}$, where the state at the end of a heat stroke also depends on the initial one). As one can observe, the non-Markovian Otto regime is characterized by positive intercycle work correlations which leads to enhancement of long-time fluctuations. In particular, perfect positive correlations (PCC = 1) are observed in the limit $\omega_H \to 0$. Conversely, the work correlations are negative for the three-stroke engine. Interestingly, the Pearson correlation coefficient is then given by a simple formula

$$\text{PCC}_{\text{3-stroke}} = -e^{-(\beta_C + \beta_H)\omega}, \qquad (48)$$

which implies that the work generated in the successive cycles becomes perfectly anti-correlated (PCC = −1) for $\omega \to 0$.

V. EXPERIMENTAL REALIZABILITY

Finally, let us discuss how heat engines with non-Markovian thermal operations can be realized experimentally. First, it was shown that the extremal thermal operation described in Sec. II C can be realized by coupling the qubit to the thermal state of a single bosonic mode

$$\rho_{\text{bos}}^{\text{th}} = \frac{e^{-\beta\Omega a^\dagger a}}{\text{Tr}(e^{-\beta\Omega a^\dagger a})}, \qquad (49)$$

with the excitation energy Ω resonant with the qubit gap ω, i.e., $\Omega = \omega$; here $a^\dagger$ and a are the bosonic creation and annihilation operators, respectively. Then, one needs to perform the energy-preserving unitary operation [86]

$$U = |g, 0\rangle\langle g, 0| + \sum_{n=1}^{\infty} (|g, n\rangle\langle e, n-1| + |e, n-1\rangle\langle g, n|), \qquad (50)$$

where g (e) denotes the ground (excited) state of the qubit and n is the number of excitations of the bosonic mode. Such an operation exchanges a single excitation between the qubit and the bosonic mode. It can be exactly realized by means of the intensity-dependent Jaynes-Cummings interaction Hamiltonian [92,93]

$$\tilde{H}_{JC} = J[\sigma_+(aa^\dagger)^{-1/2}a + \sigma_-(aa^\dagger)^{-1/2}a^\dagger], \qquad (51)$$

where J is the coupling strength while $\sigma_+ = |e\rangle\langle g|$ and $\sigma_- = |g\rangle\langle e|$ are raising and lowering operators, respectively; more specifically, a thermal operation with an arbitrary λ can be realized by varying the time of coupling between the qubit and the bosonic mode, which enables a switching between Markovian and non-Markovian regimes [86]. Bera *et al.* [94] shown that the intensity-dependent Jaynes-Cummings interaction can be realized by adding some engineered anharmonicity to the bosonic mode. Alternatively, a good approximation of the extremal thermal operation can be realized by using the standard Jaynes-Cummings interaction Hamiltonian [52,86]

$$H_{JC} = J(\sigma_+ a + \sigma_- a^\dagger). \qquad (52)$$
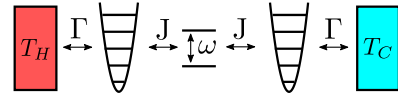

FIG. 7. A schematic idea of experimental realization of the non-Markovian heat engine. A qubit is coupled to Markovian thermal baths with temperatures T_H and T_C via auxiliary bosonic modes. Here J denotes (possibly tunable) Jaynes-Cummings coupling between the qubit and bosonic modes, while Γ is the relaxation rate of the bosonic mode.

To realize a continuous work generation, the bosonic modes need to be either resetted to the thermal state after each heat stroke or replaced with another one. The former option can be implemented by coupling the bosonic levels to Markovian thermal baths. Therefore, finally, the non-Markovian heat engine can be realized by using a qubit coupled to Markovian thermal baths through auxiliary bosonic modes (Fig. 7). In general, the coupling to auxiliary bosons needs to be tunable to enable a switching of the system-bath coupling; however, for the Otto engine the effective coupling to the bosonic modes is modulated just by tuning the qubit gap ω, since the excitation exchange is suppressed in the nonresonant regime ($\omega \neq \Omega$) [17,72,95]. An experimental realization of such a setup may be based, for example, on superconducting circuits with LC resonators playing the role of bosonic modes and the attached resistors playing the role of Markovian baths [96–98]; indeed, the Otto engine based on such an architecture has been previously theoretically proposed [17] and experimental steps towards its realization have been made [95]. Of course, realistic superconducting devices may differ from the idealistic model considered here due to, e.g., presence of the environment-induced decoherence, photon leakage from the resonators or a direct resonator-to-resonator coupling [95,96,98]. Another physical realization may be based on atoms in optical cavities [94].

VI. CONCLUSIONS

The work investigated the performance of two types of a quantum engine based on a single qubit: the Otto engine and the three-stroke engine. Interaction with the thermal baths has been described within the resource-theoretic framework by means of thermal operations, i.e., the energy-preserving unitaries acting on the system and its environment. Within this approach the fundamental limitations of heat engines employing the heat strokes based on Markovian thermalization could have been analyzed. It was shown that the Otto engine employing non-Markovian dynamics can generate more work during a single cycle then its Markovian counterpart; at the same time, both setups produce more work than the three-stroke engine, which can operate only in the non-Markovian regime. It was also demonstrated that the non-Markovian Otto engine can generate less work fluctuations than the Markovian one. Furthermore, even less fluctuations are produced by the three-stroke engine, albeit only for a low power; therefore, the relative advantage of the non-Markovian Otto engine and the three-stroke engine depends on whether the maximization of power or minimization of fluctuations is desired. Additionally, the reduction of work fluctuations in the three-stroke engine

has been demonstrated to be a result of correlations of the work generated in the successive cycles; a similar behavior has been previously investigated in the classical case [43].

On a more general level, in the spirit of some previous works [52,63], the paper confirms the particular usefulness of resource-theoretic methods for demonstrating a genuine beneficial influence of non-Markovianity on the performance of quantum nanodevices. The main advantage of this approach is its generality: it provides fundamental constraints on Markovian evolution, without assuming any specific model of microscopic dynamics. This may inspire future studies on applications of non-Markovianity in other areas of quantum technology.

ACKNOWLEDGMENTS

APPENDIX: EMBEDDABILITY OF SINGLE-QUBIT MARKOVIAN THERMAL OPERATIONS

For pedagogical purposes, let me here briefly demonstrate that every Markovian operation acting on a single qubit is also embeddable, i.e., can be realized using a time-independent Lindblad generator

$$\mathcal{L}\rho(t) = \sum_{i=1}^{4} r_i \left\{ L_i \rho(t) L_i^\dagger - \frac{1}{2}[L_i^\dagger L_i, \rho(t)] \right\}, \qquad (A1)$$

with $L_1 = |g\rangle\langle e|$, $L_2 = |e\rangle\langle g|$, $L_3 = |g\rangle\langle g|$, $L_4 = |e\rangle\langle e|$, $r_1 = \Gamma p_g^{\text{th}}$, $r_2 = \Gamma(1 - p_g^{\text{th}})$, and $r_3 = r_4 = D$; here Γ and D are the relaxation and the pure dephasing rates, respectively. The elements of the density matrix evolve then as

$$\rho_{gg}(t) = p_g(t) = p_g(0)e^{-\Gamma t} + (1 - e^{-\Gamma t})p_g^{\text{th}}, \qquad (A2)$$

$$\rho_{ee}(t) = p_e(t) = 1 - p_g(t), \qquad (A3)$$

$$\rho_{ge}(t) = \rho_{eg}^*(t) = \rho_{ge}(0)e^{-(\Gamma/2+D)t}. \qquad (A4)$$

At the same time, a thermal operation transforms the density matrix elements as [Eqs. (11) and (12)]

$$p_g(t) = p_g(0)\left(1 - \frac{\lambda}{p_g^{\text{th}}}\right) + \lambda, \qquad (A5)$$

$$\rho_{ge}(t) = \gamma \rho_{ge}(0). \qquad (A6)$$

As one can easily note, in the Markovian regime of $\lambda \leqslant p_g^{\text{th}}$ the parameters of a thermal operation can be expressed as $\lambda = (1 - e^{-\Gamma t})p_g^{\text{th}}$ and $\gamma = e^{-(\Gamma/2+D)t}$, which demonstrates the embeddability.

[1] H. T. Quan, Y.-X. Liu, C. P. Sun, and F. Nori, Quantum thermodynamic cycles and quantum heat engines, Phys. Rev. E **76**, 031105 (2007).

[2] R. Kosloff and A. Levy, Quantum heat engines and refrigerators: Continuous devices, Annu. Rev. Phys. Chem. **65**, 365 (2014).

[3] N. M. Myers, O. Abah, and S. Deffner, Quantum thermodynamic devices: From theoretical proposals to experimental reality, AVS Quantum Sci. **4**, 027101 (2022).

[4] A. E. Allahverdyan, R. Balian, and T. M. Nieuwenhuizen, Maximal work extraction from finite quantum systems, Europhys. Lett. **67**, 565 (2004).

[5] P. Skrzypczyk, A. J. Short, and S. Popescu, Work extraction and thermodynamics for individual quantum systems, Nat. Commun. **5**, 4185 (2014).

[6] Y.-H. Ma, D. Xu, H. Dong, and C.-P. Sun, Optimal operating protocol to achieve efficiency at maximum power of heat engines, Phys. Rev. E **98**, 022133 (2018).

[7] P. A. Erdman, V. Cavina, R. Fazio, F. Taddei, and V. Giovannetti, Maximum power and corresponding efficiency for two-level heat engines and refrigerators: Optimality of fast cycles, New J. Phys. **21**, 103049 (2019).

[8] E. Bäumer, M. Perarnau-Llobet, P. Kammerlander, H. Wilming, and R. Renner, Imperfect thermalizations allow for optimal thermodynamic processes, Quantum **3**, 153 (2019).

[9] O. Raz, Y. Subaşı, and R. Pugatch, Geometric Heat Engines Featuring Power That Grows with Efficiency, Phys. Rev. Lett. **116**, 160601 (2016).

[10] K. Brandner and K. Saito, Thermodynamic Geometry of Microscopic Heat Engines, Phys. Rev. Lett. **124**, 040602 (2020).

[11] P. Abiuso, H. J. D. Miller, M. Perarnau-Llobet, and M. Scandi, Geometric optimisation of quantum thermodynamic processes, Entropy **22**, 1076 (2020).

[12] T. Feldmann and R. Kosloff, Quantum four-stroke heat engine: Thermodynamic observables in a model with intrinsic friction, Phys. Rev. E **68**, 016101 (2003).

[13] M. O. Scully, Quantum Photocell: Using Quantum Coherence to Reduce Radiative Recombination and Increase Efficiency, Phys. Rev. Lett. **104**, 207701 (2010).

[14] M. O. Scully, K. R. Chapin, K. E. Dorfman, M. B. Kim, and A. Svidzinsky, Quantum heat engine power can be increased by noise-induced coherence, Proc. Natl. Acad. Sci. USA **108**, 15097 (2011).

[15] U. Harbola, S. Rahav, and S. Mukamel, Quantum heat engines: A thermodynamic analysis of power and efficiency, Europhys. Lett. **99**, 50005 (2012).

[16] K. Brandner, M. Bauer, M. T. Schmid, and U. Seifert, Coherence-enhanced efficiency of feedback-driven quantum engines, New J. Phys. **17**, 065006 (2015).

[17] B. Karimi and J. P. Pekola, Otto refrigerator based on a superconducting qubit: Classical and quantum performance, Phys. Rev. B **94**, 184503 (2016).

[18] K. Brandner and U. Seifert, Periodic thermodynamics of open quantum systems, Phys. Rev. E **93**, 062134 (2016).

[19] F. Chen, Y. Gao, and M. Galperin, Molecular heat engines: Quantum coherence effects, Entropy **19**, 472 (2017).

[20] A. Ü. C. Hardal, N. Aslan, C. M. Wilson, and Ö. E. Müstecaplıoğlu, Quantum heat engine with coupled superconducting resonators, Phys. Rev. E **96**, 062120 (2017).

[21] A. Roulet, S. Nimmrichter, J. M. Arrazola, S. Seah, and V. Scarani, Autonomous rotor heat engine, Phys. Rev. E **95**, 062131 (2017).

[22] K. Brandner, M. Bauer, and U. Seifert, Universal Coherence-Induced Power Losses of Quantum Heat Engines in Linear Response, Phys. Rev. Lett. **119**, 170602 (2017).

[23] J. Bengtsson, M. Nilsson Tengstrand, A. Wacker, P. Samuelsson, M. Ueda, H. Linke, and S. M. Reimann, Quantum Szilard Engine with Attractively Interacting Bosons, Phys. Rev. Lett. **120**, 100601 (2018).

[24] J. O. González, J. P. Palao, D. Alonso, and L. A. Correa, Classical emulation of quantum-coherent thermal machines, Phys. Rev. E **99**, 062102 (2019).

[25] V. Holubec and T. Novotný, Effects of noise-induced coherence on the fluctuations of current in quantum absorption refrigerators, J. Chem. Phys. **151**, 044108 (2019).

[26] J. Klatzow, J. N. Becker, P. M. Ledingham, C. Weinzetl, K. T. Kaczmarek, D. J. Saunders, J. Nunn, I. A. Walmsley, R. Uzdin, and E. Poem, Experimental Demonstration of Quantum Effects in the Operation of Microscopic Heat Engines, Phys. Rev. Lett. **122**, 110601 (2019).

[27] K. Ptaszyński, Coherence-enhanced constancy of a quantum thermoelectric generator, Phys. Rev. B **98**, 085425 (2018).

[28] A. Rignon-Bret, G. Guarnieri, J. Goold, and M. T. Mitchison, Thermodynamics of precision in quantum nanomachines, Phys. Rev. E **103**, 012133 (2021).

[29] J. Liu and D. Segal, Coherences and the thermodynamic uncertainty relation: Insights from quantum absorption refrigerators, Phys. Rev. E **103**, 032138 (2021).

[30] N. M. Myers, O. Abah, and S. Deffner, Quantum Otto engines at relativistic energies, New J. Phys. **23**, 105001 (2021).

[31] A. Mu, B. K. Agarwalla, G. Schaller, and D. Segal, Qubit absorption refrigerator at strong coupling, New J. Phys. **19**, 123034 (2017).

[32] G. Verley, M. Esposito, T. Willaert, and C. Van den Broeck, The unlikely Carnot efficiency, Nat. Commun. **5**, 4721 (2014).

[33] G. Verley, T. Willaert, C. Van den Broeck, and M. Esposito, Universal theory of efficiency fluctuations, Phys. Rev. E **90**, 052145 (2014).

[34] K. Proesmans and C. Van den Broeck, Stochastic efficiency: Five case studies, New J. Phys. **17**, 065004 (2015).

[35] J.-M. Park, H.-M. Chun, and J. D. Noh, Efficiency at maximum power and efficiency fluctuations in a linear Brownian heat-engine model, Phys. Rev. E **94**, 012127 (2016).

[36] V. Holubec and A. Ryabov, Work and power fluctuations in a critical heat engine, Phys. Rev. E **96**, 030102(R) (2017).

[37] V. Holubec and A. Ryabov, Cycling Tames Power Fluctuations near Optimum Efficiency, Phys. Rev. Lett. **121**, 120601 (2018).

[38] S. K. Manikandan, L. Dabelow, R. Eichhorn, and S. Krishnamurthy, Efficiency Fluctuations in Microscopic Machines, Phys. Rev. Lett. **122**, 140601 (2019).

[39] K. Ito, C. Jiang, and G. Watanabe, Universal bounds for fluctuations in small heat engines, arXiv:1910.08096.

[40] H. Vroylandt, M. Esposito, and G. Verley, Efficiency Fluctuations of Stochastic Machines Undergoing a Phase Transition, Phys. Rev. Lett. **124**, 250603 (2020).

[41] S. Saryal, M. Gerry, I. Khait, D. Segal, and B. K. Agarwalla, Universal Bounds on Fluctuations in Continuous Thermal Machines, Phys. Rev. Lett. **127**, 190603 (2021).

[42] V. Holubec and A. Ryabov, Fluctuations in heat engines, J. Phys. A: Math. Theor. **55**, 013001 (2021).

[43] G.-H. Xu and G. Watanabe, Correlation-enhanced stability of microscopic cyclic heat engines, arXiv:2111.09508.

[44] G. Watanabe and Y. Minami, Finite-time thermodynamics of fluctuations in microscopic heat engines, Phys. Rev. Research **4**, L012008 (2022).

[45] N. A. Sinitsyn, Fluctuation relation for heat engines, J. Phys. A: Math. Theor. **44**, 405001 (2011).

[46] M. Campisi, Fluctuation relation for quantum heat engines and refrigerators, J. Phys. A: Math. Theor. **47**, 245001 (2014).

[47] V. Holubec, An exactly solvable model of a stochastic heat engine: Optimization of power, power fluctuations and efficiency, J. Stat. Mech. (2014) P05022.

[48] P. Pietzonka and U. Seifert, Universal Trade-Off Between Power, Efficiency, and Constancy in Steady-State Heat Engines, Phys. Rev. Lett. **120**, 190602 (2018).

[49] J. Roßnagel, S. T. Dawkins, K. N. Tolazzi, O. Abah, E. Lutz, F. Schmidt-Kaler, and K. Singer, A single-atom heat engine, Science **352**, 325 (2016).

[50] D. von Lindenfels, O. Gräb, C. T. Schmiegelow, V. Kaushal, J. Schulz, M. T. Mitchison, J. Goold, F. Schmidt-Kaler, and U. G. Poschinger, Spin Heat Engine Coupled to a Harmonic-Oscillator Flywheel, Phys. Rev. Lett. **123**, 080602 (2019).

[51] J. P. S. Peterson, T. B. Batalhão, M. Herrera, A. M. Souza, R. S. Sarthour, I. S. Oliveira, and R. M. Serra, Experimental Characterization of a Spin Quantum Heat Engine, Phys. Rev. Lett. **123**, 240601 (2019).

[52] Á. M. Alhambra, M. Lostaglio, and C. Perry, Heat-bath algorithmic cooling with optimal thermalization strategies, Quantum **3**, 188 (2019).

[53] A. Solfanelli, A. Santini, and M. Campisi, Quantum thermodynamic methods to refrigerate a qubit on a quantum processing unit, AVS Quant. Sci. **4**, 026802 (2022).

[54] H.-P. Breuer and F. Petruccione, *The Theory of Open Quantum Systems* (Oxford University Press, Oxford, UK, 2002).

[55] Á. Rivas, S. F. Huelga, and M. B. Plenio, Quantum non-Markovianity: Characterization, quantification and detection, Rep. Prog. Phys. **77**, 094001 (2014).

[56] S. Wenderoth, H.-P. Breuer, and M. Thoss, Non-Markovian effects in the spin-boson model at zero temperature, Phys. Rev. A **104**, 012213 (2021).

[57] R. Vasile, S. Olivares, M. G. A. Paris, and S. Maniscalco, Continuous-variable quantum key distribution in non-Markovian channels, Phys. Rev. A **83**, 042321 (2011).

[58] B. Bylicka, D. Chruściński, and S. Maniscalco, Non-Markovianity and reservoir memory of quantum channels: A quantum information theory perspective, Sci. Rep. **4**, 5720 (2014).

[59] A. W. Chin, S. F. Huelga, and M. B. Plenio, Quantum Metrology in Non-Markovian Environments, Phys. Rev. Lett. **109**, 233601 (2012).

[60] W.-Z. Zhang, Y. Han, B. Xiong, and L. Zhou, Optomechanical force sensor in a non-Markovian regime, New J. Phys. **19**, 083022 (2017).

[61] S. F. Huelga, Á. Rivas, and M. B. Plenio, Non-Markovianity-Assisted Steady State Entanglement, Phys. Rev. Lett. **108**, 160402 (2012).

[62] D. Heineken, K. Beyer, K. Luoma, and W. T. Strunz, Quantum-memory-enhanced dissipative entanglement creation in nonequilibrium steady states, Phys. Rev. A **104**, 052426 (2021).

[63] G. Spaventa, S. F. Huelga, and M. B. Plenio, Capacity of non-Markovianity to boost the efficiency of molecular switches, Phys. Rev. A **105**, 012420 (2022).

[64] X. Y. Zhang, X. L. Huang, and X. X. Yi, Quantum Otto heat engine with a non-Markovian reservoir, J. Phys. A: Math. Theor. **47**, 455002 (2014).

[65] G. Thomas, N. Siddharth, S. Banerjee, and S. Ghosh, Thermodynamics of non-Markovian reservoirs and heat engines, Phys. Rev. E **97**, 062108 (2018).

[66] M. Wiedmann, J. T. Stockburger, and J. Ankerhold, Non-Markovian dynamics of a quantum heat engine: Out-of-equilibrium operation and thermal coupling control, New J. Phys. **22**, 033007 (2020).

[67] Y. Shirai, K. Hashimoto, R. Tezuka, C. Uchiyama, and N. Hatano, Non-Markovian effect on quantum Otto engine: Role of system-reservoir interaction, Phys. Rev. Research **3**, 023078 (2021).

[68] P. Abiuso and V. Giovannetti, Non-Markov enhancement of maximum power for quantum thermal machines, Phys. Rev. A **99**, 052106 (2019).

[69] P. A. Camati, J. F. G. Santos, and R. M. Serra, Employing non-Markovian effects to improve the performance of a quantum Otto refrigerator, Phys. Rev. A **102**, 012217 (2020).

[70] V. Mukherjee, A. G. Kofman, and G. Kurizki, Anti-Zeno quantum advantage in fast-driven heat machines, Commun. Phys. **3**, 8 (2020).

[71] A. Das and V. Mukherjee, Quantum-enhanced finite-time Otto cycle, Phys. Rev. Research **2**, 033083 (2020).

[72] M. Xu, J. T. Stockburger, G. Kurizki, and J. Ankerhold, Minimal quantum thermal machine in a bandgap environment: Non-Markovian features and anti-Zeno advantage, New J. Phys. **24**, 035003 (2022).

[73] M. Horodecki and J. Oppenheim, Fundamental limitations for quantum and nanoscale thermodynamics, Nat. Commun. **4**, 2059 (2013).

[74] F. G. S. L. Brandão, M. Horodecki, J. Oppenheim, J. M. Renes, and R. W. Spekkens, Resource Theory of Quantum States Out of Thermal Equilibrium, Phys. Rev. Lett. **111**, 250404 (2013).

[75] F. G. S. L. Brandão, M. Horodecki, N. H. Y. Ng, J. Oppenheim, and S. Wehner, The second laws of quantum thermodynamics, Proc. Natl. Acad. Sci. USA **112**, 3275 (2015).

[76] J. Goold, M. Huber, A. Riera, L. del Rio, and P. Skrzypczyk, The role of quantum information in thermodynamics—A topical review, J. Phys. A: Math. Theor. **49**, 143001 (2016).

[77] M. Lostaglio, An introductory review of the resource theory approach to thermodynamics, Rep. Prog. Phys. **82**, 114001 (2019).

[78] T. Feldmann and R. Kosloff, Performance of discrete heat engines and heat pumps in finite time, Phys. Rev. E **61**, 4774 (2000).

[79] M. Łobejko, P. Mazurek, and M. Horodecki, Thermodynamics of minimal coupling quantum heat engines, Quantum **4**, 375 (2020).

[80] M. Lostaglio and K. Korzekwa, Continuous thermomajorization and a complete set of laws for Markovian thermal processes, arXiv:2111.12130.

[81] L. Li, M. J. W. Hall, and H. M. Wiseman, Concepts of quantum non-Markovianity: A hierarchy, Phys. Rep. **759**, 1 (2018).

[82] S. Milz and K. Modi, Quantum stochastic processes and quantum non-Markovian phenomena, PRX Quant. **2**, 030201 (2021).

[83] P. Talkner and P. Hänggi, Colloquium: Statistical mechanics and thermodynamics at strong coupling: Quantum and classical, Rev. Mod. Phys. **92**, 041002 (2020).

[84] A. S. Trushechkin, M. Merkli, J. D. Cresser, and J. Anders, Open quantum system dynamics and the mean force Gibbs state, AVS Quant. Sci. **4**, 012301 (2022).

[85] P. Ćwikliński, M. Studziński, M. Horodecki, and J. Oppenheim, Limitations on the Evolution of Quantum Coherences: Towards Fully Quantum Second Laws of Thermodynamics, Phys. Rev. Lett. **115**, 210403 (2015).

[86] M. Lostaglio, Á. M. Alhambra, and C. Perry, Elementary thermal operations, Quantum **2**, 52 (2018).

[87] J. Kurchan, A quantum fluctuation theorem, arXiv:cond-mat/0007360.

[88] T. Monnai, Unified treatment of the quantum fluctuation theorem and the Jarzynski equality in terms of microscopic reversibility, Phys. Rev. E **72**, 027102 (2005).

[89] M. Campisi, P. Hänggi, and P. Talkner, Colloquium: Quantum fluctuation relations: Foundations and applications, Rev. Mod. Phys. **83**, 771 (2011).

[90] H. Touchette, The large deviation approach to statistical mechanics, Phys. Rep. **478**, 1 (2009).

[91] D. Chiuchiù and S. Pigolotti, Mapping of uncertainty relations between continuous and discrete time, Phys. Rev. E **97**, 032109 (2018).

[92] M. H. Naderi, M. Soltanolkotabi, and R. Roknizadeh, A theoretical scheme for generation of nonlinear coherent states in a micromaser under intensity-dependent Jaynes-Cummings model, Eur. Phys. J. D **32**, 397 (2005).

[93] J. Åberg, Catalytic Coherence, Phys. Rev. Lett. **113**, 150402 (2014).

[94] M. L. Bera, S. Julià-Farré, M. Lewenstein, and M. N. Bera, Quantum heat engines with Carnot efficiency at maximum power, Phys. Rev. Research **4**, 013157 (2022).

[95] A. Guthrie, C. D. Satrya, Y.-C. Chang, P. Menczel, F. Nori, and J. P. Pekola, A Cooper-Pair Box Architecture for Cyclic Quantum Heat Engines, Phys. Rev. Appl. **17**, 064022 (2022).

[96] A. Ronzani, B. Karimi, J. Senior, Y.-C. Chang, J. T. Peltonen, C. Chen, and J. P. Pekola, Tunable photonic heat transport in a quantum heat valve, Nat. Phys. **14**, 991 (2018).

[97] J. Senior, A. Gubaydullin, B. Karimi, J. T. Peltonen, J. Ankerhold, and J. P. Pekola, Heat rectification via a superconducting artificial atom, Commun. Phys. **3**, 40 (2020).

[98] M. Xu, J. T. Stockburger, and J. Ankerhold, Heat transport through a superconducting artificial atom, Phys. Rev. B **103**, 104304 (2021).