Stochastic heat engines beyond a unique definition of temperature

Chia-Hao Chang,[1] Chia-Jung Chang,[1] Nian-Jhu Wu,[1] Yonggun Jun,[2,*] and Cheng-Hung Chang[1,3,†]
[1]*Institute of Physics, National Yang Ming Chiao Tung University, Hsinchu 300, Taiwan*
[2]*Department of Physics and Center of Complex Systems, National Central University, Taoyuan 320, Taiwan*
[3]*Physics Division, National Center for Theoretical Sciences, Taipei 10617, Taiwan*

(Received 9 February 2023; accepted 11 October 2023; published 26 October 2023)

When Carnot and Stirling initially conceptualized heat engines, temperature unambiguously represented our everyday perception of cold and hot. However, as this energy scale is expanded to measure the strength of noise in general nonequilibrium heat baths, such as those consisting of bacteria or active particles, it takes on different definitions and connotations. This raises a fundamental question of whether and how thermodynamic conclusions beyond a unique definition of temperature would deviate from our conventional understanding. To address this inquiry, we investigate a colloidal Stirling engine governed by a large number of stochastic dynamical systems. Within experimentally accessible parameter values, we discover certain exceptional active engines that can outperform their passive counterpart, as notably claimed in a recent experiment involving a bacterial bath. Our analysis shows that such heightened performance can be attributed to either a restoring effect in noise or a significant dissipation kernel. The revealed influence of active baths on Stirling efficiency provides further insights into their impact on maximum power output, Carnot efficiency, and Curzon-Alhborn efficiency. The finding elucidates the origins of exceptional performance in stochastic heat engines, offers strategies for harvesting energies from active noises, and sheds light on the effects of nonequilibrium temperature in stochastic thermodynamics.

DOI: 10.1103/PhysRevResearch.5.043085

I. INTRODUCTION

Temperature is a concept whose meaning can be as comprehensible as that in a fairy tale or as complex as that in nonequilibrium physics. In the latter case, the interpretation of temperature is usually not unique, which poses difficulties in defining and understanding the nonequilibrium temperatures. This subtle issue has emerged in various subfields of physics, including glasses, sheared fluids, granular materials, amorphous semiconductors, turbulent fluids, chaotic systems, nuclear materials, and nanoscale systems [1–3]. Recently, it has also entered stochastic thermodynamics, where a mystery is whether and why, if any, a stochastic heat engine immersed in an active heat bath can outperform its passive counterpart. This concern arises from a noteworthy experimental observation conducted in a bacterial heat bath [4]. It raises an even more fundamental question of how to imagine the temperatures of an active bath that violate the fluctuation-dissipation theorem and what the consequences are of adopting different definitions. If we acknowledge that the applicability of thermodynamics, such as the concept of heat engine, is not limited to passive baths, then we are compelled to confront this elusive temperature or, more rigorously, this temperature-related energy scale. In contrast to the aforementioned subfields of physics, the effects of such nonequilibrium temperature are much less known in stochastic thermodynamics. This paper carries out a comprehensive survey across a wide range of systems, with the hope of filling the knowledge gap.

To this end, we consider a stochastic heat engine operating under cyclic environmental changes. These changes are typically represented by different protocols in different parameter spaces, including spaces that differ in the definition of temperature. Here, the two most widely used definitions of temperature are adopted to deduce the explicit transformation between different protocol representations. Under these representations, we analyze how the heat, work, and efficiency functionals depend on the trajectories of the engine state (Fig. 1). These trajectories are generated by a broad collection of dynamical systems, ranging from solvable theoretical models to real experimental systems. For each system, we provide experimentally accessible parameter values for attaining high-performance engines. The results summarized from these systems clarify several paradoxical arguments regarding high engine efficiency and reveal principles for achieving it. Some of those unveiled mechanisms may have played a role in existing experiments, while others could be encountered in general systems under cyclic variations of active heat baths. These results further indicate the relation between the nonequilibrium temperature of active heat baths and other measures of engine performance, like the Carnot efficiency, the maximum power of the engine, and the Curzon-Alhborn efficiency.

*yonggun@phy.ncu.edu.tw
†chchang@nycu.edu.tw

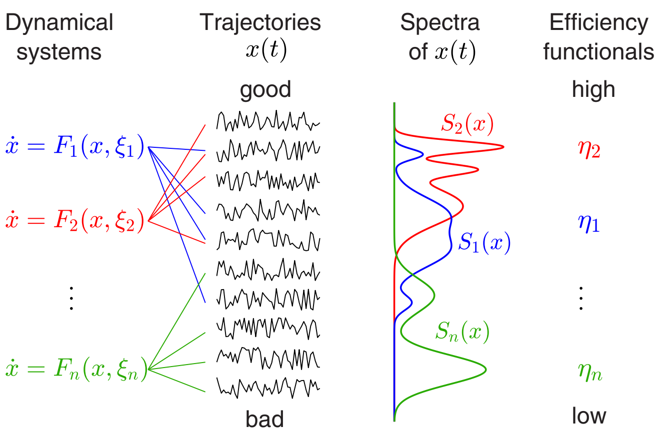

FIG. 1. The state x of a stochastic heat engine indexed by i follows a trajectory of a stochastic dynamical system $\dot{x} = F_i(x, \xi_i)$. This nonautonomous system has a cyclic variation in its parameters and the statistics of noise ξ_i. Each trajectory $x(t)$ is assigned with a functional of heat and work (Appendix A). Order the trajectories from good to bad according to whether their stochastic efficiency $\tilde{\eta}$ is high or low, which is the ratio of the work done to the heat input along a trajectory [35]. Then all trajectories of the ith system $\dot{x} = F_i(x, \xi_i)$ form a spectrum $S_i(x)$. The macroscopic efficiency η_i of the heat engine governed by $F_i(x, \xi_i)$ would be high if $S_i(x)$ contains more good trajectories. Criteria for generating such spectra are summarized from several $F_i(x, \xi_i)$ in Sec. III and discussed in Sec. IV. More details about $\tilde{\eta}$ can be found in Sec. V.

A paradigmatic example of the stochastic heat engine is made of a colloidal particle under a cyclic variation of temperature and confining potential [5–7]. If this engine is a Stirling engine, its protocol contains four processes: Isothermal compression, iso-stiffness heating, isothermal expansion, and iso-stiffness cooling (Figs. 2 and the text below Eq. (A5)). When heating and cooling are performed by changing the noise strength of a nonequilibrium bacterial heat bath, the efficiency of the engine is experimentally shown to be higher than that of its equilibrium counterpart [4]. This phenomenon was attributed to non-Gaussian statistics of particle displacement in that experiment. However, subsequent theoretical studies explained that neither non-Gaussian nor persistent noise is crucial for high efficiency [8–10]. So far, whether and under which conditions a nonequilibrium engine can surpass the efficiency of its passive version remains unsettled.

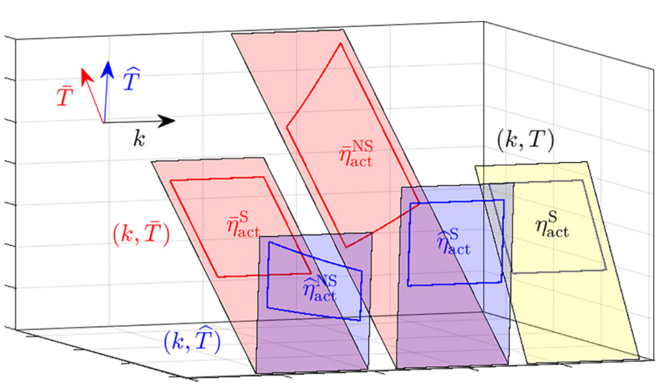

FIG. 2. The protocol of a stochastic engine can be represented by two distinct loops in two different spaces $(k, \hat{T})$ (blue) and $(k, \bar{T})$ (red) when the noise is active. These two spaces will converge to a common space (k, T) (yellow) with $T = \hat{T} = \bar{T}$ when the active noise reduces to a passive one. Here, the blue loop of efficiency $\hat{\eta}^{\mathrm{NS}}_{\mathrm{act}}$ ($\hat{\eta}^{\mathrm{S}}_{\mathrm{act}}$) in $(k, \hat{T})$ is converted into the red loop of efficiency $\bar{\eta}^{\mathrm{S}}_{\mathrm{act}}$ ($\bar{\eta}^{\mathrm{NS}}_{\mathrm{act}}$) in $(k, \bar{T})$ by $f(k) = \hat{T}/\bar{T} = (1 + k\tau/\gamma)^{-1}$, which is the TR of a colloidal heat engine subject to an Ornstein-Uhlenbeck noise, as shown before Eq. (4).

In the colloidal engine described above, there are two commonly used temperatures. One is defined by the diffusion of a free colloidal particle through the Einstein relation, denoted by $\bar{T}$ [1,3,11]. The other is by the particle variance in the equipartition theorem, denoted by $\hat{T}$ [1,3,5,8,12,13]. Suppose that this Brownian particle follows an overdamped Langevin equation

$$\gamma \dot{x} = -\partial_x V + \xi, \tag{1}$$

where x is the position of the particle, γ denotes the friction coefficient, ξ represents the random noise of a heat bath, and $V = kx^2/2$ stands for a harmonic potential of stiffness k. Then $\hat{T} = k\langle x^2\rangle/k_{\mathrm{B}}$, with k_{B} the Boltzmann constant, and $\bar{T} = \gamma D/k_{\mathrm{B}}$, with D the diffusion constant of the particle in a potential-free space. For an equilibrium heat bath, the definitions of $\hat{T}$ and $\bar{T}$ yield the same value, which corresponds to the thermal temperature T. However, for a nonequilibrium heat bath, the definitions of $\hat{T}$ and $\bar{T}$ result in different values. Therefore, we cannot have the equality relation $T = \hat{T} = \bar{T}$, and it becomes necessary to choose either $\hat{T}$ or $\bar{T}$, or other energy scales, when a discussion requires addressing the concept of temperature. In this case, both $\hat{T}$ and $\bar{T}$ are not the bona fide temperature in equilibrium physics. A heat bath or an engine is called passive (active) when the noise comes from an equilibrium (nonequilibrium) bath. In an active bath, such as that in Ref. [14], x is not necessarily Boltzmannian distributed, and the fluctuation-dissipation theorem may not hold. An exception is a non-Gaussian white active noise with $\hat{T} = \bar{T}$ [8], which extends the equipartition relation to a nonequilibrium context.

II. EFFICIENCY FUNCTIONALS IN DISTINCT (k, T) SPACES

Given a cyclic change in confinement and noise strength, an engine protocol can be plotted as a loop in the $(k, \hat{T})$ or $(k, \bar{T})$ space. A loop in $(k, \hat{T})$ can be converted into another loop in $(k, \bar{T})$ by a given temperature ratio (TR) $f(k) \equiv \hat{T}/\bar{T} = k\langle x^2\rangle/(k_{\mathrm{B}}\bar{T})$ and vice versa (Fig. 2 and Appendix A). This function quantifies the difference between $\hat{T}$ and $\bar{T}$, or equivalently the deviation of $k\langle x^2\rangle$ from $k_{\mathrm{B}}\bar{T}$. If an engine is a Stirling one in $(k, \bar{T})$, it will have a rectangular loop in that space and generally a nonrectangular loop in $(k, \hat{T})$, and vice versa. The efficiency of an active Stirling cycle $\bar{\eta}^{\mathrm{S}}_{\mathrm{act}}$ in $(k, \bar{T})$ [$\hat{\eta}^{\mathrm{S}}_{\mathrm{act}}$ in $(k, \hat{T})$] is then a function of the four corner positions of the rectangular loop. In contrast, the efficiency of an active non-Stirling cycle $\bar{\eta}^{\mathrm{NS}}_{\mathrm{act}}$ in $(k, \bar{T})$ [$\hat{\eta}^{\mathrm{NS}}_{\mathrm{act}}$ in $(k, \hat{T})$] depends on all points in the loop (Fig. 2). Substituting $\hat{T} = \bar{T} f(k)$ into formula $\hat{\eta}^{\mathrm{S}}_{\mathrm{act}}$ merely reexpresses the efficiency of a Stirling cycle specified by four rectangular corner points in $(k, \hat{T})$ in terms of other four nonrectangular corner points in $(k, \bar{T})$. Therefore, the new formula can generally not be the efficiency of a Stirling cycle $\bar{\eta}^{\rm S}_{\rm act}$ in $(k, \bar{T})$ [see more below (A5)]. When the active noise is reduced to a passive noise, then $f(k) \to 1$ and $(k, \bar{T})$ and $(k, \hat{T})$ will degenerate to an identical space (k, T) with $T = \hat{T} = \bar{T}$. In this limit, the two different loops represented in $(k, \hat{T})$ and $(k, \bar{T})$ will converge to the same loop in (k, T). If an active Stirling engine outperforms its passive counterpart, then $\hat{\eta}^{\rm S}_{\rm act} > \eta^{\rm S}_{\rm pas}$ ($\bar{\eta}^{\rm S}_{\rm act} > \eta^{\rm S}_{\rm pas}$), where the comparison is between two identical rectangular loops in $(k, \hat{T})$ and (k, T) [in $(k, \bar{T})$ and (k, T)].

Accumulative stochastic heat Q and work W along a stochastic trajectory governed by Eq. (1) can be evaluated using the stochastic energetics approach of Sekimoto [15]. In a quasistatic process, their ensemble averages along a process from state i to f are $\langle Q\rangle = \langle\int_i^f kxdx\rangle = (1/2)(k_f\langle x^2\rangle_f - k_i\langle x^2\rangle_i) - (1/2)\int_i^f \langle x^2\rangle dk$ and $\langle W\rangle = (1/2)\int_i^f \langle x^2\rangle dk$, respectively. Here, k_i (k_f) is the stiffness in the state i (f). If the temperature along a trajectory is understood as $\bar{T}$, the efficiency of an active Stirling engine is (Appendix A)

$$\bar{\eta}^{\rm S}_{\rm act} = \frac{(\bar{T}_1 - \bar{T}_2)[G(k_1) - G(k_2)]}{\bar{T}_1 f(k_2) - \bar{T}_2 f(k_1) + \bar{T}_1[G(k_1) - G(k_2)]}, \quad (2)$$

with $\bar{T}_1$ ($\bar{T}_2$) the maximum (minimum) temperature and k_1 (k_2) the maximum (minimum) stiffness of the Stirling cycle in the $(k, \bar{T})$ space. Here, $f(k)$ is the TR defined above and $G(k)$ is its derivative $G'(k) = f(k)/k$. When $f(k) = 1$ for all k, it yields $G(k) = \ln k$, for which Eq. (2) reduces to the efficiency $\eta^{\rm S}_{\rm pas}$ of a passive Stirling engine subjected to an equilibrium bath or, exceptionally, of an active Stirling engine immersed in the non-Gaussian white noise mentioned below Eq. (1). On the contrary, if the temperature in the protocol refers to $\hat{T}$, its efficiency $\hat{\eta}^{\rm S}_{\rm act}$ is always equal to $\eta^{\rm S}_{\rm pas}$ of the passive engine [8]. The condition for $\bar{\eta}^{\rm S}_{\rm act} > \eta^{\rm S}_{\rm pas}$ is (Appendix A)

$$\int_{k_2}^{k_1} \frac{1}{k}\left\{\frac{\bar{T}_2}{\bar{T}_1}[f(k_1) - f(k)] + [f(k) - f(k_2)]\right\}dk > 0. \quad (3)$$

A sufficient but not necessary condition for Eq. (3) is "a monotonically increasing $f(k)$" within the operating range $[k_2, k_1]$ of the engine. However, $f(k) \sim k\langle x^2\rangle$ should not increase too fast to lead to an increase of $\langle x^2\rangle$ with k because then a stronger confinement would counterintuitively give a larger particle fluctuation (Appendix A). Therefore, the above monotonic increasing condition can be sharpened to $f(k) \sim k^\alpha$, with $0 < \alpha < 1$. The larger the nonzero value of Eq. (3), the higher is $\bar{\eta}^{\rm S}_{\rm act}$, which, for example, can be achieved by enhancing the ratio $\bar{T}_2/\bar{T}_1$ in Eq. (3).

III. ENGINE TRAJECTORIES GENERATED BY VARIOUS DYNAMICAL SYSTEMS

A. Cosine persistent noise

As a simple example, consider the persistent noise $\xi(t)$ with a correlation $\langle\xi(t)\xi(t')\rangle = (\gamma k_{\rm B}\bar{T}/\tau)\exp(-|t - t'|/\tau)$, where τ is the correlation time. Several real systems share this property, including the Ornstein-Uhlenbeck noise, the run-and-tumble noise, and the active Brownian noise [8]. For this $\xi(t)$, $f(k) = (1 + k\tau/\gamma)^{-1} = (1 + K)^{-1} \equiv \hat{f}(K)$, where $K \equiv k\tau/\gamma$ is a dimensionless stiffness (Appendix B). Since this $f(k)$, or equivalently $\hat{f}(K)$, is a decreasing function, the persistent noise cannot give a high engine efficiency [8]. In the limit $\tau \to 0$, this noise returns to the white noise and $\bar{\eta}^{\rm S}_{\rm act}$ converges to $\eta^{\rm S}_{\rm pas}$.

A common feature of active noises is the enhanced contribution of low-frequency components in their spectrum because the persistent motion of self-propelled active particles in those noises is generically longer than that of water molecules in the thermal noise. The simplest way to embody this effect might be multiplying a cosine function of a slow angular frequency $\omega_0 \geqslant 0$ to the correlation of the persistent noise,

$$\langle\xi(t)\xi(t')\rangle = \frac{\gamma\left(1 + \omega_0^2\tau^2\right)k_{\rm B}\bar{T}}{\tau}e^{-|t-t'|/\tau}\cos[\omega_0(t - t')]. \quad (4)$$

Such $\xi(t)$ provides some restoring effect on the Brownian particle, as the oscillating polar forces of the hydrodynamic interaction in bacterial systems [16]. For this noise (Appendix B)

$$\hat{f}(K) = \frac{(c + 1)(K + 1)}{(K + 1)^2 + c}, \quad (5)$$

where $c \equiv \omega_0^2\tau^2$ and K is as defined above. $\hat{f}(K)$ is an increasing function within the range $K \in [0, \sqrt{c} - 1]$. At $\omega_0 = 0$, Eqs. (4) and (5) reduce to those of the above persistent noise. Intriguingly, the destructive role of τ in obtaining an increasing $\hat{f}(K)$ at $\omega_0 = 0$ becomes constructive when $\omega_0 > 0$.

The temporal correlation function in Eq. (4) has a similar mathematical structure as the spatial pair correlation function of liquid molecules. It can be generated by a sequence of pulses with alternating directions (Fig. 6). Although Eqs. (4) and (5) have revealed the constructive role of the oscillatory effect in noises for a large $\bar{\eta}^{\rm S}_{\rm act}$, it is highly nontrivial how to tune the shape and spacing of the pulses in this noise (see Fig. 6) to obtain a desired $\bar{T}$ in an engine protocol.

B. Oscillatory telegraphic noise

To control the $\bar{T}$ of a noise with a dominant low-frequency component as that in Eq. (4), let us turn to an oscillatory telegraphic noise ξ, which consists of a series of alternating telegraphic steps of equal time length $t_{\rm s}$ (Figs. 7 and 8). Under a constant $\bar{T}$, the magnitude of the nth step is given by $\xi^{(n)} = \bar{\xi}^{(n)} + \xi^{{\rm G}(n)}$. Here, $\bar{\xi}^{(n)} = (-1)^n\bar{\xi}$ is a periodic jump between two values $\pm\bar{\xi}$ with $\bar{\xi} > 0$ and $\xi^{{\rm G}(n)}$ denotes a Gaussian random variable of variance Ξ centered at zero, with $\langle\xi^{{\rm G}(n)}\xi^{{\rm G}(m)}\rangle = \Xi\delta^{\rm K}_{nm}$ and $\delta^{\rm K}_{nm}$ the Kronecker delta. For this piecewise smooth noise, the positions of the Brownian particle at discrete time instants, $i\varepsilon_{\rm t}$ ($i = 0, 1, 2, ...$), can be deduced from Eq. (1) (Appendix C),

$$x_{i+1} = Rx_i + w_i, \quad (6)$$

with $R \equiv 1 - k\varepsilon_{\rm t}/\gamma$ and the increment $w_i \equiv \gamma^{-1}\int_{i\varepsilon_{\rm t}}^{(i+1)\varepsilon_{\rm t}}\xi(t)dt$, where the discretization time span $\varepsilon_{\rm t} < t_{\rm s}$. If we consider an ensemble of systems whose $\xi(t)$ have a coherent phase starting with the same initial value $\bar{\xi}^{(1)} = -\bar{\xi}$ (Fig. 8), then the variance $\langle x^2\rangle$ calculated at all different x_i in Eq. (6) oscillates strongly with i. If we only pick up those x_i at the end of each telegraphic step at time instants $jt_{\rm s}$ ($j = 1, 2, 3, ...$), the variance $\langle x^2\rangle_{\rm e}$ will oscillate much less

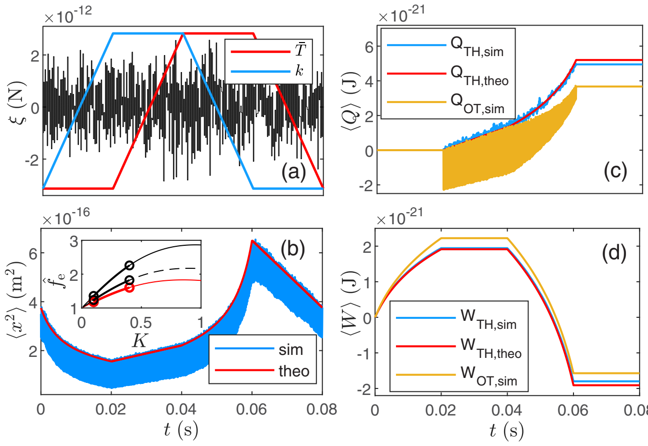

FIG. 3. A colloidal Stirling engine under an oscillatory telegraphic noise. (a) An example of this noise ξ (black) during a linear variation of k (blue) and $\bar{T}$ (red) within a Stirling cycle. (b) The variance $\langle x^2\rangle$ from a simulation (blue) and $\langle x^2\rangle_e$ from the theoretical formula in Eq. (C23) (red). The inset shows $\hat{f}_e(K)$ at $\bar{T}_2 = 200$ K (black solid), $\bar{T}_1 = 300$ K (black dashed), and $\bar{T}_1 = 400$ K (red solid). The three circles on the left- (right-) hand side are located at $K_2 = k_2t_s/(2\gamma)$ ($K_1 = k_1t_s/(2\gamma)$). (c) For the thermal noise, the simulated $\langle Q\rangle$ is $Q_{\mathrm{TH,sim}}$ (blue), which agrees well with the theoretical curve $Q_{\mathrm{TH,theo}}$ (red) given by Eq. (A4). For the oscillatory telegraphic noise ξ, the simulated $\langle Q\rangle$ is $Q_{\mathrm{OT,sim}}$ (yellow), which is shown only in the two heat input processes (iso-stiffness heating and isothermal expansion). (d) The work $\langle W\rangle$ is evaluated as in (c); however, in all four processes because we need its value in the whole cycle to calculate the efficiency. All parameter values used here can be found in Appendix C.

than $\langle x^2\rangle$. The TR of $\langle x^2\rangle_e$ is equal to (Appendix C)

$$\hat{f}_e(K) = \frac{\tanh(K)}{K}[1 + c\tanh(K)], \tag{7}$$

with $K \equiv kt_s/(2\gamma)$ and $c \equiv \bar{\xi}^2/\Xi$, where Ξ is related to $\bar{T}$ by $\Xi \equiv 2\gamma k_B\bar{T}/t_s$. Since $\hat{f}_e(K)$ derived from the less oscillating $\langle x\rangle_e$ has the same rising or falling trend as that of $\hat{f}(K)$ calculated from the oscillating $\langle x^2\rangle$, the former serves as an indicator for possible high-efficiency engines.

As an example, Fig. 3(a) demonstrates an oscillatory telegraphic noise (black) during the variation of $\bar{T}$ and k in an engine protocol. In Fig. 3(b), the oscillating $\langle x^2\rangle$ obtained from the simulations (blue) is "bounded" by the theoretically derived smooth $\langle x^2\rangle_e$ (red), as explained prior to Eq. (C28). In Fig. 3(c), the heat $\langle Q\rangle$ has a strong oscillation (yellow area) because all ξ in the ensemble average have a coherent oscillation phase. The magnitude of the oscillating $\langle Q\rangle$ is comparable to that under the thermal noise at the same temperature when only a single system in the ensemble is considered. At the beginning of the iso-stiffness heating, $\langle Q\rangle$ has a rapid initial drop to a large negative value (left edge of the yellow area) when the particle moves from a large $x > 0$ toward $x = 0$. Notably, in stochastic thermodynamics, the heat along a single trajectory could be negative even in a heat input process. In Fig. 3(d), the oscillation of work $\langle W\rangle$ is much less than that of $\langle Q\rangle$. Averaged over 10^4 cycles, all with the same phase in ξ, we obtain $(\eta^S_{\mathrm{pas}}, \bar{\eta}^S_{\mathrm{act}}) \approx (0.37, 0.43)$ for $(\bar{T}_2, \bar{T}_1) = (200, 400)$ and $\approx(0.20, 0.20)$ for $(\bar{T}_2, \bar{T}_1) = (300, 400)$. The former performs better, as reflected in the steeper $\hat{f}_e(K)$ at $\bar{T}_2 = 200$ K, compared to that at 300 K, in the inset of Fig. 3(b). After randomizing the phase of ξ as in the equal *a priori* probability in the microcanonical ensemble, $\langle Q\rangle$ in Fig. 3(c) becomes a less oscillating curve, as that of $\langle W\rangle$ in Fig. 3(d). See an example of high efficiency under such a randomized phase of the oscillatory telegraphic noise in Fig. 9.

C. Discretization induced high efficiency

If the oscillatory effect is absent, $\bar{\xi} = 0$, and the length of the telegraphic step t_s is as short as the time step ε_t, Eq. (7) will be replaced by $\hat{f}_{e,0}(K) = 1/(1-K)$, where $K \equiv k\varepsilon_t/(2\gamma)$ (Appendix C). This is an even simpler increasing function. However, if $\varepsilon_t \to 0$ is the limit to approach a real system, it yields $K \to 0$ and $\hat{f}_{e,0}(K) \to 1$ for any given finite range $[k_2, k_1]$ in Eq. (3). That is, the efficiency of the active engine will become infinitely close to that of its passive counterpart. Therefore, if $K \sim \varepsilon_t$, an increasing $\hat{f}(K)$ does not promise a high efficiency. This specious high efficiency could be mistakenly concluded from the discrete data ($\varepsilon_t > 0$) recorded in numerous experiments, as illustrated in the feedback system below.

D. Optical feedback trap

The dynamics of a colloidal particle immersed in water of temperature T and confined in an arbitrary real potential can be mimicked by the optical feedback trap (OFT) technique [17,18]. The idea is to shift the center $x_L(t)$ of an optical tweezers of stiffness k_{ot} in the OFT to a proper position to let the particle in $x(t)$ experience an instant optical force, as if it were in the real potential. The dynamics of this particle fulfills the equation (Appendix D),

$$\gamma\dot{x}(t) + k_{ot}[x(t) - x_L(t)] = \xi(t). \tag{8}$$

Here, $\xi(t)$ represents the thermal noise, which is Gaussian and white, with $\langle\xi(t)\rangle = 0$ and $\langle\xi(t)\xi(t')\rangle = 2\gamma k_BT\delta(t-t')$, where $\delta(t-t')$ is the Dirac delta function. If the real potential is harmonic and has stiffness k, ideally $x_L(t)$ should instantly vary with $x(t)$ according to the relation $x_L(t) = -\alpha x(t)$, where $\alpha \equiv -1 + k/k_{ot}$ is the feedback gain. However, due to an inevitable feedback delay, the actual $x_L(t) = -\alpha x(t-t_d)$ contains a small delay time t_d (Appendix D). With this delay, the experimentally measured x are discrete acquisition positions x_i that satisfy the difference equation of Eq. (8),

$$x_{i+1} = x_i - \frac{k_{ot}}{\gamma}(x_i + \alpha x_{i-l})\varepsilon_t + \xi^x_i. \tag{9}$$

Here, l is the number of delay steps, which corresponds to the delay time $t_d = l\varepsilon_t$, and ξ^x_i is the Gaussian fluctuation of the particle position with $\langle\xi^x_i\rangle = 0$ and $\langle\xi^x_i\xi^x_j\rangle = 2k_BT\varepsilon_t\delta^K_{ij}/\gamma$, where δ^K_{ij} is the Kronecker delta as before Eq. (6) [19]. For $l = 1$, the TR of Eq. (9) is (Appendix D)

$$\hat{f}(K) = \frac{-2(K-c_1)}{(K-c_2)(K-c_3)}, \tag{10}$$

where $c_1 \equiv \beta - 1$, $c_2 \equiv \beta + 1$, $c_3 \equiv 2\beta - 2$, $K \equiv k\varepsilon_t/\gamma$, and $\beta \equiv k_{ot}\varepsilon_t/\gamma$. Since ε_t is in practice very small, both K and

$\beta \ll 1$. In this regime, Eq. (10) is clearly an increasing function. It is especially apparent at $\alpha = 0$, for which Eq. (10) becomes $\hat{f}_0(K) = 2/(2-K)$ (Appendix D). However, reducing the data acquisition time, $\varepsilon_t \to 0$, will cause $\hat{f}_0(K) \to 1$ and $\hat{f}(K) \to 1$ in Eq. (10). Therefore, the increasing function in Eq. (10) and $\hat{f}_0(K)$ cannot lead to a high-efficiency engine, as explained in Sec. III C. Nevertheless, it is too early to conclude that delay cannot contribute to high efficiency since Eq. (10) is a consequence of $t_d = \varepsilon_t$, because of which the influence of delay might have disappeared after we take the limit $\varepsilon_t \to 0$.

E. Stochastic delay differential equation

To inspect a pure delay effect on engine efficiency without the influence of data discretization, let us add a delay time t_d to the continuous equation in Eq. (1) to form a stochastic delay differential equation,

$$\gamma \dot{x}(t) = -kx(t - t_d) + \xi(t), \tag{11}$$

where ξ is the same as in Eq. (8). For $t_d > 0$, the system is out of equilibrium and does not satisfy the fluctuation-dissipation theorem [20]. This equation has been applied to study the behavior of an active Brownian particle [21], the emerging structures of multiple active particles [22], and systems in many other contexts [23,24]. With the known variance of the steady state of Eq. (11) [25], one obtains its $\hat{T}$ and TR (Appendix E),

$$\hat{f}(K) = \frac{1 + \sin(K)}{\cos(K)}, \tag{12}$$

when $K \equiv kt_d/\gamma < \pi/2$. The high efficiency indicated by the rising branch of Eq. (12) can be understood from the first-order Taylor expansion of Eq. (11) in a small t_d, corresponding to a small K [20],

$$\gamma \dot{x}(t) = -\frac{kx(t)}{1-K} + \frac{\xi(t)}{1-K}. \tag{13}$$

This equation has a TR (Appendix E)

$$\hat{f}(K) = \frac{1}{1-K}, \tag{14}$$

which is simpler than Eq. (12) and has a clear increasing trend. Compared to Eq. (11) at $t_d = 0$, Eq. (13) has a stronger potential confinement and stochastic force, which compete with each other and lead to a larger $\langle x^2 \rangle$.

The OFT equation (8) is a more complicated stochastic delay differential equation than Eq. (11). It has a nondelay term $k_{ot}x(t)$ in addition to the delay term $k_{ot}x_L(t)$, where the latter plays the same role as $x(t - t_d)$ in Eq. (11). However, the variance $\langle x^2 \rangle$ of an equation like Eq. (8) can also be derived [26,27]. It gives rise to the TR (Appendix D)

$$\hat{f}(K) = \frac{K[(K-c)\sin(g(K)) + g(K)]}{g(K)[c + (K-c)\cos(g(K))]}, \tag{15}$$

when $K > 2c$ and $0 \leqslant g(K) < \cos^{-1}(c/(c-K)) \leqslant \pi$, where $K \equiv kt_d/\gamma$, $c \equiv k_{ot}t_d/\gamma$, and $g(K) \equiv \sqrt{K(K-2c)}$. One can easily obtain a rising regime in Eq. (15) when K is slightly larger than $2c$ (Appendix D).

F. Memory kernels in the position term

The increasing trend of $\hat{f}_{e,0}(K)$ mentioned in Sec. III C comes from $\varepsilon_t \neq 0$, which means that the magnitude of a previous force will last for a period of ε_t in Eq. (6). The increasing trend of $\hat{f}(K)$ in Eq. (12) is caused by $t_d \neq 0$ in Eq. (11), indicative of the influence of an early force on the dynamical system. Both pertain to memory effects, which leads us to examine the dynamics of a colloidal particle governed by the generalized Langevin equation

$$\gamma \dot{x}(t) + k \int_{-\infty}^{t} K_M(t - t')x(t')dt' = \xi(t), \tag{16}$$

where $K_M(t - t')$ is a memory kernel. This model can capture the behavior of a Brownian particle when its potential energy from a confining well can be converted into the energy stored in the compressed background media. More applications of the model can be found in the context of integrodifferential equations of Volterra type [28]. Taking the Fourier transform on both sides of Eq. (16), we obtain the power spectrum of x, $S_x(\omega) = S_\xi(\omega)/\{\omega^2\gamma^2 + k^2\tilde{K}_M(\omega)\tilde{K}_M(-\omega) + i\omega\gamma k[\tilde{K}_M(-\omega) - \tilde{K}_M(\omega)]\}$ and its variance $\langle x^2 \rangle = \int_{-\infty}^{\infty} S_x(\omega)d\omega/(2\pi)$, with $S_\xi(\omega)$ the power spectrum of ξ (Appendix F).

Given $K_M(t) = \exp(-|t|/\tau_p)/\tau_p$ for $t \geqslant 0$ and 0 elsewhere, as well as $\langle \xi(t)\xi(t') \rangle = 2\gamma k_B \bar{T} \delta(t - t')$, the corresponding $\langle x^2 \rangle$ gives an increasing TR (Appendix F),

$$\hat{f}(K) = K + 1, \tag{17}$$

where $K = k\tau_p/\gamma$. It shows that when the noise is uncorrelated, a simple memory effect on the position of the particle is sufficient to enhance the engine performance to surpass its passive counterpart.

G. Dissipation kernels in the velocity term

In contrast to the memory kernel in the position term in Eq. (16), an even more well-known generalized Langevin equation has a dissipation kernel $K_D(t - t')$ in the velocity term [29,30],

$$\gamma \int_{-\infty}^{t} K_D(t - t')\dot{x}(t')dt' + kx = \xi(t). \tag{18}$$

It has been used to study several active [31] and self-propelled particles [32] for persistent random walks, as well as Brownian colloidal heat engines in viscoelastic baths [33,34]. Such dissipation kernel has also been suspected to be the reason for the high efficiency of the bacterial-driven heat engine [8]. In analogy to the calculations below Eq. (16), the variance $\langle x^2 \rangle$ of Eq. (18) can be obtained from its power spectrum $S_x(\omega) = S_\xi(\omega)/\{k^2 + \omega^2\gamma^2\tilde{K}_D(\omega)\tilde{K}_D(-\omega) + i\omega\gamma k[\tilde{K}_D(\omega) - \tilde{K}_D(-\omega)]\}$ (Appendix G).

If $K_D(t) = \exp(-|t|/\tau_v)/\tau_v$ for $t \geqslant 0$ and 0 elsewhere, as well as $\langle \xi(t)\xi(t') \rangle = 2\gamma k_B \bar{T} \exp(-|t - t'|/\tau)/(2\tau)$, it yields (case 1 in Appendix G).

$$\hat{f}(K) = \frac{c(K + 1 - c)}{(1 - c)(K + 1)(K + c)}, \tag{19}$$

where $c \equiv \tau_v/(\tau_v + \tau)$ and $K \equiv k\tau_v/\gamma$. This $\hat{f}(K)$ is an increasing function when $K \in [0, c - 1 + \sqrt{c(2c - 1)}]$

if $c > c_0 \equiv (-1+\sqrt{5})/2$, corresponding to $\tau_v/\tau > (-1+\sqrt{5})/(3-\sqrt{5})$. In the limit $\tau_v \to 0$, Eq. (19) returns to the $f(k)$ calculated by $\langle x^2\rangle = k_B T/[k(1+k\tau/\gamma)]$ in Eq. (15) of Ref. [8]. With the same $K_D(t)$ but different correlation $\langle \xi(t)\xi(t')\rangle = 2a\gamma k_B \bar{T}\,\mathrm{sinc}(a(t-t'))$, where $a > 0$ and $\mathrm{sinc}(x) = \sin(\pi x)/(\pi x)$, the TR becomes (case 2 in Appendix G)

$$\hat{f}(K) = \frac{2(1-K)}{\pi}\left[cK + (1-K^2)\tan^{-1}\left(\frac{c}{K}\right)\right], \tag{20}$$

where $c = \pi a \tau_v$ and $K \equiv k\tau_v/(k\tau_v + \gamma)$ increases with k and lies within (0,1). This $\hat{f}(K)$ is also an increasing function, at least at small K when $\tan^{-1}(c/K) \approx \pi/2$, for which the rising trend appears when $c > c_0 \equiv (\pi + \sqrt{\pi^2+16})/4$.

Both Eqs. (19) and (20) can become an increasing function in some range of K when c exceeds a threshold c_0, or equivalently when τ_v is larger than some critical value. That is, the correlation in the dissipation kernel K_D must be sufficiently strong, in comparison with that of the injection kernel $\langle \xi(t)\xi(t')\rangle$ to find high-efficiency engines. Interestingly, both conditions $c > c_0$ for Eqs. (19) and (20) do not depend on viscosity γ. Moreover, regardless of whether $\langle \xi(t)\xi(t')\rangle$ has a long tail in Eq. (19) or a sharp cutoff frequency in Eq. (20), a rising regime can be found in these TRs. Therefore, such a regime is highly expected for other $\langle \xi(t)\xi(t')\rangle$ with more general intermediate tails.

IV. ORIGINS OF HIGH EFFICIENCY

The high-efficiency engines found above can be summarized into two categories: (i) microscopic restoring effects of noise and (ii) strong memory or dissipation kernels. Category (i) includes the oscillatory effect, which leads to Eqs. (5) and (7), while (ii) contains the retarding effect, which causes Eqs. (12) and (15), as well as various kernels that induce Eqs. (17), (19), and (20). Mathematically, why these two categories have high efficiencies is that an additional parameter is introduced in $\langle x^2\rangle$ to let $f(k)$ behave as k^α elucidated below Eq. (3). For example, in category (i), introducing $\bar{\xi}$ into the telegraphic noise before Eq. (6) creates the term $c\tanh(K)$ in Eq. (7) to avoid the monotonic decline of $\tanh(K)/K$. In category (ii), introducing t_d into Eq. (11) generates the terms $\sin(K)$ and $\cos(K)$ in Eq. (12) to turn $\hat{f}(K)$ into an increasing function. The increasing trends of most of these $\hat{f}(K)$ are depicted in Fig. 4.

The examples studied above provide a lot of useful information to understand the efficiency of stochastic engines. First, according to the expressions of $\langle Q\rangle$ and $\langle W\rangle$ before Eq. (2), it is tempting to regard a small $k_f\langle x^2\rangle_f - k_i\langle x^2\rangle_i$ and a large $\int_i^f \langle x^2\rangle dk$ for general stiffness and temperature as the condition for a dynamical system of x to generate good trajectories (Fig. 1) because they would give a large ratio $\langle W\rangle/\langle Q\rangle$ for a cycle. Unfortunately, high efficiency is usually not due to a large ratio $\langle W\rangle/\langle Q\rangle$ in all individual processes, but a compromise between large and small ratios in different processes, as seen in Figs. 3(c) and 3(d). Certainly, one can question whether some perfect dynamical system could generate large $\langle W\rangle/\langle Q\rangle$ for all processes. Second, based on the mismatch between energy injection and dissipation, $\bar{T} < \hat{T}$ is hypothesized to be a condition for finding high efficiencies [8]. It looks trivial because if an engine cycle in a high-temperature range $[\hat{T}_1, \hat{T}_2]$ in the $(k, \hat{T})$ space is represented in a low-temperature range $[\bar{T}_1, \bar{T}_2]$ in the $(k, \bar{T})$ space, the low heat input of the latter would intuitively lead to higher efficiency. However, such a comparison is not fair because it would not be between two Stirling engines, as explained in Fig. 2. Nevertheless, as seen in Fig. 4, the rising regimes of all studied $\hat{f}(K)$ indeed have $\bar{T} < \hat{T}$, or equivalently $\hat{f}(K) > 1$. Third, the restoring effect of noise could intuitively suppress particle diffusion and its associated $\bar{T}$, but would not greatly change the long-term value of the confined variance $\langle x^2\rangle$ and its related $\hat{T}$. Thus, this effect seems to be beneficial for creating $\bar{T} < \hat{T}$. However, this impression is challenged by the oscillatory factor $\bar{\xi}$ before Eq. (6), which can influence $\hat{T}_e$ in Eq. (C24), but not $\bar{T}_e = \bar{T}$ in Eqs. (C17) and (C28). Lastly, note that taking different interpretations of temperature neither produces new trajectories x in Fig. 1 nor modifies the relationship between these trajectories and the functionals of heat, work, and efficiency. It merely affects the temperature values assigned to the trajectories evolved from a thermodynamic process, as seen in Eqs. (A2) and (A3). Ultimately, diverse interpretations of temperature lead to a divergence in the efficiency value of an engine [see more examples below the paragraph of Eq. (A13) in Appendix A].

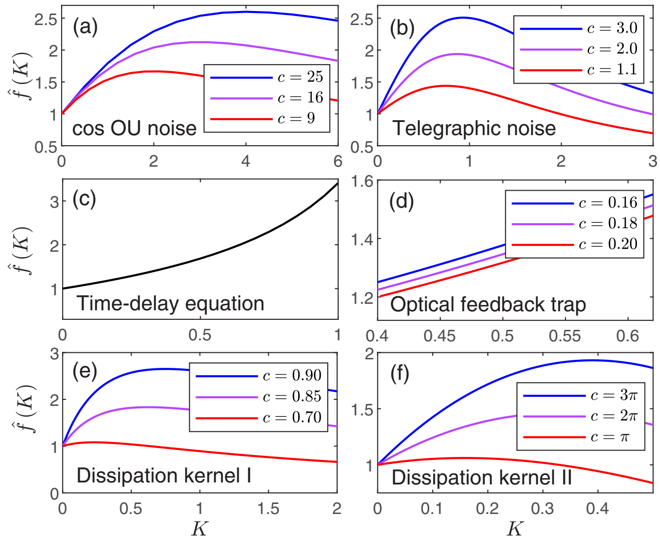

FIG. 4. (a), (b), (c), (d), (e), and (f) are the TRs $\hat{f}(K)$ of six stochastic dynamical systems in Eqs. (5), (7), (12), (15), (19), and (20) within experimentally accessible parameter ranges collected in Appendix B, C, D, E, and G.

V. DISCUSSION

In addition to indicating high efficiencies, $f(k)$ offers further insight into how active noise impacts more general thermodynamic problems in heat engines. Firstly, the difference between the active Carnot efficiency $\bar{\eta}^{\mathrm{C}}_{\mathrm{act}}$ and the active Stirling efficiency $\bar{\eta}^{\mathrm{S}}_{\mathrm{act}}$ can be expressed as

$$\bar{\eta}^{\mathrm{C}}_{\mathrm{act}} - \bar{\eta}^{\mathrm{S}}_{\mathrm{act}} = \left(\eta^{\mathrm{C}}_{\mathrm{pas}}\right)^2 H(k_1, k_2), \tag{21}$$

which is linearly proportional to the square of the passive Carnot efficiency $\eta_{\mathrm{pas}}^{\mathrm{C}}$ [see Eq. (A18)]. Here, the function $H(k_1, k_2)$ is solely determined by $f(k)$ and the maximum k_1 (minimum k_2) of stiffness. As $H(k_1, k_2) > 0$, the discrepancy in Eq. (21) is always positive. Secondly, if the engine operation time is sufficiently large but not infinite, the power of the engine can be expressed in terms of $f(k)$ as well [see Eq. (A22)],

$$P \approx -\frac{1}{\tau}\frac{k_{\mathrm{B}}(\bar{T}_2 - \bar{T}_1)}{2}\int_{k_2}^{k_1}\frac{f(k)}{k}dk. \tag{22}$$

Thirdly, the efficiency at maximum power is also influenced by active noises, which can be illustrated by the low dissipation Carnot engine [see Eq. (A23)]. In this engine, the heat exchange during the isothermal process at low (high) temperature T_{c} (T_{h}) is $Q_{\mathrm{c}} = Q_{\mathrm{c}}^{\infty} - T_{\mathrm{c}}\Sigma_{\mathrm{c}}/\tau_{\mathrm{c}}$ ($Q_{\mathrm{h}} = Q_{\mathrm{h}}^{\infty} - T_{\mathrm{h}}\Sigma_{\mathrm{h}}/\tau_{\mathrm{h}}$), where τ_{c} (τ_{h}) is its corresponding operation time and Q_{c}^{∞} (Q_{h}^{∞}) is the heat for $\tau_{\mathrm{c}} \to \infty$ ($\tau_{\mathrm{h}} \to \infty$). The efficiency at maximum power of the passive Carnot engine is then [see Eq. (A26)],

$$\eta^* = \eta_{\mathrm{C}} H'(T_{\mathrm{c}}, \Sigma_{\mathrm{c}}, T_{\mathrm{h}}, \Sigma_{\mathrm{h}}), \tag{23}$$

which is proportional to the efficiency of the conventional Carnot efficiency η_{C}. Here, H' is a function of T_{c}, T_{h}, Σ_{c}, and Σ_{h}. Although the dependence of heat on $f(k)$ can alter the efficiency in Eq. (2), it interestingly does not influence η^*. This implies that an active noise affects the Curzon-Ahlborn efficiency solely through the variations in Σ_{c} and Σ_{h} with $f(k)$. Fourthly, the probability of finding a stochastic efficiency $\tilde{\eta}$ among all long trajectories in Fig. 1 is given by [see Eq. (A14)]

$$P_i(\tilde{\eta}) \sim e^{-tJ(\tilde{\eta})}, \tag{24}$$

where t is the time span of the trajectory and $J(\tilde{\eta})$ is the large deviation function [35]. If $P_i(\tilde{\eta})$ can still be expressed in terms of $f(k)$, the latter should be embedded in $J(\tilde{\eta})$.

The current colloidal heat engine (HE) shares some similarities with the chemical reaction network (CN) of D-ribose isomerization under temperature switching in Ref. [36,37]. The latter is a kinetic model of three isomers, with their population ratio being tuned by temperature. The HE (CN) undergo deterministic (stochastic) cyclic transformations between a cold and a hot temperature, denoted as T_{c} and T_{h}, respectively. The state of HE is denoted by the location of the colloidal particle on the position coordinate x_{p}, whereas the state of CN characterized by the conformation of furanose is represented on the reaction coordinate x_{r}. The heat exchange along a state trajectory of HE on x_{p} can generate mechanical work, while that of CN on the abstract axis x_{r} is irrelevant to work, but to entropy production. The state of HE typically follows a quasisteady cycle, whereas that of CN alternately repeats two relaxation processes toward different equilibrium states. Both the HE and CN fall within the category of nonequilibrium systems. However, nonequilibrium in the HE arises from the violation of the fluctuation-dissipation theorem, characterized by a nonzero energy flux between injection and dissipation kernels. Instead, nonequilibrium in the CN results from the violation of detailed balance, indicated by a nonzero population flux between chemical states. When the parameter k_{D} for mass transport in the CN is small, the switches between T_{c} and T_{h} are slow and resemble those in traditional Carnot or Stirling heat engines. Conversely, when k_{D} is sufficiently large, the switches occur so rapidly that the populations of different isomers in the CN reach a nonequilibrium steady distribution, which is reminiscent of the Brownian gyrator [38], a type of heat engine without temperature switching.

An even broader class of nonequilibrium system is that of chemically active systems (CAS), which encompass active matter [39] and various life-related chemical processes [40]. In these systems, a constant flow of energy or matter propels chemical reactions, generates mechanical forces, and/or initiates molecular motion, which prevents the system from thermalizing toward an equilibrium state. The active noises of HE will be a member of CAS when they originate from active matter comprised of self-propelled agents, as opposed to passive matter like water molecules in passive engines. While these agents can be bird flocks, fish schools, and Janus particles within the category of active matter, they usually refer to bacterium swarms in the context of the microscopic HE [4]. In addition, the CN is also a type of CAS because the continuous injection and removal of energy during temperature switching result in non-Boltzmannian population distributions among distinct conformational states. Although both HE and CN are nonequilibrium problems, how they are rendered nonequilibrium by temperature is different. Another recent intriguing study of CAS involves the measurement of local temperature and heat transport within living cells [41–43], which is possible as it is based on the local equilibrium condition. However, there has been concern regarding the limits of length and time scales for the validity of that condition [40], which is estimated to be approximately 10^2 nm and 10^{-6} s inside a living cell [44]. Above these critical values, passive thermal fluctuations will be dominated by active ones, the latter originating from local stochastic events of chemical reactions and having different meanings in temperature. In this case, a tracer particle inside a cell will yield a value for $\hat{T}$ from its position variance and another value for $\bar{T}$ from diffusion. The discrepancy between the values of $\hat{T}$ and $\bar{T}$ is similar to that observed in the colloidal heat engine analyzed in this study.

VI. CONCLUSIONS

In the field of statistical physics, there has been a recent trend towards expanding its scope to incorporate active matter systems. However, the nonequilibrium nature of these systems gives rise to nonuniqueness in defining thermodynamic concepts and relations, which are difficult to extrapolate from their equilibrium counterparts [1]. In this study, active matter serves as a heat bath connected to a heat engine governed by a series of stochastic dynamical systems. Under the two most frequently considered interpretations of temperature $\hat{T}$ and $\bar{T}$, our analysis reveals whether, when, and why the engine efficiency of an active bath can unexpectedly surpass that of its passive equilibrium counterpart. Physically, such high performance can be attributed to either a restoring effect in noise or a strong dissipation kernel. Mathematically, it arises from the enlarged dimensionality in the parameter space of the heat engine. Interestingly, the heat engine efficiency once again plays a role in the history of understanding temperatures, apart from its early contribution in finding the absolute temperature [45,46]. In general active heat baths, other temperature measures beyond $\hat{T}$ and $\bar{T}$ are expected, such as the active particle velocity-dependent temperature [47]. The interpretation, understanding, or potential unification of the diverse consequences stemming from these energy scales will be a challenging task for future statistical physics. The current study sets an example for research in this direction.

ACKNOWLEDGMENTS

The authors acknowledge the support of the Ministry of Science and Technology, Taiwan, with Grants No. 108-2112-M-008-009- (Y.J.) and No. 109-2112-M-009-008- (C.-H.C.), and the support of NCTS of Taiwan.

APPENDIX A: EFFICIENCIES OF STIRLING ENGINES IN AN ACTIVE BATH

The temperature of a stochastically moving Brownian particle has two commonly adopted definitions. One is $\bar{T}$ defined by the mean-square displacement $\langle x^2(t)\rangle = 2Dt = 2(k_B\bar{T}/\gamma)t$ averaged over an ensemble of particle trajectories $x(t)$, with D the diffusion constant of the particle, k_B the Boltzmann constant, and γ the friction coefficient, which are related by the Einstein relation $D = k_B\bar{T}/\gamma$. The other is $\hat{T}$ defied by the equipartition theorem, $\langle x^2\rangle = k_B\hat{T}/k$, via the variance $\langle x^2\rangle$ of the stationary distribution of the particle when it is confined in a harmonic potential of stiffness k.

Since $\bar{T}$ is defined by the diffusivity of a tracer particle surrounded by background particles, such as bacteria in an active heat bath or water molecules in a passive heat bath, this temperature increases with the activity of the background particles, like the enhanced self-propulsion force of bacteria or the intensified agitation of water molecules. On the other hand, $\hat{T}$ is defined by the "effective pressure" $\langle x^2\rangle$ under a confining potential of stiffness k. This temperature not only increases with the activity of the background particles, as $\bar{T}$, but also with the "effective volume" $1/k$ [48]. Although $\hat{T}$ seems to be more commonly used in the literature, $\bar{T}$ is arguably a more natural definition of temperature because it does not depend on a system parameter like k [8].

Although $\bar{T}$ and $\hat{T}$ are identical for the thermal noise (a passive noise), they are generally different from each other for an active noise. In the latter case, these two temperatures can be related by [8]

$$\hat{T} = f(k)\bar{T}. \tag{A1}$$

In stochastic energetics [15], if a Brownian particle is subjected to a force $F(x)$ and moves along a stochastic trajectory from an initial state i to a final state f, then the stochastic heat received by the particle is $Q \equiv -\int_i^f F(x)dx$. For $F(x) = -kx$, the heat will be $Q = \int_i^f kx\, dx = [\frac{kx^2}{2}]_i^f - \int_i^f \frac{x^2}{2}dk$. In the quasistatic regime, its ensemble average can be expressed in terms of $\hat{T}$ or $\bar{T}$ and represented in the $(k, \hat{T})$ or $(k, \bar{T})$ space,

$$\begin{aligned}\langle Q\rangle &= \frac{1}{2}\big(k_f\langle x^2\rangle_f - k_i\langle x^2\rangle_i\big) - \frac{1}{2}\int_i^f k\langle x^2\rangle\frac{dk}{k}\\ &= \underbrace{\frac{k_B}{2}(\hat{T}_f - \hat{T}_i) - \frac{k_B}{2}\int_i^f \hat{T}\frac{dk}{k}}_{(k,\hat{T})\text{ space}}\\ &= \underbrace{\frac{k_B}{2}[f(k_f)\bar{T}_f - f(k_i)\bar{T}_i] - \frac{k_B}{2}\int_i^f f(k)\bar{T}\frac{dk}{k}}_{(k,\bar{T})\text{ space}}.\end{aligned} \tag{A2}$$

Here, k_i (k_f) is the stiffness in the initial (final) state. Furthermore, the stochastic work done on the particle is $W = \int_i^f \frac{\partial U}{\partial k}\frac{dk}{dt}dt = \int_i^f \frac{1}{2}x^2dk$ [15], when $U = -\int F(x)dx = \frac{kx^2}{2}$. In the quasistatic regime, its ensemble average can be expressed in terms of $\hat{T}$ or $\bar{T}$ and represented in the $(k, \hat{T})$ or $(k, \bar{T})$ space,

$$\begin{aligned}\langle W\rangle &= \frac{1}{2}\int_i^f k\langle x^2\rangle\frac{dk}{k}\\ &= \underbrace{\frac{k_B}{2}\int_i^f \hat{T}\frac{dk}{k}}_{(k,\hat{T})\text{ space}} = \underbrace{\frac{k_B}{2}\int_i^f f(k)\bar{T}\frac{dk}{k}}_{(k,\bar{T})\text{ space}}.\end{aligned} \tag{A3}$$

In Eqs. (A2) and (A3), $\langle x^2\rangle$ and $\hat{T}$ generally vary with k and the noise strength. If the noise strength is properly controlled to let $k\langle x^2\rangle$ remain invariant during the change from state i to f, the process is iso-$\hat{T}$, which means isothermal at constant $\hat{T} = k\langle x^2\rangle/k_B$. Likewise, if $\bar{T} = \hat{T}/f(k)$ is kept unchanged, the process is iso-$\bar{T}$.

According to Eqs. (A2) and (A3), an iso-$\hat{T}$ and an iso-k process in the $(k, \hat{T})$ space have the explicit expressions

$$\begin{aligned}&\begin{cases}\langle Q\rangle = -\frac{k_B\hat{T}}{2}\ln\left(\frac{k_f}{k_i}\right)\\ \langle W\rangle = \frac{k_B\hat{T}}{2}\ln\left(\frac{k_f}{k_i}\right)\end{cases} \text{iso-}\hat{T}\text{ in space }(k,\hat{T}),\\ &\begin{cases}\langle Q\rangle = \frac{k_B}{2}(\hat{T}_f - \hat{T}_i)\\ \langle W\rangle = 0\end{cases} \quad\text{iso-}k\text{ in space }(k,\hat{T}).\end{aligned} \tag{A4}$$

In analogy, an iso-$\bar{T}$ and an iso-k process in the $(k, \bar{T})$ space will give

$$\begin{aligned}&\begin{cases}\langle Q\rangle = \frac{k_B\bar{T}}{2}[f(k_f) - f(k_i) - G(k_f) + G(k_i)],\\ \langle W\rangle = \frac{k_B\bar{T}}{2}[G(k)]_i^f, \quad \text{iso-}\bar{T}\text{ in space }(k,\bar{T}),\end{cases}\\ &\begin{cases}\langle Q\rangle = \frac{k_B}{2}(\bar{T}_f - \bar{T}_i)f(k),\\ \langle W\rangle = 0\end{cases} \quad\text{iso-}k\text{ in space }(k,\bar{T}),\end{aligned} \tag{A5}$$

where $G(k) = \int dk f(k)/k$. Notice that the iso-$\bar{T}$ process in (A5) comes from (A2) and (A3). It cannot be derived by simply inserting $\hat{T} = f(k)\bar{T}$ into the iso-$\hat{T}$ process in (A4) because this would only reexpress an iso-$\hat{T}$ process in the $(k, \hat{T})$ space by a generally "non"-iso-$\bar{T}$ process in the $(k, \bar{T})$ space.

A Stirling engine consists of four processes (Fig. 5): In the iso-T compression (A $\to$ B), k rises from k_2 to k_1 at T_2. In the iso-k heating (B $\to$ C), T rises from T_2 to T_1 at k_1. In the

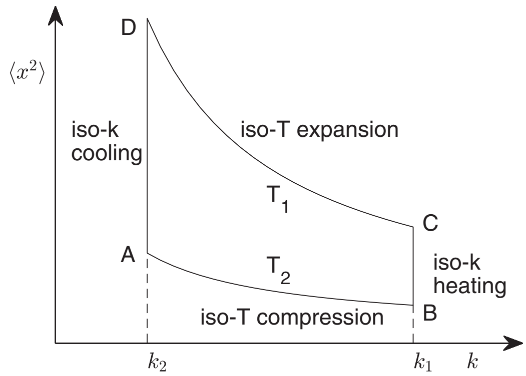

FIG. 5. The four processes in the protocol of a Stirling engine.

iso-T expansion (C $\to$ D), k falls from k_1 to k_2 at T_1. In the iso-k cooling (D $\to$ A), T falls from T_1 to T_2 at k_2. The cycle formed by these four processes is represented by a loop in the $(k, \langle x^2\rangle)$ space in Fig. 5. If iso-T there means iso-$\hat{T}$ (iso-$\bar{T}$), the loop should have a rectangular shape in the $(k, \hat{T})$ space ($(k, \bar{T})$ space), as the three rectangles depicted in Fig. 2. In a passive bath, $\hat{T} = \bar{T}$ and iso-$\hat{T}$ and iso-$\bar{T}$ processes become identical, while in an active bath, generically $\hat{T} \neq \bar{T}$.

For a Stirling engine in the $(k, \hat{T})$ space, its four processes have the expressions of heat and work from Eq. (A4),

$$\begin{aligned}
&\langle Q_{\rm AB}\rangle = -\frac{k_{\rm B}\hat{T}_2}{2}\ln a < 0, \quad \langle Q_{\rm BC}\rangle = \frac{k_{\rm B}}{2}(\hat{T}_1 - \hat{T}_2) > 0,\\
&\langle Q_{\rm CD}\rangle = \frac{k_{\rm B}\hat{T}_1}{2}\ln a > 0, \quad \langle Q_{\rm DA}\rangle = -\frac{k_{\rm B}}{2}(\hat{T}_1 - \hat{T}_2) < 0,\\
&\langle W_{\rm AB}\rangle = \frac{k_{\rm B}\hat{T}_2}{2}\ln a > 0, \quad \langle W_{\rm CD}\rangle = -\frac{k_{\rm B}\hat{T}_1}{2}\ln a < 0,\\
&\langle W_{\rm BC}\rangle = \langle W_{\rm DA}\rangle = 0, \qquad \text{(A6)}
\end{aligned}$$

with $a \equiv k_1/k_2 > 1$. For a Stirling engine in the $(k, \bar{T})$ space, the heat and work in Eq. (A5) indicate

$$\begin{aligned}
&\langle Q_{\rm AB}\rangle = \frac{k_{\rm B}\bar{T}_2}{2}[f(k_1) - f(k_2) - G(k_1) + G(k_2)],\\
&\langle Q_{\rm BC}\rangle = \frac{k_{\rm B}}{2}(\bar{T}_1 - \bar{T}_2)f(k_1) > 0,\\
&\langle Q_{\rm CD}\rangle = -\frac{k_{\rm B}\bar{T}_1}{2}[f(k_1) - f(k_2) - G(k_1) + G(k_2)],\\
&\langle Q_{\rm DA}\rangle = -\frac{k_{\rm B}}{2}(\bar{T}_1 - \bar{T}_2)f(k_2) < 0,\\
&\langle W_{\rm AB}\rangle = \frac{k_{\rm B}\bar{T}_2}{2}[G(k_1) - G(k_2)] > 0,\\
&\langle W_{\rm CD}\rangle = -\frac{k_{\rm B}\bar{T}_1}{2}[G(k_1) - G(k_2)] < 0,\\
&\langle W_{\rm BC}\rangle = \langle W_{\rm DA}\rangle = 0. \qquad \text{(A7)}
\end{aligned}$$

Among the eight terms in Eq. (A7), only the signs of $\langle Q_{\rm AB}\rangle$ and $\langle Q_{\rm CD}\rangle$ are less trivial. However, due to $\langle Q_{\rm AB}\rangle = -(\bar{T}_2/\bar{T}_1)\langle Q_{\rm CD}\rangle$, the conditions for $\langle Q_{\rm AB}\rangle < 0$ and $\langle Q_{\rm CD}\rangle > 0$ are equivalent. This generic case occurs if and only if

$$G(k_1) - G(k_2) = \int_{k_2}^{k_1} \frac{f(k)}{k}dk > f(k_1) - f(k_2). \qquad \text{(A8)}$$

When $f(k) = \hat{T}/\bar{T} = k\langle x^2\rangle/(k_{\rm B}\bar{T})$ decreases with k, the inequality in Eq. (A8) is trivial because $\int_{k_2}^{k_1}\frac{f(k)}{k}dk > 0$ and $f(k_1) - f(k_2) < 0$. For example, $f(k) = (1 + k\tau/\gamma)^{-1}$ in Sec. III A or Eq. (B10) has

$$\begin{aligned}
&\int_{k_2}^{k_1} \frac{f(k)}{k}dk - f(k_1) + f(k_2)\\
&= \int_{k_2}^{k_1} \frac{1}{1 + k\tau/\gamma}\frac{1}{k}dk - \frac{1}{1 + k_1\tau/\gamma} + \frac{1}{1 + k_2\tau/\gamma}\\
&= \ln\left(\frac{\tau + \gamma/k_2}{\tau + \gamma/k_1}\right) + \frac{\gamma\tau(k_1 - k_2)}{(\gamma + k_2\tau)(\gamma + k_1\tau)} > 0 \qquad \text{(A9)}
\end{aligned}$$

because either of the two terms in the last line is positive.

In opposite, if $f(k)$ increases with k, it needs to rise sufficiently fast to see a violation of Eq. (A8). For instance, for $f(k) = ck^\alpha$ within $[k_2, k_1]$ with $\alpha,\ c > 0$,

$$\begin{aligned}
&\int_{k_2}^{k_1} \frac{f(k)}{k}dk - [f(k_1) - f(k_2)]\\
&= \int_{k_2}^{k_1} ck^{\alpha-1}dk - c\big(k_1^\alpha - k_2^\alpha\big)\\
&= c\left(\frac{1}{\alpha} - 1\right)\big(k_1^\alpha - k_2^\alpha\big). \qquad \text{(A10)}
\end{aligned}$$

The condition for Eq. (A10) <0 is $\alpha > 1$. Hence, at least $f(k)$ must rise to the extent to have $\alpha > 1$ within $[k_2, k_1]$ to see $\langle Q_{\rm AB}\rangle > 0$ and $\langle Q_{\rm CD}\rangle < 0$. However, in this case, the fluctuation of the particle $\langle x^2\rangle$ in $f(k) = k\langle x^2\rangle/(k_{\rm B}\bar{T})$ will increase with k, which is counterintuitive and never seen in all systems tested below, including Eqs. (B8), (C25), (D15), (E4), (G13), and (G23) (see A8 in Supplemental Material [49]). All these systems have $\int_{K_2}^{K_1}\frac{\hat{f}(K)}{K}dK > \hat{f}(K_1) - \hat{f}(K_2)$, where $\hat{f}(K)$ is the function $f(k)$ with k replaced by $K = \text{const} \times k$ or $K \equiv (1 + \text{const} \times k^{-1})^{-1}$. Such an inequality of $\hat{f}(K)$ implies the inequality in Eq. (A8), which leads to $\langle Q_{\rm AB}\rangle < 0$ and $\langle Q_{\rm CD}\rangle > 0$.

Under the generic condition given in Eq. (A8), we have $\langle Q_{\rm AB}\rangle < 0$, $\langle Q_{\rm BC}\rangle > 0$, $\langle Q_{\rm CD}\rangle > 0$, and $\langle Q_{\rm DA}\rangle < 0$, for which the efficiency of an active Stirling engine in the $(k, \bar{T})$ space can be deduced from Eq. (A7),

$$\begin{aligned}
\bar{\eta}_{\rm act}^{\rm S} &= -\frac{\langle W\rangle}{\langle Q_{\rm input}\rangle}\\
&= -\frac{\langle W_{\rm AB}\rangle + \langle W_{\rm BC}\rangle + \langle W_{\rm CD}\rangle + \langle W_{\rm DA}\rangle}{\langle Q_{\rm BC}\rangle + \langle Q_{\rm CD}\rangle}\\
&= \frac{(\bar{T}_1 - \bar{T}_2)\int_{k_2}^{k_1}\frac{f(k)}{k}dk}{\bar{T}_1 f(k_2) - \bar{T}_2 f(k_1) + \bar{T}_1\int_{k_2}^{k_1}\frac{f(k)}{k}dk}\\
&= (\bar{T}_1 - \bar{T}_2)\left[\frac{\bar{T}_1 f(k_2) - \bar{T}_2 f(k_1)}{\int_{k_2}^{k_1}\frac{f(k)}{k}dk} + \bar{T}_1\right]^{-1}. \qquad \text{(A11)}
\end{aligned}$$

The third equality is equal to Eq. (2), and as known in Ref. [8]. On the contrary, if $\langle Q_{\rm AB}\rangle > 0$, $\langle Q_{\rm BC}\rangle > 0$, $\langle Q_{\rm CD}\rangle < 0$, and $\langle Q_{\rm DA}\rangle < 0$ should happen, $\langle Q_{\rm input}\rangle = \langle Q_{\rm BC}\rangle + \langle Q_{\rm CD}\rangle$ in Eq. (A11) will be replaced by $\langle Q_{\rm BC}\rangle + \langle Q_{\rm AB}\rangle = \langle Q_{\rm BC}\rangle - (\bar{T}_2/\bar{T}_1)\langle Q_{\rm CD}\rangle$.

For $f(k) = 1$, Eq. (A11) reduces to the efficiency of a passive Stirling engine

$$\eta^{\rm S}_{\rm pas} = (\bar{T}_1 - \bar{T}_2)\left[\frac{\bar{T}_1 - \bar{T}_2}{\int_{k_2}^{k_1}\frac{1}{k}dk} + \bar{T}_1\right]^{-1}. \tag{A12}$$

Therefore, the condition for $\bar{\eta}^{\rm S}_{\rm act} > \eta^{\rm S}_{\rm pas}$ is

$$\begin{aligned}&\frac{\bar{T}_1 f(k_2) - \bar{T}_2 f(k_1)}{\int_{k_2}^{k_1}\frac{f(k)}{k}dk} < \frac{\bar{T}_1 - \bar{T}_2}{\int_{k_2}^{k_1}\frac{1}{k}dk}\\ &\Leftrightarrow \int_{k_2}^{k_1}\{\bar{T}_1 f(k_2) - \bar{T}_2 f(k_1) - [\bar{T}_1 - \bar{T}_2]f(k)\}\frac{dk}{k} < 0\\ &\Leftrightarrow \int_{k_2}^{k_1}\left\{\frac{\bar{T}_2}{\bar{T}_1}[f(k_1) - f(k)] + [f(k) - f(k_2)]\right\}\frac{dk}{k} > 0.\end{aligned} \tag{A13}$$

For $T_1 \gg T_2$, it reduces to $\int_{k_2}^{k_1}[f(k) - f(k_2)]\frac{dk}{k} > 0$ [8]. A sufficient but not necessary simple and useful condition for Eq. (A13) is a "monotonically increasing $f(k)$" in the range $[k_2, k_1]$ where the engine operates.

To illustrate the reason why diverse interpretations of temperature lead to a divergence in the efficiency value, claimed prior to the Conclusions, here let us consider some examples. Suppose a thermodynamic process is generated by varying the stiffness k in a desired manner. Repeating the same process multiple times leads to an ensemble of data x, which yields a unique value of $\langle x^2\rangle$ and $\langle W\rangle = \frac{1}{2}\int_i^f k\langle x^2\rangle\frac{dk}{k}$ in the first expression of Eq. (A3), say $\langle W\rangle = 3$. If we take the second expression there, $\langle W\rangle = \frac{k_{\rm B}}{2}\int_i^f \hat{T}\frac{dk}{k}$, suppose it leads to $\hat{T} = 1$. Then we will obtain $\bar{T} = \frac{1}{2}$, if we take the third expression, $\langle W\rangle = \frac{k_{\rm B}}{2}\int_i^f f(k)\bar{T}\frac{dk}{k}$, in case $f(k) \equiv 2$. That is, $\langle W\rangle = 3$ is the work along a process at $\hat{T} = 1$ in the $(k, \hat{T})$ space or the work along a process at $\bar{T} = \frac{1}{2}$ in the $(k, \bar{T})$ space. To obtain a process at $\bar{T} = 1$ in the $(k, \bar{T})$ space, we need to vary k in a different way. It would generally change $\langle x^2\rangle$ and subsequently its work to become a different value, say $\langle W\rangle = 5$. That is, even though the process path at $\hat{T} = 1$ in the $(k, \hat{T})$ space and at $\bar{T} = 1$ in the $(k, \bar{T})$ space are the same, they have different works $\langle W\rangle = 3$ and 5, respectively. Likewise, even though two cycles in the $(k, \hat{T})$ and $(k, \bar{T})$ spaces are the same, their $\langle W\rangle$, $\langle Q\rangle$, and subsequently their efficiencies are generally different. As an example of "the same cycle", consider a Stirling cycle in the $(k, \hat{T})$ space, which is a rectangle specified by four corner points $(k_2, \hat{T}_2)$, $(k_1, \hat{T}_2)$, $(k_1, \hat{T}_1)$, and $(k_2, \hat{T}_1)$, whose efficiency is a function $\hat{\eta}(k_1, k_2, \hat{T}_1, \hat{T}_2)$ of $k_1, k_2, \hat{T}_1$, and $\hat{T}_2$. Furthermore, a Stirling cycle in the $(k, \bar{T})$ space is also a rectangle specified by another four corner points $(k_2', \bar{T}_2)$, $(k_1', \bar{T}_2)$, $(k_1', \bar{T}_1)$, and $(k_2', \bar{T}_1)$, whose efficiency is generally a different function $\bar{\eta}(k_1', k_2', \hat{T}_1, \hat{T}_2)$. To compare their performance, the difference between these two formula, $\hat{\eta}(k_1, k_2, \hat{T}_1, \hat{T}_2) - \bar{\eta}(k_1', k_2', \bar{T}_1, \bar{T}_2)$, is calculated under the condition $k_1 = k_1'$, $k_2 = k_2'$, $\hat{T}_1 = \bar{T}_1$, and $\hat{T}_2 = \bar{T}_2$. That means, these two engine performances are compared under the same four corner points and subsequently "the same rectangular cycle" in the $(k, \bar{T})$ and $(k, \hat{T})$ spaces. It is similar to comparing Eq. (19) with Eq. (4) under the condition $k_1 = k_1'$ and $k_2 = k_2'$ in Zakine's paper [8].

The discussion of the influence of the active noise on the macroscopic efficiency in Eq. (A11) can be extended to that on the stochastic efficiency in Fig. 1. By integrating over all trajectories with the same value of the stochastic efficiency $\tilde{\eta}$ in $S_i(x)$ in Fig. 1, one can obtain the statistics of efficiencies $P_i(\tilde{\eta})$. When the trajectories are long,

$$P_i(\tilde{\eta}) \sim e^{-tJ(\tilde{\eta})}, \tag{A14}$$

where t is the time span of the trajectory and $J(\tilde{\eta})$ denotes a large deviation function in the large deviation theory [35]. This $P_i(\tilde{\eta})$ behaves asymptotically like a delta function peaked at the macroscopic efficiency $\tilde{\eta} = \eta_i$. If $J(\tilde{\eta})$ can be expressed in terms of $f(k)$ or other characteristics of the heat bath, it would provide information on how an active noise can affect the statistics of $\tilde{\eta}$.

To compare the Stirling efficiency with the Carnot efficiency under an active noise, let us extend the argument presented below Eq. (4) in Zakine's paper [8]. If a perfect regenerator can absorb all the heat released during the isochoric cooling D → A and inject it back into the engine during the isochoric heating B → C, then $\langle Q_{\rm input}\rangle = \langle Q_{\rm BC}\rangle + \langle Q_{\rm CD}\rangle$ in Eq. (A11) will reduce to $\langle Q_{\rm input}\rangle = \langle Q_{\rm CD}\rangle$. Consequently, the efficiency $\bar{\eta}^{\rm S}_{\rm act}$ of the active Stirling engine in Eq. (A11) will be replaced by

$$\begin{aligned}\bar{\eta}^{\rm C}_{\rm act} &= -\frac{\langle W\rangle}{\langle Q_{\rm input}\rangle}\\ &= -\frac{\langle W_{\rm AB}\rangle + \langle W_{\rm BC}\rangle + \langle W_{\rm CD}\rangle + \langle W_{\rm DA}\rangle}{\langle Q_{\rm CD}\rangle}\\ &= -\frac{\bar{T}_2[G(k_1) - G(k_2)] - \bar{T}_1[G(k_1) - G(k_2)]}{-\bar{T}_1[f(k_1) - f(k_2)] + \bar{T}_1[G(k_1) - G(k_2)]}\\ &= \frac{\left(1 - \frac{\bar{T}_2}{\bar{T}_1}\right)[G(k_1) - G(k_2)]}{-[f(k_1) - f(k_2)] + [G(k_1) - G(k_2)]}\\ &= \eta^{\rm C}_{\rm pas}\frac{\int_{k_2}^{k_1}\frac{f(k)}{k}dk}{f(k_2) - f(k_1) + \int_{k_2}^{k_1}\frac{f(k)}{k}dk}.\end{aligned} \tag{A15}$$

Here, $\eta^{\rm C}_{\rm pas} \equiv 1 - \frac{\bar{T}_2}{\bar{T}_1}$ has the same form as the efficiency $1 - \frac{T_2}{T_1}$ of the passive Carnot engine in the (k, T) space, while $\bar{\eta}^{\rm C}_{\rm act}$ is the efficiency of the active Carnot engine in the $(k, \bar{T})$ space. Since $\langle Q_{\rm BC}\rangle > 0$,

$$\bar{\eta}^{\rm C}_{\rm act} = -\frac{\langle W_{\rm AB}\rangle + \langle W_{\rm BC}\rangle + \langle W_{\rm CD}\rangle + \langle W_{DA}\rangle}{\langle Q_{\rm CD}\rangle} \tag{A16}$$

is an upper bound of

$$\bar{\eta}^{\rm S}_{\rm act} = -\frac{\langle W_{\rm AB}\rangle + \langle W_{\rm BC}\rangle + \langle W_{\rm CD}\rangle + \langle W_{DA}\rangle}{\langle Q_{\rm BC}\rangle + \langle Q_{\rm CD}\rangle}. \tag{A17}$$

When $f(k) \equiv 1$, one obtains $\bar{\eta}^{\rm C}_{\rm act} = \eta^{\rm C}_{\rm pas}$, which recovers the special case of the passive noise described below Eq. (4) in Ref. [8].

To calculate the deviation between the Stirling and the Carnot efficiency, it is more meaningful to compare $\bar{\eta}^{\rm S}_{\rm act}$ and $\bar{\eta}^{\rm C}_{\rm act}$, both pertaining to the active noise, rather than comparing $\bar{\eta}^{\rm S}_{\rm act}$ and $\eta^{\rm C}_{\rm pas}$ between the active and the passive noise.

Together with Eq. (A11), that deviation will be

$$\bar{\eta}^{\mathrm{C}}_{\mathrm{act}} - \bar{\eta}^{\mathrm{S}}_{\mathrm{act}} = \eta^{\mathrm{C}}_{\mathrm{pas}} \frac{\int_{k_2}^{k_1} \frac{f(k)}{k} dk}{f(k_2) - f(k_1) + \int_{k_2}^{k_1} \frac{f(k)}{k} dk} - \eta^{\mathrm{C}}_{\mathrm{pas}} \frac{\int_{k_2}^{k_1} \frac{f(k)}{k} dk}{f(k_2) - \frac{\bar{T}_2}{\bar{T}_1} f(k_1) + \int_{k_2}^{k_1} \frac{f(k)}{k} dk}$$
$$= \left(\eta^{\mathrm{C}}_{\mathrm{pas}}\right)^2 f(k_1) \int_{k_2}^{k_1} \frac{f(k)}{k} dk \times \frac{\left[f(k_2) - \frac{\bar{T}_2}{\bar{T}_1} f(k_1) + \int_{k_2}^{k_1} \frac{f(k)}{k} dk\right]^{-1}}{\left[f(k_2) - f(k_1) + \int_{k_2}^{k_1} \frac{f(k)}{k} dk\right]}, \tag{A18}$$

which is positive as mentioned above.

The revealed effect of active baths on the Stirling efficiency discussed above also indicates the effect on the power of the engine. For a heat engine with a four-process cycle, suppose the operating time for the ith process is $\tau_i < \infty$, where $i = 1, \dots, 4$. Their heat and work,

$$W_i \approx W_i^{\infty} + \frac{A_i}{\tau_i} \quad \text{and} \quad Q_i \approx Q_i^{\infty} - \frac{B_i}{\tau_i}, \tag{A19}$$

can be generalized from Ref. [50]. It yields the power

$$P = -\frac{W}{\tau} = -\frac{\sum_{i=1}^4 W_i}{\tau} \approx -\frac{\sum_{i=1}^4 W_i^{\infty}}{\tau} - \frac{\sum_{i=1}^4 A_i}{\tau^2}, \tag{A20}$$

where $\tau = \sum_{i=1}^4 \tau_i$. In the example of the Stirling engine,

$$\sum_{i=1}^4 W_i^{\infty} = (\bar{T}_2 - \bar{T}_1)\frac{k_B}{2} \int_{k_2}^{k_1} \frac{f(k)}{k} dk < 0 \tag{A21}$$

has been derived in Eq. (A7). However, the dependence of $\sum_{i=1}^4 A_i$, and consequently P, on $f(k)$ is unknown. Nevertheless, if τ is sufficiently large, such that $-\frac{\sum_{i=1}^4 A_i}{\tau^2}$ in P is negligible, as in the quasistatic regime, then

$$P \approx -\frac{1}{\tau} \frac{k_B(\bar{T}_2 - \bar{T}_1)}{2} \int_{k_2}^{k_1} \frac{f(k)}{k} dk. \tag{A22}$$

In this case, we know how to adjust the function $f(k)$ of an active noise to enhance the power P and surpass the power of its passive counterpart. It can be achieved when the $f(k)$ of the active noise is an increasing function, starting from $f(0) = 1$, like that in Fig. 4, within the operating range $[k_2, k_1]$ of the engine so that $\int_{k_2}^{k_1} \frac{f(k)}{k} dk > \int_{k_2}^{k_1} \frac{1}{k} dk$.

To understand the maximum power under active noises, let us use the low dissipation Carnot engine to explain it [50]. According to that study, the heat exchanges during the isothermal processes at low and high temperatures, T_c and T_h, are

$$Q_c = Q_c^{\infty} - \frac{B_c}{\tau_c} + \cdots = T_c\left(-\Delta S - \frac{\Sigma_c}{\tau_c} + \cdots\right),$$
$$Q_h = Q_h^{\infty} - \frac{B_h}{\tau_h} + \cdots = T_h\left(\Delta S - \frac{\Sigma_h}{\tau_h} + \cdots\right). \tag{A23}$$

Here, B_c, B_h, Σ_c, and Σ_h are some coefficients, τ_c (τ_h) denotes the operation time during the cold (hot) isothermal process, and S stands for the entropy during those processes. Imposing the condition of zero derivatives, $\frac{\partial P}{\partial \tau_h} = \frac{\partial P}{\partial \tau_c} = 0$, for the maximum P^* of the power

$$P = -\frac{W}{\tau_h + \tau_c} = \frac{Q_h + Q_c}{\tau_h + \tau_c} = \frac{(T_h - T_c)\Delta S - \frac{T_h \Sigma_h}{\tau_h} - \frac{T_c \Sigma_c}{\tau_c}}{\tau_h + \tau_c} \tag{A24}$$

yields the durations

$$\tau_h^* = \frac{2T_h \Sigma_h}{(T_h - T_c)\Delta S}\left(1 + \sqrt{\frac{T_c \Sigma_c}{T_h \Sigma_h}}\right),$$
$$\tau_c^* = \frac{2T_c \Sigma_c}{(T_h - T_c)\Delta S}\left(1 + \sqrt{\frac{T_h \Sigma_h}{T_c \Sigma_c}}\right) \tag{A25}$$

of P^*. The efficiency $\eta = -\frac{W}{Q_h} = 1 + \frac{Q_c}{Q_h}$ at τ_h^* and τ_c^* is then

$$\eta^* = \frac{\eta_C\left(1 + \sqrt{\frac{T_c \Sigma_c}{T_h \Sigma_h}}\right)}{\left(1 + \sqrt{\frac{T_c \Sigma_c}{T_h \Sigma_h}}\right)^2 + \frac{T_c}{T_h}\left(1 - \frac{\Sigma_c}{\Sigma_h}\right)}. \tag{A26}$$

When $\Sigma_c = \Sigma_h$, η^* recovers the Curzon-Ahlborn efficiency $\eta_{CA} = 1 - \sqrt{\frac{T_c}{T_h}}$.

For active noises, the condition $\frac{\partial P}{\partial \tau_h} = \frac{\partial P}{\partial \tau_c} = 0$ to maximize the power P is the same, although their Σ_c, Σ_h, and ΔS in Eq. (A24) will generally be different from those of the passive noise. For the Curzon-Ahlborn efficiency under active noises, first notice that τ_h, τ_c, and ΔS are independent variables for general P and η. However, due to $\tau_h^* \sim \frac{1}{\Delta S}$ and $\tau_c^* \sim \frac{1}{\Delta S}$ in Eq. (A25), P^* and η^* interestingly become independent of ΔS. It implies that the variation of η^* with $f(k)$ is not determined by the variations of $Q_c^{\infty} = T_c \Delta S$ and $Q_h^{\infty} = T_h \Delta S$ with $f(k)$, as indicated in Eq. (A7). This analysis shows that the impact of an active noise on the Curzon-Ahlborn efficiency occurs in the variations of Σ_c and Σ_h with $f(k)$. The entire argument for the Carnot engine applies in a similar manner to the Stirling engine, both with and without a perfect regenerator.

APPENDIX B: PARTICLE SUBJECT TO THE OSCILLATORY ORNSTEIN-UHLENBECK NOISE

Let us consider the Langevin equation

$$\dot{x} + \frac{k}{\gamma} x = \frac{\xi}{\gamma}, \tag{B1}$$

where k and γ are as defined before Eq. (A1) and ξ represents a noise. Suppose that noise is a Gaussian colored noise with the correlation

$$\langle\xi(t)\xi(t')\rangle = \bar{C}e^{-|t-t'|/\tau}\cos[\omega_0(t-t')]$$
$$= 2\gamma(1+\omega_0^2\tau^2)k_B\bar{T}\underbrace{\frac{e^{-|t-t'|/\tau}}{2\tau}}_{\delta\text{-function if }\tau\to 0}\cos[\omega_0(t-t')], \tag{B2}$$

where

$$\bar{C} \equiv \frac{\gamma(1+\omega_0^2\tau^2)k_B\bar{T}}{\tau}. \tag{B3}$$

Therein τ and ω_0 characterize the correlation and the oscillation effect of $\langle\xi(t)\xi(t')\rangle$, respectively. $\bar{T}$ denotes the temperature defined by diffusion, as explained before Eq. (A1), when a free Brownian particle is subject to the colored noise ξ. To show that $\bar{C}$ is indeed related to such $\bar{T}$ by Eq. (B3), let us consider the mean-square displacement of Eq. (B1) at $k=0$,

$$\langle x^2(t)\rangle = \frac{1}{\gamma^2}\int_0^t\int_0^t\langle\xi(s)\xi(s')\rangle ds ds'$$
$$= \frac{1}{\gamma^2}\int_0^t\int_0^t\bar{C}e^{-|s-s'|/\tau}\cos[\omega_0(s-s')]ds ds'$$
$$= \frac{\bar{C}}{\gamma^2}\int_0^t\left\{\int_0^{s'}e^{(s-s')/\tau}\cos[\omega_0(s-s')]ds\right.$$
$$\left.+\int_{s'}^t e^{-(s-s')/\tau}\cos[\omega_0(s-s')]ds\right\}ds'$$
$$= \frac{\bar{C}}{\gamma^2}\int_0^t\left\{\int_{-s'}^0 e^{s/\tau}\cos(\omega_0 s)ds\right.$$
$$\left.+\int_0^{t-s'}e^{-s/\tau}\cos(\omega_0 s)ds\right\}ds'$$
$$= \frac{\bar{C}}{\gamma^2}\frac{1}{(1+\omega_0^2\tau^2)^2}\big[-2\tau^2+2\omega_0^2\tau^4+2\omega_0^2\tau^3 t$$
$$+2\tau t+2\tau^2 e^{-t/\tau}\cos(\omega_0 t)-4\omega_0\tau^3 e^{-t/\tau}\sin(\omega_0 t)$$
$$-2\omega_0^2\tau^4 e^{-t/\tau}\cos(\omega_0 t)\big], \tag{B4}$$

where the last equality has been confirmed by a symbolic calculation (see B4 within the Supplemental Material [49]). At large t, Eq. (B4) behaves as

$$\langle x^2(t)\rangle \approx 2\frac{\bar{C}(\omega_0^2\tau^3+\tau)}{\gamma^2(1+\omega_0^2\tau^2)^2}t \equiv 2Dt, \tag{B5}$$

with $D \equiv \bar{C}\tau\gamma^{-2}(1+\omega_0^2\tau^2)^{-1}$ the diffusion constant under the colored noise ξ. Expressing it in the form of the Einstein relation $D = k_B\bar{T}/\gamma$, mentioned before Eq. (A1), shows that $\bar{C}$ is indeed related to $\bar{T}$ by Eq. (B3).

For $k>0$, recall that the Fourier transforms of $f(t)\cos(bt)$ and $e^{-|t|/\tau}$ are $F(f(t)\cos(bt)) = [\hat{f}(\omega-b)+\hat{f}(\omega+b)]/2$ and $F(e^{-|t|/\tau}) = (2/\tau)[(1/\tau)^2+\omega^2]^{-1}$, respectively, where $\hat{f}(\omega)$ is the Fourier transform of $f(t)$. Therefore, the Fourier transform of Eq. (B2) at $t'=0$ is

$$F(\langle\xi(t)\xi(0)\rangle)$$
$$= \bar{C}F(e^{-|t|/\tau}\cos(\omega_0 t))$$
$$= \frac{\bar{C}}{2}\left[\frac{2/\tau}{(1/\tau)^2+(\omega-\omega_0)^2}+\frac{2/\tau}{(1/\tau)^2+(\omega+\omega_0)^2}\right]$$
$$= \tau\bar{C}\left(\frac{1}{1+(\omega-\omega_0)^2\tau^2}+\frac{1}{1+(\omega+\omega_0)^2\tau^2}\right).$$

Subsequently, the power spectrum of x in Eq. (B1) is [for more discussion, see Eq. (F6)]

$$S_x(\omega) = \frac{F(\langle\xi(t)\xi(0)\rangle)}{k^2+\omega^2\gamma^2}$$
$$= \frac{1}{(k^2+\omega^2\gamma^2)}\frac{2\tau\bar{C}[1+(\omega^2+\omega_0^2)\tau^2]}{[1+(\omega-\omega_0)^2\tau^2][1+(\omega+\omega_0)^2\tau^2]}$$
$$= 2\tau\bar{C}\left(\frac{\gamma}{k\tau}\right)\left(\frac{k\tau}{\gamma}\right)\frac{(\tau^2/\gamma^2)}{(k^2\tau^2/\gamma^2+\omega^2\tau^2)}$$
$$\times\frac{1+\omega^2\tau^2+\omega_0^2\tau^2}{[1+(\omega\tau-\omega_0\tau)^2][1+(\omega\tau+\omega_0\tau)^2]}\left(\frac{1}{\tau}\right)\tau$$
$$= \frac{2\tau\bar{C}}{\gamma k}\frac{K}{K^2+w^2}\frac{1+w^2+w_0^2}{[1+(w-w_0)^2][1+(w+w_0)^2]}\tau, \tag{B6}$$

where $K\equiv k\tau/\gamma$, $w\equiv\omega\tau$, and $w_0\equiv\omega_0\tau$. As a result, the variance of x is (see B7 within the Supplemental Material [49])

$$\langle x^2\rangle = \frac{1}{2\pi}\int_{-\infty}^{\infty}S_x(\omega)d\omega$$
$$= \frac{1}{k}\frac{\tau\bar{C}}{\pi\gamma}K\int_{-\infty}^{\infty}\frac{1+w^2+w_0^2}{(K^2+w^2)[1+(w-w_0)^2][1+(w+w_0)^2]}dw$$
$$= \frac{1}{k}\frac{\tau\bar{C}}{\pi\gamma}K\frac{(K+1)\pi}{K(K^2+2K+1+w_0^2)}$$
$$= \frac{1}{k}\frac{\tau}{\gamma}\frac{\gamma(1+\omega_0^2\tau^2)k_B\bar{T}}{\tau}\frac{(K+1)}{(K+1)^2+w_0^2}$$
$$= \frac{k_B\bar{T}}{k}\frac{(1+w_0^2)(K+1)}{(K+1)^2+w_0^2} = \frac{k_B\bar{T}}{k}\frac{(c+1)(K+1)}{(K+1)^2+c}, \tag{B7}$$

where $dw\equiv d(\omega\tau)$ and $c\equiv w_0^2>0$. Consequently,

$$f(k) = \frac{\hat{T}}{\bar{T}} = \frac{k\langle x^2\rangle}{k_B\bar{T}} = \frac{(c+1)(K+1)}{(K+1)^2+c} \equiv \hat{f}(K). \tag{B8}$$

Since K is proportional to k, the condition for an increasing $f(k)$ is equivalent to that for an increasing $\hat{f}(K)$.

In the limit $\omega_0\to 0$, Eqs. (B2) and (B7) reduce to

$$\langle\xi(t)\xi(t')\rangle = \frac{\gamma k_B\bar{T}}{\tau}e^{-|t-t'|/\tau},$$
$$\langle x^2\rangle = \frac{k_B\bar{T}}{k(1+K)} = \frac{k_B\bar{T}}{k(1+k\tau/\gamma)} \tag{B9}$$

of a persistent noise [8]. Notably, due to a different expression of the Langevin equation in Ref. [8], the $\langle\eta_p(t)\eta_p(t')\rangle$ in Eq. (12) of that paper corresponds to $\langle\xi(t)\xi(t')\rangle/\gamma^2$ here. With Eq. (B9), Eq. (B8) is simplified to

$$f(k) = \frac{1}{1+K} \equiv \hat{f}(K), \tag{B10}$$

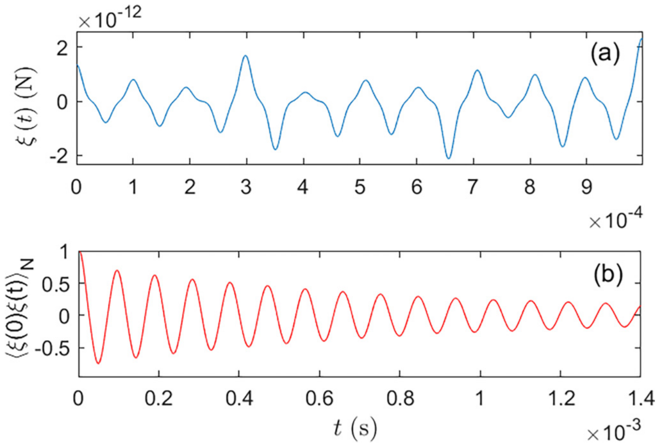

FIG. 6. A noise with a dominant low frequency. The noise composed of a series of alternating peaks in (a) has a normalized autocorrelation in (b) similar to Eq. (B2).

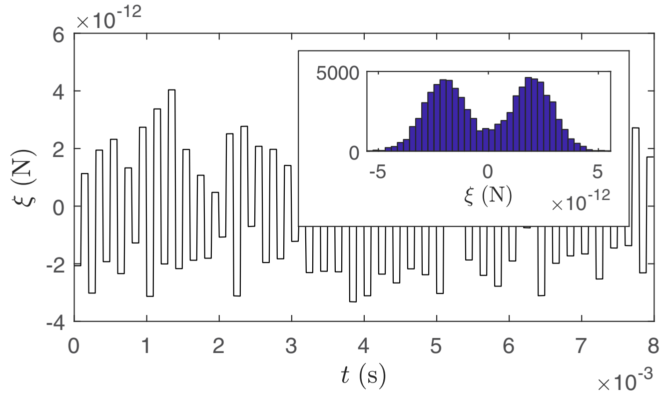

FIG. 7. An oscillatory telegraphic noise, whose magnitudes have a double-Gaussian distribution with two peaks symmetrically located at two nonzero positions (inset).

which is a decreasing function of K. However, for $\omega_0 > 0$, $\hat{f}(K)$ in Eq. (B8) can be an increasing function in the region where its slope $d\hat{f}(K)/dK = -(c+1)[(K+1)^2 - c][(K+1)^2 + c]^{-2} > 0$, which has its maximum value $(c-1)/(c+1)$ at $K = 0$. This derivative is positive when $(K+1)^2 < c$, or equivalently $K < \sqrt{c} - 1$, because both $K+1$ and $\sqrt{c} = w_0 > 0$. Since $K = k\tau/\gamma$ and $w_0 = \omega_0\tau$, it implies $k < (\omega_0 - 1/\tau)\gamma$, which is the range for finding an increasing $\hat{f}(K)$ in Eq. (B8).

In experiments of colloidal engines, typical parameter values include $\gamma = 10^{-8}$ kg/s, $k \in [1, 4] \times 10^{-5}$ N/m, and the data acquisition time 10^{-5} s, which is less than the particle relaxation time $t_{\rm relax} \equiv \gamma/k \in [2.5 \times 10^{-4}, 10^{-3}]$ s. For $(\omega_0, \tau) = (10^4 s^{-1}, 10^{-3}\mathrm{s})$, since $k \leqslant (\omega_0 - 1/\tau)\gamma = 9 \times 10^{-5}$ N/m, one can find an increasing $\hat{f}(K)$. Even for weaker correlation times, such as $\tau = 5 \times 10^{-4}$ s, $\tau = 4 \times 10^{-4}$ s, and $\tau = 3 \times 10^{-4}$ s, an increasing trend of $\hat{f}(K)$ can still be clearly seen, as the blue ($c = 25$), violet ($c = 16$), and red ($c = 9$) lines, respectively, shown in Fig. 4(a). In this plot, k is extended to 2×10^{-3} N/m to show how $\hat{f}(K)$ changes from a rising to a falling function.

As an example of ξ in Eq. (B2), Fig. 6(a) shows a noise composed of a series of alternating peaks, whose shapes, heights, and spacings between two neighboring peaks are all Gaussian distributed. Its correlation in (b) is similar to that in Eq. (B2). If the peaks in (a) become denser, the noise will have a larger ω_0, which enhances the chance of finding an increasing $\hat{f}(K)$ and a high-efficiency engine.

APPENDIX C: PARTICLE SUBJECT TO THE OSCILLATORY TELEGRAPHIC NOISE

An oscillatory telegraphic noise consists of a series of telegraphic steps that have piecewise constant noise magnitudes (Fig. 7). Its magnitudes form a double-Gaussian distribution (inset in Fig. 7). If ξ in Eq. (B1) is of this kind of noise, the change of the particle position from x_i at time t_i to x_{i+1} at time t_{i+1} under a given k is an integration of Eq. (B1), $x_{i+1} - x_i = -\gamma^{-1}\int_{t_i}^{t_{i+1}} kx(t)\,dt + \gamma^{-1}\int_{t_i}^{t_{i+1}} \xi(t)dt$, where $i = 0,\ 1,\ 2, \ldots$. If the span $\varepsilon_{\rm t} = t_{i+1} - t_i$ are identical and small for all i, then the range $[t_i, t_{i+1}] = [i\varepsilon_{\rm t}, (i+1)\varepsilon_{\rm t}]$ and $\int_{t_i}^{t_{i+1}} kx(t)\,dt \approx kx_i\varepsilon_{\rm t}$. When $\xi(t)$ has a constant magnitude ξ_i in that range, $w_i \equiv \gamma^{-1}\int_{t_i}^{t_{i+1}} \xi(t)dt = \gamma^{-1}\int_{i\varepsilon_{\rm t}}^{(i+1)\varepsilon_{\rm t}} \xi_i dt = \varepsilon_{\rm t}\xi_i/\gamma$. Therefore,

$$x_{i+1} = Rx_i + w_i, \tag{C1}$$

with $R \equiv 1 - k\varepsilon_{\rm t}/\gamma$. Consequently,

$$\begin{aligned}
x_1 &= Rx_0 + w_0\\
&\overset{\times R^N}{\Longrightarrow} R^N x_1 = R^{N+1}x_0 + R^N w_0,\\
x_2 &= Rx_1 + w_1\\
&\overset{\times R^{N-1}}{\Longrightarrow} R^{N-1}x_2 = R^N x_1 + R^{N-1}w_1,\\
&\quad\vdots\\
x_{N-1} &= Rx_{N-2} + w_{N-2}\\
&\overset{\times R^2}{\Longrightarrow} R^2 x_{N-1} = R^3 x_{N-2} + R^2 w_{N-2},\\
x_N &= Rx_{N-1} + w_{N-1}\\
&\overset{\times R^1}{\Longrightarrow} Rx_N = R^2 x_{N-1} + Rw_{N-1}.
\end{aligned} \tag{C2}$$

Summing up all equations behind the arrows and starting with $x_0 = 0$, one obtains $Rx_N = Rw_{N-1} + R^2 w_{N-2} + \cdots + R^{N-1}w_1 + R^N w_0$ and accordingly

$$\begin{aligned}
x_N &= w_{N-1} + Rw_{N-2} + \cdots + R^{N-2}w_1 + R^{N-1}w_0\\
&= \sum_{i=1}^{N} R^{i-1} w_{N-i} = \sum_{i=1}^{N} R^{i-1}\frac{\varepsilon_{\rm t}\xi_{N-i}}{\gamma}\\
&= \frac{\varepsilon_{\rm t}}{\gamma}\sum_{i=1}^{N} R^{i-1}\xi_{N-i}.
\end{aligned} \tag{C3}$$

If we reexpress the position x_N at time $t = N\varepsilon_{\rm t}$ as $x(t)$, the ensemble average of the product $x(t)x(t')$ will be

$$\langle x(t)x(t')\rangle = \langle x_N x_{N'}\rangle = \frac{{\varepsilon_{\rm t}}^2}{\gamma^2}\sum_{i=1}^{N}\sum_{j=1}^{N'} M_{ij}, \tag{C4}$$

with $M_{ij} \equiv R^{(i+j-2)}\langle\xi_{N-i}\xi_{N-j}\rangle$ the (i, j)th entry of an $N \times N'$ matrix $\mathbf{M}$. Therefore, $\langle x(t)x(t')\rangle$ is the sum of all entries of

$\mathbf{M}$. For a given $t = N\varepsilon_{\mathrm{t}}$ and $t' = N'\varepsilon_{\mathrm{t}}$, the value of R and the sizes N and N' of $\mathbf{M}$ depend on the value of the numerical parameter ε_{t}. However, $\langle x(t)x(t')\rangle$ should be almost invariant of that parameter, when ε_{t} is sufficiently small. For $N = N'$, $\mathbf{M}$ becomes a square matrix of dimension N,

$$\mathbf{M} = \begin{bmatrix} \langle\xi_{N-1}\xi_{N-1}\rangle & R\langle\xi_{N-1}\xi_{N-2}\rangle & R^2\langle\xi_{N-1}\xi_{N-3}\rangle & & & \\ R\langle\xi_{N-2}\xi_{N-1}\rangle & R^2\langle\xi_{N-2}\xi_{N-2}\rangle & & \cdots & \cdots & \\ R^2\langle\xi_{N-3}\xi_{N-1}\rangle & & & & & \\ & \vdots & & \ddots & \vdots & \\ & & & & & R^{2N-4}\langle\xi_2\xi_0\rangle \\ & \vdots & & \cdots & R^{2N-4}\langle\xi_1\xi_1\rangle & R^{2N-3}\langle\xi_1\xi_0\rangle \\ & & & R^{2N-4}\langle\xi_0\xi_2\rangle & R^{2N-3}\langle\xi_0\xi_1\rangle & R^{2N-2}\langle\xi_0\xi_0\rangle \end{bmatrix}. \tag{C5}$$

Let the magnitude of the nth telegraphic step of the oscillatory telegraphic noise in Fig. 8 be

$$\xi^{(n)} = \bar{\xi}^{(n)} + \xi^{\mathrm{G}(n)} \quad (n = 1,\ 2,\ 3, \ldots). \tag{C6}$$

Here, $\bar{\xi}^{(n)} = (-1)^n\bar{\xi}$ is a deterministic term, which oscillates between two values $\pm\bar{\xi}$, depicted by the two dashed lines in that plot. $\xi^{\mathrm{G}(n)}$ is a Gaussian random variable, which causes deviations from the dashed lines. The correlation of $\xi^{\mathrm{G}(n)}$ is $\langle\xi^{\mathrm{G}(n)}\xi^{G(m)}\rangle = \Xi\delta^K_{nm}$, which gives the variance $\Xi = \langle\xi^{\mathrm{G}(n)2}\rangle$, where δ^K_{nm} is the Kronecker delta. Thus, the correlation between two telegraphic steps is $\langle\xi^{(n)}\xi^{(m)}\rangle = (-1)^{n+m}\bar{\xi}^2 + \Xi\delta^K_{nm}$. If we use this noise to drive a Stirling engine in the $(k, \bar{T})$ space, the temperature T in Fig. 8 refers to $\bar{T} = \bar{T}_{\mathrm{e}}$ in Eq. (C28), where $\bar{T}_{\mathrm{e}} = \Xi t_{\mathrm{s}}/(2\gamma k_{\mathrm{B}})$ in Eq. (C17).

Suppose that an oscillatory telegraphic noise has n_{s} telegraphic steps in each of the four processes in an engine cycle and each step is discretized as n_{d} spans of length ε_{t}. Figure 8 shows an example of $(n_{\mathrm{d}}, n_{\mathrm{s}}) = (3,\ 4)$ in a cycle. Therein the noise magnitudes $\xi_i = \bar{\xi}_i + \xi^{\mathrm{G}}_i$ at the ith time instant ($i = 0,\ 1,\ 2, \ldots$) are denoted by red stars. During the two isothermal processes, if the ith instant belongs to the nth telegraphic step, ξ_i, $\bar{\xi}_i$, and ξ^{G}_i have magnitudes $\xi^{(n)}$, $\bar{\xi}^{(n)}$, and $\xi^{\mathrm{G}(n)}$ in Eq. (C6), respectively. During the iso-k heating and cooling processes, ξ_i in a telegraphic step should, in principle, vary slightly with the temperature T, as seen in $\xi_{12} \sim \xi_{23}$ and $\xi_{36} \sim \xi_{47}$. Specifically, ξ_i should deviate more from (come closer to) $\bar{\xi}_i = +\bar{\xi}$ or $-\bar{\xi}$ when T becomes larger (smaller). However, since the variation of T is very small within a short tilted telegraphic step, in simulations all ξ_i in the nth step can be approximated by the same $\xi^{(n)}$ of a flat step, whose $\xi^{\mathrm{G}(n)}$ is determined by the T at the beginning of the tilted step.

For $(n_{\mathrm{d}}, n_{\mathrm{s}}) = (2,\ 4)$ and $N = 8$, Eq. (C3) gives

$$\begin{aligned} x_8 &= w_7 + Rw_6 + \cdots + R^6 w_1 + R^7 w_0 \\ &= \left(\frac{\xi_7\varepsilon_{\mathrm{t}}}{\gamma}\right) + R\left(\frac{\xi_6\varepsilon_{\mathrm{t}}}{\gamma}\right) + \cdots + R^7\left(\frac{\xi_0\varepsilon_{\mathrm{t}}}{\gamma}\right) \\ &= \frac{\varepsilon_{\mathrm{t}}}{\gamma}\big[\big(\bar{\xi}_7 + \xi^{\mathrm{G}}_7\big) + R\big(\bar{\xi}_6 + \xi^{\mathrm{G}}_6\big) + +R^2\big(\bar{\xi}_5 + \xi^{\mathrm{G}}_5\big) \\ &\quad + R^3\big(\bar{\xi}_4 + \xi^{\mathrm{G}}_4\big) + R^4\big(\bar{\xi}_3 + \xi^{\mathrm{G}}_3\big) + R^5\big(\bar{\xi}_2 + \xi^{\mathrm{G}}_2\big) \\ &\quad + R^6\big(\bar{\xi}_1 + \xi^{\mathrm{G}}_1\big) + R^7\big(\bar{\xi}_0 + \xi^{\mathrm{G}}_0\big)\big] \\ &= \frac{\varepsilon_{\mathrm{t}}}{\gamma}\big[\big(+\bar{\xi} + \xi^{\mathrm{G}}_7\big) + R\big(+\bar{\xi} + \xi^{\mathrm{G}}_6\big) + R^2\big(-\bar{\xi} + \xi^{\mathrm{G}}_5\big) \\ &\quad + R^3\big(-\bar{\xi} + \xi^{\mathrm{G}}_4\big) + R^4\big(+\bar{\xi} + \xi^{\mathrm{G}}_3\big) + R^5\big(+\bar{\xi} + \xi^{\mathrm{G}}_2\big) \\ &\quad + R^6\big(-\bar{\xi} + \xi^{\mathrm{G}}_1\big) + R^7\big(-\bar{\xi} + \xi^{\mathrm{G}}_0\big)\big], \end{aligned} \tag{C7}$$

with $\xi^{\mathrm{G}}_0 = \xi^{\mathrm{G}}_1 = \xi^{\mathrm{G}(1)}$, $\xi^{\mathrm{G}}_2 = \xi^{\mathrm{G}}_3 = \xi^{\mathrm{G}(2)}$, $\xi^{\mathrm{G}}_4 = \xi^{\mathrm{G}}_5 = \xi^{\mathrm{G}(3)}$, and $\xi^{\mathrm{G}}_6 = \xi^{\mathrm{G}}_7 = \xi^{\mathrm{G}(4)}$. Its $\mathbf{M}$ in Eq. (C5) is then the 8×8

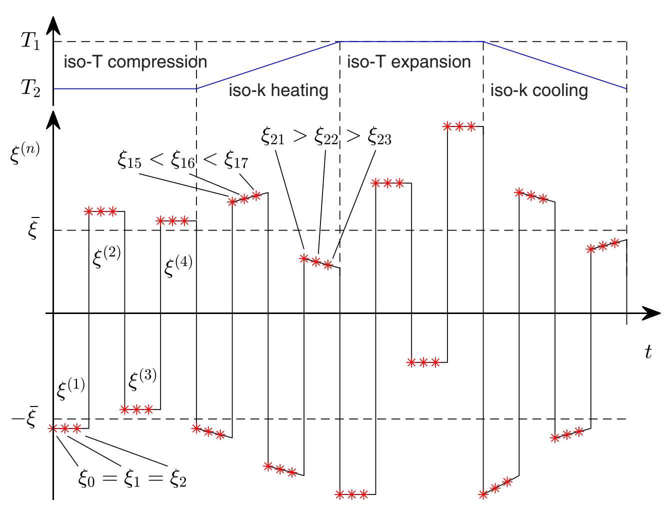

FIG. 8. An oscillatory telegraphic noise of four telegraphic steps ($n_{\mathrm{s}} = 4$) and three discretized spans ($n_{\mathrm{d}} = 3$) in each of the four processes of a Stirling cycle (Fig. 5). The noise magnitudes ξ_0, ξ_1, and ξ_2 in the first telegraphic step have the same value $\xi^{(1)}$ in Eq. (C6), the noise magnitudes ξ_3, ξ_4, and ξ_5 in the second step have the same value $\xi^{(2)}$, and so on. The ξ presented here is regarded as having a phase zero. If ξ_i at t_i are replaced by ξ_{i+1} at t_{i+1}, the phase of ξ is shifted 60°.

matrix

$$
\begin{aligned}
\mathbf{M} &= \begin{bmatrix}
\langle\xi_7\xi_7\rangle & R\langle\xi_7\xi_6\rangle & R^2\langle\xi_7\xi_5\rangle & R^3\langle\xi_7\xi_4\rangle & R^4\langle\xi_7\xi_3\rangle & R^5\langle\xi_7\xi_2\rangle & R^6\langle\xi_7\xi_1\rangle & R^7\langle\xi_7\xi_0\rangle \\
R\langle\xi_6\xi_7\rangle & R^2\langle\xi_6\xi_6\rangle & R^3\langle\xi_6\xi_5\rangle & R^4\langle\xi_6\xi_4\rangle & R^5\langle\xi_6\xi_3\rangle & R^6\langle\xi_6\xi_2\rangle & R^7\langle\xi_6\xi_1\rangle & R^8\langle\xi_6\xi_0\rangle \\
R^2\langle\xi_5\xi_7\rangle & R^3\langle\xi_5\xi_6\rangle & R^4\langle\xi_5\xi_5\rangle & R^5\langle\xi_5\xi_4\rangle & R^6\langle\xi_5\xi_3\rangle & R^7\langle\xi_5\xi_2\rangle & R^8\langle\xi_5\xi_1\rangle & R^9\langle\xi_5\xi_0\rangle \\
R^3\langle\xi_4\xi_7\rangle & R^4\langle\xi_4\xi_6\rangle & R^5\langle\xi_4\xi_5\rangle & R^6\langle\xi_4\xi_4\rangle & R^7\langle\xi_4\xi_3\rangle & R^8\langle\xi_4\xi_2\rangle & R^9\langle\xi_4\xi_1\rangle & R^{10}\langle\xi_4\xi_0\rangle \\
R^4\langle\xi_3\xi_7\rangle & R^5\langle\xi_3\xi_6\rangle & R^6\langle\xi_3\xi_5\rangle & R^7\langle\xi_3\xi_4\rangle & R^8\langle\xi_3\xi_3\rangle & R^9\langle\xi_3\xi_2\rangle & R^{10}\langle\xi_3\xi_1\rangle & R^{11}\langle\xi_3\xi_0\rangle \\
R^5\langle\xi_2\xi_7\rangle & R^6\langle\xi_2\xi_6\rangle & R^7\langle\xi_2\xi_5\rangle & R^8\langle\xi_2\xi_4\rangle & R^9\langle\xi_2\xi_3\rangle & R^{10}\langle\xi_2\xi_2\rangle & R^{11}\langle\xi_2\xi_1\rangle & R^{12}\langle\xi_2\xi_0\rangle \\
R^6\langle\xi_1\xi_7\rangle & R^7\langle\xi_1\xi_6\rangle & R^8\langle\xi_1\xi_5\rangle & R^9\langle\xi_1\xi_4\rangle & R^{10}\langle\xi_1\xi_3\rangle & R^{11}\langle\xi_1\xi_2\rangle & R^{12}\langle\xi_1\xi_1\rangle & R^{13}\langle\xi_1\xi_0\rangle \\
R^7\langle\xi_0\xi_7\rangle & R^8\langle\xi_0\xi_6\rangle & R^9\langle\xi_0\xi_5\rangle & R^{10}\langle\xi_0\xi_4\rangle & R^{11}\langle\xi_0\xi_3\rangle & R^{12}\langle\xi_0\xi_2\rangle & R^{13}\langle\xi_0\xi_1\rangle & R^{14}\langle\xi_0\xi_0\rangle
\end{bmatrix} \\
&= \begin{bmatrix}
(\bar\xi^2+\Xi) & R(\bar\xi^2+\Xi) & -\bar\xi^2R^2 & -\bar\xi^2R^3 & \bar\xi^2R^4 & \bar\xi^2R^5 & -\bar\xi^2R^6 & -\bar\xi^2R^7 \\
R(\bar\xi^2+\Xi) & R^2(\bar\xi^2+\Xi) & -\bar\xi^2R^3 & -\bar\xi^2R^4 & \bar\xi^2R^5 & \bar\xi^2R^6 & -\bar\xi^2R^7 & -\bar\xi^2R^8 \\
-\bar\xi^2R^2 & -\bar\xi^2R^3 & R^4(\bar\xi^2+\Xi) & R^5(\bar\xi^2+\Xi) & -\bar\xi^2R^6 & -\bar\xi^2R^7 & \bar\xi^2R^8 & \bar\xi^2R^9 \\
-\bar\xi^2R^3 & -\bar\xi^2R^4 & R^5(\bar\xi^2+\Xi) & R^6(\bar\xi^2+\Xi) & -\bar\xi^2R^7 & -\bar\xi^2R^8 & \bar\xi^2R^9 & \bar\xi^2R^{10} \\
\bar\xi^2R^4 & \bar\xi^2R^5 & -\bar\xi^2R^6 & -\bar\xi^2R^7 & R^8(\bar\xi^2+\Xi) & R^9(\bar\xi^2+\Xi) & -\bar\xi^2R^{10} & -\bar\xi^2R^{11} \\
\bar\xi^2R^5 & \bar\xi^2R^6 & -\bar\xi^2R^7 & -\bar\xi^2R^8 & R^9(\bar\xi^2+\Xi) & R^{10}(\bar\xi^2+\Xi) & -\bar\xi^2R^{11} & -\bar\xi^2R^{12} \\
-\bar\xi^2R^6 & -\bar\xi^2R^7 & \bar\xi^2R^8 & \bar\xi^2R^9 & -\bar\xi^2R^{10} & -\bar\xi^2R^{11} & R^{12}(\bar\xi^2+\Xi) & R^{13}(\bar\xi^2+\Xi) \\
-\bar\xi^2R^7 & -\bar\xi^2R^8 & \bar\xi^2R^9 & \bar\xi^2R^{10} & -\bar\xi^2R^{11} & -\bar\xi^2R^{12} & R^{13}(\bar\xi^2+\Xi) & R^{14}(\bar\xi^2+\Xi)
\end{bmatrix} \\
&= \begin{bmatrix}
(\bar\xi^2+\Xi)\mathbf{E} & -\bar\xi^2R^2\mathbf{E} & \bar\xi^2R^4\mathbf{E} & -\bar\xi^2R^6\mathbf{E} \\
-\bar\xi^2R^2\mathbf{E} & (\bar\xi^2+\Xi)R^4\mathbf{E} & -\bar\xi^2R^6\mathbf{E} & \bar\xi^2R^8\mathbf{E} \\
\bar\xi^2R^4\mathbf{E} & -\bar\xi^2R^6\mathbf{E} & (\bar\xi^2+\Xi)R^8\mathbf{E} & -\bar\xi^2R^{10}\mathbf{E} \\
-\bar\xi^2R^6\mathbf{E} & \bar\xi^2R^8\mathbf{E} & -\bar\xi^2R^{10}\mathbf{E} & (\bar\xi^2+\Xi)R^{12}\mathbf{E}
\end{bmatrix} \\
&= \bar\xi^2\begin{bmatrix}
\mathbf{E} & -R^2\mathbf{E} & R^4\mathbf{E} & -R^6\mathbf{E} \\
-R^2\mathbf{E} & R^4\mathbf{E} & -R^6\mathbf{E} & R^8\mathbf{E} \\
R^4\mathbf{E} & -R^6\mathbf{E} & R^8\mathbf{E} & -R^{10}\mathbf{E} \\
-R^6\mathbf{E} & R^8\mathbf{E} & -R^{10}\mathbf{E} & R^{12}\mathbf{E}
\end{bmatrix} + \Xi\begin{bmatrix}
\mathbf{E} & 0 & 0 & 0 \\
0 & R^4\mathbf{E} & 0 & 0 \\
0 & 0 & R^8\mathbf{E} & 0 \\
0 & 0 & 0 & R^{12}\mathbf{E}
\end{bmatrix},
\end{aligned}
\tag{C8}
$$

where the 2×2 submatrix $\mathbf{E} \equiv \left[\begin{smallmatrix} R^0 & R^1 \\ R^1 & R^2 \end{smallmatrix}\right]$ gives a contribution from a telegraphic step with $n_\mathrm{d} = 2$.

For general n_d and even number of n_s, the noise has $n_\mathrm{s}/2$ pairs of forward and backward telegraphic steps. In this case, the matrix in Eq. (C8) can be extended to

$$
\mathbf{M} = \bar\xi^2\begin{bmatrix}
\mathbf{E} & -R^{n_\mathrm{d}}\mathbf{E} & \dots & (-R^{n_\mathrm{d}})^{n_\mathrm{s}-1}\mathbf{E} \\
-R^{n_\mathrm{d}}\mathbf{E} & (-R^{n_\mathrm{d}})^2\mathbf{E} & & \vdots \\
\vdots & & \ddots & (-R^{n_\mathrm{d}})^{2n_\mathrm{s}-3}\mathbf{E} \\
(-R^{n_\mathrm{d}})^{n_\mathrm{s}-1}\mathbf{E} & \dots & (-R^{n_\mathrm{d}})^{2n_\mathrm{s}-3}\mathbf{E} & (-R^{n_\mathrm{d}})^{2n_\mathrm{s}-2}\mathbf{E}
\end{bmatrix} + \Xi\begin{bmatrix}
\mathbf{E} & 0 & 0 & 0 \\
0 & R^{2n_\mathrm{d}}\mathbf{E} & 0 & 0 \\
0 & 0 & \ddots & 0 \\
0 & 0 & 0 & R^{(2n_\mathrm{s}-2)n_\mathrm{d}}\mathbf{E}
\end{bmatrix},
\tag{C9}
$$

where

$$
\mathbf{E} \equiv \begin{bmatrix}
R^0 & R^1 & \dots & R^{n_\mathrm{d}-2} & R^{n_\mathrm{d}-1} \\
R^1 & R^2 & \dots & R^{n_\mathrm{d}-1} & R^{n_\mathrm{d}} \\
\vdots & & \ddots & & \vdots \\
R^{n_\mathrm{d}-2} & R^{n_\mathrm{d}-1} & & R^{2n_\mathrm{d}-4} & R^{2n_\mathrm{d}-3} \\
R^{n_\mathrm{d}-1} & R^{n_\mathrm{d}} & \dots & R^{2n_\mathrm{d}-3} & R^{2n_\mathrm{d}-2}
\end{bmatrix}.
$$

Removing the $n_\mathrm{d} \times n_\mathrm{d}$ submatrix $\mathbf{E}$ from the $n_\mathrm{d} n_\mathrm{s} \times n_\mathrm{d} n_\mathrm{s}$ matrix $\mathbf{M}$ yields an $n_\mathrm{s} \times n_\mathrm{s}$ matrix

$$\mathbf{M}' \equiv \bar{\xi}^2 \underbrace{\begin{bmatrix} 1 & -R^{n_\mathrm{d}} & \dots & (-R^{n_\mathrm{d}})^{n_\mathrm{s}-1} \\ -R^{n_\mathrm{d}} & (-R^{n_\mathrm{d}})^2 & & \vdots \\ \vdots & & \ddots & (-R^{n_\mathrm{d}})^{2n_\mathrm{s}-3} \\ (-R^{n_\mathrm{d}})^{n_\mathrm{s}-1} & \dots & (-R^{n_\mathrm{d}})^{2n_\mathrm{s}-3} & (-R^{n_\mathrm{d}})^{2n_\mathrm{s}-2} \end{bmatrix}}_{\equiv \mathbf{M}_\mathrm{I}} + \Xi \underbrace{\begin{bmatrix} 1 & 0 & & 0 & 0 \\ 0 & R^{2n_\mathrm{d}} & & 0 & 0 \\ 0 & 0 & \ddots & & 0 \\ 0 & 0 & & 0 & R^{(2n_\mathrm{s}-2)n_\mathrm{d}} \end{bmatrix}}_{\mathbf{M}_\mathrm{II}}$$

$$= \bar{\xi}^2 \mathbf{M}_\mathrm{I} + \Xi \mathbf{M}_\mathrm{II}, \tag{C10}$$

with $\mathbf{M}_\mathrm{I}$ and $\mathbf{M}_\mathrm{II}$ two additional submatrixes. Let the sum of all entries in $\mathbf{M}$, $\mathbf{E}$, $\mathbf{M}'$, $\mathbf{M}_\mathrm{I}$, and $\mathbf{M}_\mathrm{II}$, be $S_\mathbf{M}$, $S_\mathbf{E}$, $S_{\mathbf{M}'}$, $S_{\mathbf{M}_\mathrm{I}}$, and $S_{\mathbf{M}_\mathrm{II}}$, respectively. Clearly, they are related by

$$S_\mathbf{M} = S_{\mathbf{M}'} S_\mathbf{E} = (\bar{\xi}^2 S_{\mathbf{M}_\mathrm{I}} + \Xi S_{\mathbf{M}_\mathrm{II}}) S_\mathbf{E}. \tag{C11}$$

Since $1 + r + \cdots + r^n = (1 - r^{n+1})/(1 - r)$ for $r < 1$, we have

$$S_\mathbf{E} \equiv (1 + R + R^2 + \cdots + R^{n_\mathrm{d}-1})^2 = \begin{cases} {n_\mathrm{d}}^2 & \text{for } R = 1 \\ \left(\frac{1-R^{n_\mathrm{d}}}{1-R}\right)^2 & \text{for } R < 1, \end{cases} \tag{C12}$$

$$S_{\mathbf{M}_\mathrm{I}} \equiv (1 + (-R^{n_\mathrm{d}}) + \cdots + (-R^{n_\mathrm{d}})^{n_\mathrm{s}-1})^2 = \begin{cases} \frac{1+(-1)^{n_\mathrm{s}-1}}{2} & \text{for } R = 1 \\ \left(\frac{1-(-R^{n_\mathrm{d}})^{n_\mathrm{s}}}{1+R^{n_\mathrm{d}}}\right)^2 & \text{for } R < 1 \Rightarrow |-R^{n_\mathrm{d}}| < 1, \end{cases} \tag{C13}$$

$$S_{\mathbf{M}_\mathrm{II}} = 1 + R^{2n_\mathrm{d}} + R^{4n_\mathrm{d}} + \cdots + R^{(2n_\mathrm{s}-2)n_\mathrm{d}} = \begin{cases} n_\mathrm{s} & \text{for } R = 1 \\ \frac{1-R^{2n_\mathrm{s}n_\mathrm{d}}}{1-R^{2n_\mathrm{d}}} & \text{for } R < 1, \end{cases} \tag{C14}$$

where $S_{\mathbf{M}_\mathrm{I}} = 0$ for $R = 1$ in Eq. (C13) if n_s there is an even number. Let $\langle x^2(t)\rangle_\mathrm{e}$ be the variance of x calculated from the matrices in Eqs. (C8) and (C9), for which t is at the end of a pair of telegraphic steps and n_s is an even number. Due to Eqs. (C4) and (C11)~(C14), one obtains

$$\begin{aligned} \langle x^2(t)\rangle_\mathrm{e} &= \frac{{\varepsilon_\mathrm{t}}^2}{\gamma^2} S_\mathbf{M} = \frac{{\varepsilon_\mathrm{t}}^2}{\gamma^2} (\bar{\xi}^2 S_{\mathbf{M}_\mathrm{I}} + \Xi S_{\mathbf{M}_\mathrm{II}}) S_\mathbf{E} \\ &= \frac{{\varepsilon_\mathrm{t}}^2}{\gamma^2} \left[\bar{\xi}^2 \begin{cases} \frac{1+(-1)^{n_\mathrm{s}-1}}{2} \\ \left(\frac{1-(-R^{n_\mathrm{d}})^{n_\mathrm{s}}}{1+R^{n_\mathrm{d}}}\right)^2 \end{cases} + \Xi \begin{cases} n_\mathrm{s} \\ \frac{1-R^{2n_\mathrm{s}n_\mathrm{d}}}{1-R^{2n_\mathrm{d}}} \end{cases} \right] \times \begin{cases} {n_\mathrm{d}}^2 \\ \left(\frac{1-R^{n_\mathrm{d}}}{1-R}\right)^2 \end{cases} \\ &= \begin{cases} \left[\bar{\xi}^2 \frac{1+(-1)^{n_\mathrm{s}-1}}{2} + \Xi n_\mathrm{s}\right] {n_\mathrm{d}}^2 \frac{{\varepsilon_\mathrm{t}}^2}{\gamma^2} \\ \left[\bar{\xi}^2 \left(\frac{1-(-R^{n_\mathrm{d}})^{n_\mathrm{s}}}{1+R^{n_\mathrm{d}}}\right)^2 + \Xi \frac{1-R^{2n_\mathrm{s}n_\mathrm{d}}}{1-R^{2n_\mathrm{d}}}\right] (1 - R^{n_\mathrm{d}})^2 \left(\frac{\varepsilon_\mathrm{t}}{1-R}\right)^2 \frac{1}{\gamma^2} \end{cases} \\ &= \begin{cases} \left[\bar{\xi}^2 \frac{1+(-1)^{n_\mathrm{s}-1}}{2} + \Xi \left(\frac{t}{t_\mathrm{s}}\right)\right] \frac{t_\mathrm{s}^2}{\gamma^2} & \text{for } R = 1 \\ \left[\bar{\xi}^2 (1 - (-1)^{n_\mathrm{s}} R^N)^2 \left(\frac{1-R^{n_\mathrm{d}}}{1+R^{n_\mathrm{d}}}\right)^2 + \Xi \left(1 - R^{2N}\right) \left(\frac{1-R^{n_\mathrm{d}}}{1+R^{n_\mathrm{d}}}\right)\right] \frac{1}{k^2} & \text{for } R < 1, \end{cases} \end{aligned} \tag{C15}$$

with $n_s = t/t_s$, $n_d\varepsilon_t = t_s$, and $n_s n_d = N$, where

$$\left(\frac{\varepsilon_t}{1-R}\right)^2\frac{1}{\gamma^2} = \left(\frac{\varepsilon_t}{k\varepsilon_t/\gamma}\right)^2\frac{1}{\gamma^2} = \frac{1}{k^2}$$

has been used in the last equality of Eq. (C15). Here, t and t_s are two physical parameters, while $n_d \to \infty$ and $\varepsilon_t \to 0$ are two numerical parameters, which have disappeared in Eq. (C15). The long-time behavior of $\langle x^2(t)\rangle_e$ in Eq. (C15) is as follows.

For $R = 1$, the first part of Eq. (C15) at $t \gg 1$ behaves asymptotically as

$$\begin{aligned}\langle x^2(t)\rangle_{e,\text{free}} &\equiv \Xi\left(\frac{t}{t_s}\right)\frac{t_s^2}{\gamma^2} = 2\left(\frac{\Xi t_s}{2\gamma^2}\right)t\\ &= 2Dt \equiv 2\left(\frac{k_B\bar{T}_e}{\gamma}\right)t,\end{aligned} \tag{C16}$$

where D and $\bar{T}_e$ are defined by diffusion, as in the case of D and $\bar{T}$ before Eq. (A1). Therefore, $\bar{T}_e$ is related to the noise characteristics Ξ and t_s and the mean-square displacement $\langle x^2(t)\rangle_{e,\text{free}}$ by

$$\bar{T}_e = \frac{\Xi t_s}{2\gamma k_B} = \frac{\gamma}{2k_B t}\langle x^2(t)\rangle_{e,\text{free}}. \tag{C17}$$

Here, the second equality has been confirmed numerically. Interestingly, the temperature $\bar{T}_e$ perceived by a stochastic heat engine can be controlled by Ξ or t_s, but not by $\bar{\xi}$.

For $R < 1$, consider the limit $(n_d, \varepsilon_t) \to (\infty, 0)$ under a constant $t_s = n_d\varepsilon_t$ and $t = N\varepsilon_t = n_s n_d \varepsilon_t = n_s t_s$. Using $\exp x = \lim_{n\to\infty}(1+\frac{x}{n})^n$ for all real x and $R = 1 - k\varepsilon_t/\gamma = 1 + (-kt_s/\gamma)/n_d = 1 + X/n_d$, with the dimensionless quantity $X \equiv -kt_s/\gamma$, we obtain

$$R^{n_d} = \left(1+\frac{X}{n_d}\right)^{n_d} \xrightarrow{n_d\to\infty} e^X = e^{-kt_s/\gamma}, \tag{C18}$$

$$\begin{aligned}R^N &= \left(1-\frac{k}{\gamma}\varepsilon_t\right)^{t/\varepsilon_t} = \left[\left(1-\frac{k/\gamma}{1/\varepsilon_t}\right)^{1/\varepsilon_t}\right]^t\\ &\xrightarrow{\varepsilon_t\to 0} [e^{-k/\gamma}]^t = e^{-kt/\gamma} \xrightarrow{t\to\infty} 0,\end{aligned} \tag{C19}$$

$$R^{2N} = [R^N]^2 \xrightarrow{\varepsilon_t\to 0} [e^{-kt/\gamma}]^2 = e^{-2kt/\gamma} \xrightarrow{t\to\infty} 0, \tag{C20}$$

$$\frac{1-R^{n_d}}{1+R^{n_d}} \xrightarrow{n_d\to\infty} \tanh\left(-\frac{X}{2}\right) = \tanh\left(\frac{kt_s}{2\gamma}\right), \tag{C21}$$

where Eq. (C18) has been used to derive Eq. (C21). Therefore, in the limit $(n_d, \varepsilon_t) \to (\infty, 0)$, the second part of Eq. (C15) for $R < 1$ becomes

$$\begin{aligned}\langle x^2(t)\rangle_e = \bigg[&\bar{\xi}^2(1-(-1)^{n_s}e^{-kt/\gamma})^2\tanh^2\left(\frac{kt_s}{2\gamma}\right)\\ &+ Xi(1-e^{-2kt/\gamma})\tanh\left(\frac{kt_s}{2\gamma}\right)\bigg]\frac{1}{k^2}.\end{aligned} \tag{C22}$$

At $t \to \infty$, Eq. (C22) reduces to

$$\langle x^2\rangle_{e,\text{well}} \equiv \left[\bar{\xi}^2\tanh^2\left(\frac{kt_s}{2\gamma}\right) + \Xi\tanh\left(\frac{kt_s}{2\gamma}\right)\right]\frac{1}{k^2}, \tag{C23}$$

which implies

$$\begin{aligned}\hat{T}_e &= k\langle x^2\rangle_{e,\text{well}}/k_B\\ &= \left[\bar{\xi}^2\tanh^2\left(\frac{kt_s}{2\gamma}\right) + \Xi\tanh\left(\frac{kt_s}{2\gamma}\right)\right]\frac{1}{kk_B}.\end{aligned} \tag{C24}$$

Unlike that $\bar{T}_e$ in Eq. (C17) can only be varied by Ξ and t_s, all the three parameters Ξ, t_s, and $\bar{\xi}$ can be used to manipulate $\hat{T}_e$ in Eq. (C24). This allows us to use $\bar{\xi}$ to tune $\hat{T}_e$ without changing $\bar{T}_e$. Numerically, the second equality in Eq. (C24) has been confirmed by our simulation.

With Eqs. (C17) and (C24), one obtains

$$\begin{aligned}f_e(k) &= \frac{\hat{T}_e}{\bar{T}_e} = \frac{\left[\bar{\xi}^2\tanh^2\left(\frac{kt_s}{2\gamma}\right) + \Xi\tanh\left(\frac{kt_s}{2\gamma}\right)\right]\frac{1}{kk_B}}{\Xi t_s/(2\gamma k_B)}\\ &= \frac{1}{kt_s/(2\gamma)}\left[\frac{\bar{\xi}^2}{\Xi}\tanh^2\left(\frac{kt_s}{2\gamma}\right) + \tanh\left(\frac{kt_s}{2\gamma}\right)\right]\\ &= \frac{1}{K}[c\tanh^2(K) + \tanh(K)]\\ &= \frac{\tanh(K)}{K}[1 + c\tanh(K)] \equiv \hat{f}_e(K).\end{aligned} \tag{C25}$$

Here, $K \equiv kt_s/(2\gamma)$ and $c \equiv \bar{\xi}^2/\Xi$ are two dimensionless parameters, where $\Xi = 2\gamma k_B\bar{T}/t_s$ due to Eq. (C17), with $\bar{T} = \bar{T}_e$ as explained in Eq. (C28) below. Interestingly, $\hat{f}_e(K)$ depends more crucially on the ratio $\bar{\xi}^2/\Xi$ than the individual values of $\bar{\xi}^2$ and Ξ. Since $\tanh(K) \approx K$ at small K, Eq. (C25) behaves as $\hat{f}_e(K) \approx 1 + cK$ in that regime. A larger c will enhance the increasing trend of $\hat{f}_e(K)$, which is beneficial for improving the engine efficiency. Furthermore, a smaller t_s will pull the value of $K = kt_s/(2\gamma)$ back to a smaller number, where the slope of $\hat{f}_e(K)$ is steeper, which would also improve the efficiency.

In Eq. (C25), $\bar{\xi} > 0$ is crucial for obtaining an increasing $\hat{f}_e(K)$. If $\bar{\xi} = 0$, Eq. (C25) will reduce to a decreasing function $\hat{f}_e(K) = \tanh(K)/K$. However, if $\bar{\xi} = 0$ is combined with $n_d = 1$ (only one red star in each telegraphic step in Fig. 8), taking $t \to \infty$ for $\langle x^2(t)\rangle_e$ in Eq. (C15) will give

$$\Xi\left(\frac{t}{t_s}\right)\frac{t_s^2}{\gamma^2} = 2\left(\frac{\Xi t_s}{2\gamma^2}\right)t \equiv 2\left(\frac{k_B\bar{T}_e}{\gamma}\right)t \quad \text{for } R = 1,$$

$$\Xi(1-R^{2N})\left(\frac{1-R}{1+R}\right)\frac{1}{k^2} \xrightarrow{N\to\infty} \left(\frac{1-R}{1+R}\right)\frac{\Xi}{k^2} \equiv \frac{k_B\hat{T}_e}{k} \quad \text{for } R < 1. \tag{C26}$$

It yields $\bar{T}_e = \Xi t_s/(2\gamma k_B)$, the same as Eq. (C17), and $\hat{T}_e = (1-R)(1+R)^{-1}\Xi/(kk_B)$, different from Eq. (C24). Together with $R = 1 - k\varepsilon_t/\gamma$, $n_d = t_s/\varepsilon_t = 1$, and $K \equiv kt_s/(2\gamma) = k\varepsilon_t/(2\gamma)$, they lead to an increasing function of K,

$$\begin{aligned}f_{e,0}(k) &= \frac{\hat{T}_e}{\bar{T}_e} = \left(\frac{1-R}{1+R}\right)\frac{\Xi}{kk_B}\frac{2\gamma k_B}{\Xi t_s}\\ &= \left(\frac{k\varepsilon_t/\gamma}{2-k\varepsilon_t/\gamma}\right)\frac{1}{k\varepsilon_t/(2\gamma)}\\ &= \frac{2}{2-k\varepsilon_t/\gamma} = \frac{1}{1-K} \equiv \hat{f}_e(K),\end{aligned} \tag{C27}$$

which is even simpler than Eq. (C25). However, this increasing $f_{e,0}(k)$ cannot provide a high-efficiency engine because $K \sim \varepsilon_t$, as explained in Sec. III C.

Recall that Eq. (C25) is a result of $\mathbf{M}$ in Eq. (C9), which is derived under three conditions: (i) k is fixed or changes quasistatically, (ii) n_s is an even number such that t contains $n_s/2$ pairs of backward and forward telegraphic steps, and (iii) ξ starts with the initial value $\bar{\xi}^{(1)} = -\bar{\xi}$ in Eq. (C6). At that t, a particle in the overdamped regime has been pushed to a positive extreme position, $x > 0$, far away from the equilibrium point at $x = 0$ and thus has a large $\langle x^2\rangle$. If n_s is an odd number, the noise just completes a backward step at t, so that the particle will be at a negative extreme position, $x < 0$, again far from $x = 0$. In this case, $\langle x^2(t)\rangle_{\mathrm{e,free}}$, $\bar{T}_e$, $\langle x^2\rangle_{\mathrm{e,well}}$, and $\hat{T}_e$ in Eqs. (C16), (C17), (C23), and (C24), respectively, remain the same. Due to symmetry, these results still hold when the initial value $\bar{\xi}^{(1)} = -\bar{\xi}$ is replaced by $+\bar{\xi}$. At other t, which contain incomplete telegraphic steps, $|x|$ is generically smaller than the absolute values of the above extreme positions. Therefore, its variance in the well, $\langle x^2\rangle_{\mathrm{well}}$ and $\hat{T} = k\langle x^2\rangle_{\mathrm{well}}/k_B$ are generically smaller than $\langle x^2\rangle_{\mathrm{e,well}}$ in Eq. (C23) and $\hat{T}_e$ in Eq. (C24), respectively. Its variance in a potential-free space, $\langle x^2(t)\rangle_{\mathrm{free}}$, is also different from $\langle x^2(t)\rangle_{\mathrm{e,free}}$. However, at large t, the former has the same asymptotic behavior as the latter in Eq. (C16). Therefore, the temperatures defined by $\langle x^2(t)\rangle_{\mathrm{free}}$ and $\langle x^2(t)\rangle_{\mathrm{e,free}}$ are identical,

$$\bar{T} = \bar{T}_e. \tag{C28}$$

Putting all together, although $\hat{f}(K) = \hat{T}/\bar{T}$ is different from $\hat{f}_e(K) = \hat{T}_e/\bar{T}_e$, they have the same rising or falling trend. An example of $\langle x^2\rangle_{\mathrm{well}}$ and $\langle x^2\rangle_{\mathrm{e,well}}$ can be seen in Fig. 3(b), where the former denoted by $\langle x^2\rangle$ (blue) is, as expected above, "bounded" by the latter denoted by $\langle x^2\rangle_e$ (red).

So far, the results are obtained when the telegraphic noises ξ in all systems of the ensemble have the same phase starting with the initial value $\bar{\xi}^{(1)} = -\bar{\xi}$. When we randomize the phase of ξ as in the equal *a priori* probability in the microcanonical ensemble, the phase-averaged $\langle x^2(t)\rangle_{\mathrm{free}}$, $\langle x^2\rangle_{\mathrm{well}}$, $\hat{T}$, and $\hat{f}(K)$ become much smoother, which are "bounded" by, and follow the same trend as, the above $\langle x^2(t)\rangle_{\mathrm{e,free}}$, $\langle x^2\rangle_{\mathrm{e,well}}$, $\hat{T}_e$, and $\hat{f}_e(K)$, respectively. Notice that $\bar{T}$ and $\bar{T}_e$ in Eq. (C28) still have the same value after the phase is randomized because these temperatures depend only on the long-time spread of x and not on the phase. As $\bar{T}$ does not oscillate with ξ, it is an appropriate temperature in the engine protocol.

In colloidal engine experiments, typical or accessible ranges of parameter values include $\gamma = 10^{-8}$ kg/s, $k \in [1,\ 4] \times 10^{-5}$ N/m, and $(\bar{T}_2, \bar{T}_1) = (200,\ 400)$ K. To create these $\bar{T}$, let us consider the length of the telegraphic step, $t_s = n_d\varepsilon_t = 2 \times 10^{-4}$ s. It is of the same order as the relaxation time $t_{\mathrm{relax}} \equiv \gamma/k \in 10^{-8}/([1,\ 4] \times 10^{-5})$ s $= [2.5,\ 10] \times 10^{-4}$ s. Besides, $K = kt_s/(2\gamma) \in [1,\ 4] \times 10^{-5} \times 2 \times 10^{-4}/(2 \times 10^{-8}) = [0.1,\ 0.4]$ and $\Xi = 2\gamma k_B\bar{T}/t_s \in 2 \times 10^{-8} \times 1.38 \times 10^{-23} \times [200,\ 400]/(2 \times 10^{-4})\mathrm{N}^2 = [2.76,\ 5.52] \times 10^{-25}$ N^2. Taking $(\varepsilon_t,\ n_d,\ n_s,\ n_{\mathrm{cycle}}) = (10^{-5}\ \mathrm{s},\ 20, 10^2,\ 10^4)$, both $\varepsilon_t/t_{\mathrm{relax}} \in 10^{-5}/([2.5,\ 10] \times 10^{-4}) = [4, 1] \times 10^{-2}$ and $t_{\mathrm{relax}}/t_{\mathrm{process}} = t_{\mathrm{relax}}/(n_d n_s \varepsilon_t) \in [2.5,\ 10] \times 10^{-4}/(20 \times 10^2 \times 10^{-5}) = [1.25,\ 5.00] \times 10^{-2}$ are sufficiently small in simulations. For $\bar{\xi} = 1$ pN $= 10^{-12}$ N, its $\bar{\xi}^2 = 10^{-24}$ N^2 is larger than $\langle \xi^2_{\mathrm{thermal}}(t)\rangle = 2\gamma k_B T = 8.28 \times 10^{-29}$ N^2 of the thermal noise $\xi_{\mathrm{thermal}}(t)$ of water molecules at $T = 300$ K.

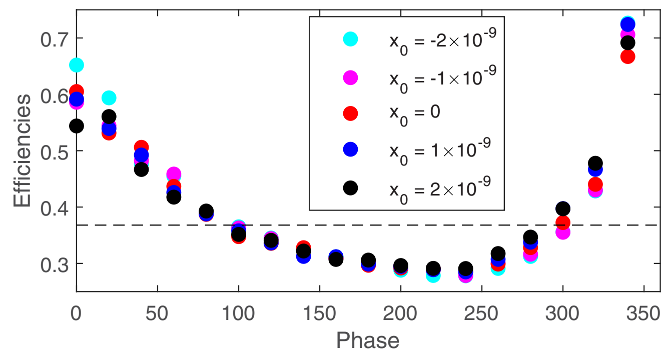

FIG. 9. Engine efficiencies under oscillatory telegraphic noises at different phases. The 90 points in the plot represent the efficiencies of an active engine driven by the oscillatory telegraphic noise of 18 uniformly distributed phases starting with five different initial particle positions x_0, each of which is averaged over 10^4 single cycles in an ensemble. The efficiencies of the same color have the same x_0. The dashed line denotes the efficiency 0.368 of the passive engine.

This $\bar{\xi}$ is comparable to the experimentally measured thrust forces of flagellar motors, including 0.57 pN for peritrichous (multiflagellated) *Escherichia coli*, 0.37 pN for *Salmonella typhimurium*, and 0.029 pN for the amphitrichous magnetotactic bacterium *Magnetospirillum magneticum* [51].

Taking these parameter values, we obtain Fig. 3, where the efficiencies of the passive and active Stirling engine are $(\eta^S_{\mathrm{pas}},\ \bar{\eta}^S_{\mathrm{act}}) = (0.37,\ 0.43)$. Since $c = \bar{\xi}^2/\Xi = \bar{\xi}^2/(2\gamma k_B\bar{T}/t_s) \in 10^{-24}/([2.76,\ 5.52] \times 10^{-25}) = [1.81,\ 3.62]$, its $\hat{f}_e(K)$ should have a rising branch. Indeed, if $c = 3.0$, 2.0, and 1.1, $\hat{f}_e(K)$ behaves as the blue, violet, and red lines, respectively, in Fig. 4(b). If $(\bar{T}_2, \bar{T}_1) = (200,\ 400)$ K is replaced by $(300,\ 400)$ K, the efficiency of the active engine will reduce to that of its passive counterpart, with $(\eta^S_{\mathrm{pas}},\ \bar{\eta}^S_{\mathrm{act}}) = (0.20,\ 0.20)$.

In the above simulations, all engine efficiencies are calculated by the long-time average over a large number of successive cycles, as commonly evaluated in experiments. Therein the phase of the noise in each cycle is zero, as shown in Fig. 8, while the initial position of the Brownian particle is set to zero, $x_0 = 0$ μm, only in the first cycle. These efficiencies can be compared with those calculated by the ensemble average. To this end, let us prepare 90 ensembles of single-cycle simulations from 18 uniformly distributed phases and 5 different initial particle positions x_0. Their efficiencies are indicated by 90 points in Fig. 9, where $\bar{\xi} = 1.6$ pN is used. These efficiencies are not always higher than the efficiency 0.368 of the passive engine (dashed line). However, taking the average of 18 phases for each of the five individual $x_0 = (-2 \times 10^{-9},\ -10^{-9}, 0,\ 10^{-9}, 2 \times 10^{-9})$ μm gives five efficiencies (0.399 0.393 0.392 0.398 0.395), all of which surpass the efficiency of the passive engine.

APPENDIX D: PARTICLE DYNAMICS IN THE OPTICAL FEEDBACK TRAP

In the OFT technique, one mimics the motion of a Brownian particle in an arbitrary potential $U(x)$ by instantaneously shifting a simpler potential $U_o(x)$ to some position, such that the two different potentials are the same, $U(x) = U_o(x)$, at the instant position x of the particle. The continuously moving potential $U_o(x)$ creates a constant virtual potential $U_v(x)$ to approximate the desired real fixed potential $U(x)$. In experiments, $U_o(x)$ is commonly a harmonic potential of stiffness k_{ot} given by an optical tweezers.

When a colloidal particle of mass m is immersed in water and confined by a harmonic potential $U(x) = kx^2/2$ of stiffness k, the particle position $x(t)$ will be governed by a Langevin equation $m\ddot{x}(t) = -\gamma\dot{x}(t) - kx(t) + \xi(t) \approx 0$, if the system is in the overdamped regime. Here, γ is the friction coefficient and $\xi(t) = \sqrt{2\gamma k_B T}\Gamma(t) = \sqrt{2\gamma^2 D}\Gamma(t)$ is the stochastic force of thermal noise, where T and D are the temperature of the water and the diffusion constant of the particle, respectively, and $\Gamma(t)$ denotes a Gaussian white noise of unit variance. In OFT, if the real force $-dU(x)/dx = -kx(t)$ is generated by the force, $-k_{ot}[x(t) - x_L(t)]$, of an optical tweezers centered at $x_L(t)$, they satisfy the relation $-kx(t) = -k_{ot}[x(t) - x_L(t)]$. It yields $x_L(t) = -\alpha x(t)$, where $\alpha \equiv -1 + k/k_{ot}$. Therefore, ideally, the particle should be governed by the Langevin dynamics $\gamma\dot{x}(t) + kx(t) = \gamma\dot{x}(t) + k_{ot}[x(t) + \alpha x(t)] = \xi(t)$. However, in practice, the location of the optical tweezers can only be adjusted according to an inevitably slightly earlier particle position, $x_L(t - t_d) = -\alpha x(t - t_d)$, where t_d is a feedback delay time. Thus, the real particle dynamics satisfies

$$\gamma\dot{x}(t) + k_{ot}x(t) + \alpha k_{ot}x(t - t_d) = \xi(t). \tag{D1}$$

If we express the mean-square displacement of Eq. (D1) as $\langle x^2(t)\rangle = 2\bar{D}t = 2(k_B\bar{T}/\gamma)t$, it defines a diffusion constant $\bar{D}$ and a corresponding temperature $\bar{T}$, as that before Eq. (A1), for $t_d > 0$. By definition, $\bar{D}$ and $\bar{T}$ will reduce to the above D and T only at $t_d \to 0$. But, in reality, $\bar{D} = D$ and $\bar{T} = T$ for any $t_d > 0$ because $\bar{D}$ and $\bar{T}$ are both defined at $k = 0$. In this case, $\alpha = -1$ and Eq. (D1) reduces to $\gamma\dot{x}(t) + k_{ot}[x(t) - x(t - t_d)] = \xi(t)$, which is used to mimic a free particle governed by the Langevin equation $\gamma\dot{x}(t) = \xi(t)$. Since the real potential is absent in the latter equation, the optical tweezers in the former one is redundant and can be removed by setting $k_{ot} = 0$. Therefore, the $\bar{D}$ and $\bar{T}$ of the former equation are the same as the D and T of the latter one even for $t_d > 0$.

Suppose that the particle positions x_i in experiments are measured at discrete time instants $i\varepsilon_t$, with $i = 0, 1, 2, \ldots$, where ε_t is the data acquisition time. Then they follow a discretized version of Eq. (D1), $\gamma(x_{i+1} - x_i)/\varepsilon_t + k_{ot}(x_i + \alpha x_{i-l}) = \xi_i$ [19], when the delay time in Eq. (D1) is $t_d = l\varepsilon_t$ for $l = 0, 1, 2, \ldots$. Here, $\xi_i = \sqrt{2\gamma k_B T/\varepsilon_t}\Gamma_i = \sqrt{2\gamma^2 D/\varepsilon_t}\Gamma_i$, where Γ_i has the same statistics as $\Gamma(t)$ defined before Eq. (D1). Therefore,

$$\begin{aligned} x_{i+1} &= x_i - \frac{k_{ot}}{\gamma}(x_i + \alpha x_{i-l})\varepsilon_t + \frac{\varepsilon_t}{\gamma}\xi_i \\ &= (1-\beta)x_i - \alpha\beta x_{i-l} + \xi_i^x, \end{aligned} \tag{D2}$$

with $\beta \equiv k_{ot}\varepsilon_t/\gamma$, the position fluctuation $\xi_i^x \equiv \varepsilon_t\xi_i/\gamma = \sqrt{2(k_B T/\gamma)\varepsilon_t}\Gamma_i = \sqrt{2D\varepsilon_t}\Gamma_i$, and $\langle\xi_i^x\xi_j^x\rangle = 2D\varepsilon_t\delta_{ij}^K$, where δ_{ij}^K is the Kronecker delta. For $l = 1$, or equivalently $t_d = \varepsilon_t$, Eq. (D2) reduces to

$$x_{i+1} = (1-\beta)x_i - \alpha\beta x_{i-1} + \xi_i^x, \tag{D3}$$

which has the variance

$$\begin{aligned} \langle x_{i+1}{}^2\rangle &= (1-\beta)^2\langle x_i{}^2\rangle + \alpha^2\beta^2\langle x_{i-1}{}^2\rangle \\ &\quad + \langle\xi_i^{x2}\rangle - 2\alpha\beta(1-\beta)\langle x_i x_{i-1}\rangle \\ &\quad - 2\alpha\beta\langle x_{i-1}\xi_i^x\rangle + 2(1-\beta)\langle x_i\xi_i^x\rangle. \end{aligned} \tag{D4}$$

Multiplying Eq. (D3) with x_i and taking the average on all terms gives

$$\langle x_{i+1}x_i\rangle = (1-\beta)\langle x_i{}^2\rangle - \alpha\beta\langle x_i x_{i-1}\rangle + \langle x_i\xi_i^x\rangle. \tag{D5}$$

Considering $\langle x_{i-1}\xi_i^x\rangle = \langle x_i\xi_i^x\rangle = 0$ due to causality, as well as $\langle x_{i+1}{}^2\rangle = \langle x_i^2\rangle = \langle x_{i-1}{}^2\rangle \equiv \langle x^2\rangle$ and $\langle x_{i+1}x_i\rangle = \langle x_i x_{i-1}\rangle$ for a steady state, Eqs. (D4) and (D5) become

$$\begin{aligned} \langle x^2\rangle &= (1-\beta)^2\langle x^2\rangle + \alpha^2\beta^2\langle x^2\rangle + \langle\xi_i^{x2}\rangle \\ &\quad - 2\alpha\beta(1-\beta)\langle x_i x_{i-1}\rangle, \end{aligned} \tag{D6}$$

$$\langle x_i x_{i-1}\rangle = \frac{(1-\beta)}{(1+\alpha\beta)}\langle x^2\rangle, \tag{D7}$$

respectively. Inserting Eq. (D7) and $\langle\xi_i^{x2}\rangle = 2D\varepsilon_t$ into Eq. (D6) gives

$$\begin{aligned} \langle x^2\rangle &= \frac{2D\varepsilon_t}{1-(1-\beta)^2 - \alpha^2\beta^2 + \frac{2\alpha\beta(1-\beta)^2}{(1+\alpha\beta)}} \\ &= \frac{2D\varepsilon_t(1+\alpha\beta)}{\beta(1-\alpha\beta)(1+\alpha)(2+\alpha\beta-\beta)} \\ &= \frac{2\frac{k_B T}{\gamma}\varepsilon_t(1-\beta+K)}{\frac{k_{ot}\varepsilon_t}{\gamma}(1+\beta-K)\frac{k}{k_{ot}}(2-\beta+K-\beta)} \\ &= \frac{-2k_B T(K-\beta+1)}{k(K-\beta-1)(K-2\beta+2)} \\ &= \frac{-2(K-c_1)}{(K-c_2)(K-c_3)}\frac{k_B T}{k}, \end{aligned} \tag{D8}$$

with $c_1 \equiv \beta - 1$, $c_2 \equiv \beta + 1$, $c_3 \equiv 2\beta - 2$, and $K \equiv k\varepsilon_t/\gamma$, where $\alpha\beta = (-1 + k/k_{ot})(k_{ot}\varepsilon_t/\gamma) = -k_{ot}\varepsilon_t/\gamma + k\varepsilon_t/\gamma = -\beta + K$ has been used. Notice that due to feedback delay, Eq. (D8) has deviated from the variance $\langle x^2\rangle = k_B T/k$ in the absence of delay. This deviation depends on k_{ot} and t_d, where $t_d = \varepsilon_t$ has been assumed before Eq. (D3). With Eq. (D8), it follows that

$$f(k) = \frac{k\langle x^2\rangle}{k_B\bar{T}} = \frac{-2(K-c_1)}{(K-c_2)(K-c_3)} \equiv \hat{f}(K), \tag{D9}$$

where $\bar{T} = T$, as explained below Eq. (D1), has been used.

Since the nonzero ε_t is in practice very small, both K and $\beta \ll 1$. Therefore, the real K is far from the zero point c_1 and the two singularities at c_2 and c_3 in Eq. (D9). In this regime, $\hat{f}(K)$ is clearly a rising function of K. For instance, when $\alpha = 0$, which is equivalent to $k = k_{ot}$ and $\beta = K$, the real potential is equal to the optical tweezers potential. In this case, $c_1 = K - 1$, $c_2 = K + 1$, and $c_3 = 2K - 2$ in Eq. (D9), which yields an increasing function

$$\hat{f}_0(K) = \frac{2}{2-K}. \tag{D10}$$

For the experimentally accessible values $\gamma = 10^{-8}$ kg/s, $k_{ot} = 3 \times 10^{-5}$ N/m, $k \in [1, 4] \times 10^{-5}$ N/m, and $\varepsilon_t =$

10^{-5} s, both $\beta = k_{\rm ot}\varepsilon_{\rm t}/\gamma$ and $K = k\varepsilon_{\rm t}/\gamma$ are positive and $\ll 1$, so that $c_1 = -1^+$, $c_2 = 1^+$, $c_3 = -2^+$, and $K = 0^+$, where a^+ denotes a number slightly larger than a. Therefore,

$$\frac{d\hat{f}(K)}{dK} = 2\frac{K^2 - 2c_1K + (c_1c_2 + c_1c_3 - c_2c_3)}{(K - c_2)^2(K - c_3)^2}$$
$$\approx 2\frac{0^2 - 2(-1)0 + (-1 + 2 + 2)}{1^2 2^2} = \frac{3}{2} > 0,$$

indicating that $\hat{f}(K)$ is an increasing function. However, $K \equiv k\varepsilon_{\rm t}/\gamma$ in Eqs. (D9) and (D10) shrinks to zero as $\varepsilon_{\rm t} \to 0$. For the same reason as for Eq. (C27), the rising $\hat{f}(K)$ in Eq. (D9) and $\hat{f}_0(K)$ in Eq. (D10) do not lead to a high-efficiency engine.

Beyond the discrete dynamics of a tiny $t_{\rm d}$ in Eq. (D3), let us rewrite the original continuous dynamics Eq. (D1) with an arbitrary $t_{\rm d}$ as the form

$$\dot{y}(t) + \gamma_1 y(t) + \gamma_2 y(t - t_{\rm d}) = \Gamma(t). \tag{D11}$$

Here, $y(t) \equiv \gamma x(t)/A$, $\gamma_1 \equiv k_{\rm ot}/\gamma$, $\gamma_2 \equiv \alpha k_{\rm ot}/\gamma = \alpha\gamma_1 = (-1 + k/k_{\rm ot})k_{\rm ot}/\gamma = -k_{\rm ot}/\gamma + k/\gamma = -\gamma_1 + k/\gamma$, and $\Gamma(t) = \xi(t)/A$ is the Gaussian white noise of unit variance before Eq. (D1), where $A \equiv \sqrt{2\gamma k_{\rm B}T}$. Under the conditions $\gamma_2 > \gamma_1 \geqslant 0$ and $0 \leqslant t_{\rm d}\sqrt{\gamma_2{}^2 - \gamma_1^2} < \cos^{-1}(-\gamma_1/\gamma_2) \leqslant \pi$, the variance of y is [26,27]

$$\langle y^2\rangle = \frac{\gamma_2 \sin\left(\sqrt{\gamma_2{}^2 - \gamma_1^2}t_{\rm d}\right) + \sqrt{\gamma_2{}^2 - \gamma_1^2}}{2\sqrt{\gamma_2{}^2 - \gamma_1^2}\left[\gamma_1 + \gamma_2\cos\left(\sqrt{\gamma_2{}^2 - \gamma_1^2}t_{\rm d}\right)\right]}. \tag{D12}$$

For $\gamma_1 > \gamma_2 \geqslant 0$ and $\gamma_1 = \gamma_2 \geqslant 0$, a similar formula can be derived [27]. The first condition for Eq. (D12) means $\alpha\gamma_1 > \gamma_1 \geqslant 0$, which occurs when $\alpha = -1 + k/k_{\rm ot} = (k - k_{\rm ot})/k_{\rm ot} > 1$ or equivalently $k > 2k_{\rm ot}$. The second condition means $0 \leqslant t_{\rm d}\sqrt{(\gamma_2 + \gamma_1)(\gamma_2 - \gamma_1)} = t_{\rm d}\sqrt{(k/\gamma)(-2\gamma_1 + k/\gamma)} = \sqrt{(kt_{\rm d}/\gamma)(kt_{\rm d}/\gamma - 2k_{\rm ot}t_{\rm d}/\gamma)} = \sqrt{K(K - 2c)} \equiv g(K) < \cos^{-1}(c/(c - K)) \leqslant \pi$, with $K \equiv kt_{\rm d}/\gamma$, $c \equiv k_{\rm ot}t_{\rm d}/\gamma = t_{\rm d}\gamma_1$, and $K - c = t_{\rm d}\gamma_2$. In terms of K and c, the first condition above is $K > 2c$ and Eq. (D12) can be reexpressed as

$$\langle y^2\rangle = \frac{t_{\rm d}}{2}\frac{t_{\rm d}\gamma_2 \sin\left(\sqrt{t_{\rm d}{}^2\gamma_2{}^2 - t_{\rm d}{}^2\gamma_1^2}\right) + \sqrt{t_{\rm d}{}^2\gamma_2{}^2 - t_{\rm d}{}^2\gamma_1^2}}{\sqrt{t_{\rm d}{}^2\gamma_2{}^2 - t_{\rm d}{}^2\gamma_1^2}\left[t_{\rm d}\gamma_1 + t_{\rm d}\gamma_2\cos\left(\sqrt{t_{\rm d}{}^2\gamma_2{}^2 - t_{\rm d}{}^2\gamma_1^2}\right)\right]}$$
$$= \frac{1}{k}\frac{\gamma K[(K - c)\sin(g(K)) + g(K)]}{2g(K)[c + (K - c)\cos(g(K))]}, \tag{D13}$$

and subsequently

$$\langle x^2\rangle = \frac{A^2}{\gamma^2}\langle y^2\rangle = \frac{2k_{\rm B}T}{\gamma}\langle y^2\rangle$$
$$= \frac{1}{k}\frac{k_{\rm B}TK[(K - c)\sin(g(K)) + g(K)]}{g(K)[c + (K - c)\cos(g(K))]}. \tag{D14}$$

Notice that K here is normally much larger than the K in Eq. (D8) because $t_{\rm d}$ here is generically much larger than $\varepsilon_{\rm t}$ there. Since $\bar{T} = T$, with the same reason mentioned below Eq. (D1), it follows from $f(k) = k\langle x^2\rangle/(k_{\rm B}\bar{T})$ that

$$f(k) = \frac{K[(K - c)\sin(g(K)) + g(K)]}{g(K)[c + (K - c)\cos(g(K))]} \equiv \hat{f}(K). \tag{D15}$$

Under the first condition for Eq. (D12), $k > 2k_{\rm ot}$, consider a K slightly larger than $2c$, such that $K - 2c \equiv e \approx 0^+$ and $g(K) = \sqrt{(2c + e)e} \approx 0^+$. Taylor expanding Eq. (D15) with respect to such a small $g(K)$ gives

$$\hat{f}(K) = \frac{K\left\{(c + e)\left[g(K) - \frac{(g(K))^3}{6} + O^5\right] + g(K)\right\}}{g(K)\left\{c + (c + e)\left[1 - \frac{(g(K))^2}{2} + O^4\right]\right\}}$$
$$= \frac{K\left\{(c + e)\left[1 - \frac{(g(K))^2}{6} + O^4\right] + 1\right\}}{\left\{c + (c + e)\left[1 - \frac{(g(K))^2}{2} + O^4\right]\right\}}$$
$$\approx \frac{(c + 1)}{2c}K, \tag{D16}$$

with $O^4 \equiv O((g(K))^4)$ and $O^5 \equiv O((g(K))^5)$. Equation (D16) is indeed an increasing function of K.

For $\gamma = 10^{-8}$ kg/s, $k \in [2.1,\ 4.0] \times 10^{-5}$ N/m, and $k_{\rm ot} = 10^{-5}$ N/m, if we take $t_{\rm d} = 10^{-4}$ s, which is longer than the data acquisition time 10^{-5} s, it would give $K = kt_{\rm d}/\gamma = [0.21, 0.40]$. If $t_{\rm d} = 1.6 \times 10^{-4}$, 1.8×10^{-4}, and 2.0×10^{-4} s, it yields $c = k_{\rm ot}t_{\rm d}/\gamma = 0.16$, 0.18, and 0.20, respectively, for Eq. (D15), as depicted in Fig. 4(d). These curves do not start with $K = 0$ because we need $K > 2c$, or equivalently $k > 2k_{\rm ot}$, to let $g(K)$ in Eq. (D15) be a real number.

APPENDIX E: PARTICLE DYNAMICS IN THE STOCHASTIC DELAY DIFFERENTIAL EQUATION

If a delay time $t_{\rm d}$ is added to Eq. (B1), the latter becomes

$$\gamma\dot{x}(t) = -kx(t - t_{\rm d}) + \xi(t). \tag{E1}$$

Suppose that the noise here has the same strength $\xi(t) = A\Gamma(t)$ as that in Eq. (D1), where $A = \sqrt{2\gamma k_{\rm B}T} = \sqrt{2\gamma^2 D}$ is as defined below Eq. (D11). Then $\xi(t)dt = A\Gamma(t)dt = AdW(t)$, with $W(t)$ a Wiener process, where $\langle W(t)\rangle = 0$ and $\langle W^2(t)\rangle = t$. Adopting these notations, Eq. (E1) becomes

$$\dot{x}(t) = -\kappa x(t - t_{\rm d}) + \zeta(t), \tag{E2}$$

where $\kappa = k/\gamma$, $\zeta = \xi/\gamma$, and $\zeta(t)dt = \sigma dW(t)$, with $\sigma = \sqrt{2k_{\rm B}T/\gamma} = \sqrt{2D}$. For general $\kappa \geqslant 0$ and $\pi/(2\kappa) > t_{\rm d} \geqslant 0$, the variance of the steady state of x in Eq. (E2) is [25]

$$\langle x^2\rangle = \frac{\sigma^2}{2\kappa}\left[\frac{1 + \sin(\kappa t_{\rm d})}{\cos(\kappa t_{\rm d})}\right] = \frac{2k_{\rm B}T/\gamma}{2k/\gamma}\left[\frac{1 + \sin(kt_{\rm d}/\gamma)}{\cos(kt_{\rm d}/\gamma)}\right]$$
$$= \frac{k_{\rm B}T}{k}\frac{1 + \sin(K)}{\cos(K)}, \tag{E3}$$

where $K \equiv kt_{\rm d}/\gamma = \kappa t_{\rm d} < \pi/2$. $\langle x^2\rangle$ varies with $t_{\rm d}$ and has its minimum at $t_{\rm d} = 0$. It becomes larger when $t_{\rm d} > 0$ because the restoring force of the potential well to prevent the particle from moving away from the potential minimum at $x = 0$ becomes less instantaneous and less efficient. Thus, the $\langle x^2\rangle$ associated $\hat{T}$ at $t_{\rm d} > 0$ is also larger than that at $t_{\rm d} = 0$. From Eq. (E3) it follows that

$$f(k) = \frac{k\langle x^2\rangle}{k_{\rm B}\bar{T}} = \frac{1 + \sin(K)}{\cos(K)} \equiv \hat{f}(K), \tag{E4}$$

which is an increasing function within $K \in [0, \frac{\pi}{2})$, where $\bar{T} = T$ below Eq. (D1) has been used.

TABLE I. Examples of Fourier transforms.

$\underbrace{h(t)}_{\text{dim}=\mathrm{T}^{-1}}$	$\frac{\delta(t)}{1_\mathrm{T}}$	$\frac{e^{\frac{-\|t\|}{\tau}}}{2\tau}$	$\frac{\mathrm{sinc}(at)}{1_\mathrm{T}} \equiv \frac{\sin(\pi at)}{1_\mathrm{T}\pi at}$	$\begin{bmatrix}\frac{\delta(t)}{1_\mathrm{T}}\\ 0\end{bmatrix}$	$\begin{bmatrix}\frac{e^{\frac{-\|t\|}{\tau}}}{\tau}\\ 0\end{bmatrix}$
$\underbrace{F(h(t))}_{\text{dim. less}}$	1	$\frac{1}{1+\omega^2\tau^2}$	$\frac{1}{1_\mathrm{T}\|a\|}\,\mathrm{rect}(\frac{\omega}{2\pi a})$	1	$\frac{1}{1+i\omega\tau}$

For $\gamma = 10^{-8}$ kg/s and $k \in [0,\ 4] \times 10^{-5}$ N/m, $t_\mathrm{d} = 5 \times 10^{-5}$ and 2.5×10^{-4} s give $K = kt_\mathrm{d}/\gamma \in [0,\ 0.2]$ and $[0,\ 1]$, respectively. If we extend the range of k for the first t_d to cover the range $K \in [0,\ 1]$, the same as the second t_d, they show an identical $\hat{f}(K)$, as depicted in Fig. 4(c).

Rewriting Eq. (E2) as $\dot{x}(t) = -\kappa x(t - t_\mathrm{d}) + \sqrt{2D}\Gamma(t)$ and expanding it to the first order in a small t_d, it yields [20]

$$\begin{aligned}\dot{x}(t) &= -\frac{\kappa x(t)}{1-\kappa t_\mathrm{d}} + \frac{\sqrt{2D}\Gamma(t)}{1-\kappa t_\mathrm{d}} = -\frac{kx(t)}{\gamma - kt_\mathrm{d}} + \frac{\xi(t)}{\gamma - kt_\mathrm{d}}\\ &= \frac{1}{\gamma}\left(-\frac{kx(t)}{1-K} + \frac{\xi(t)}{1-K}\right),\end{aligned} \tag{E5}$$

with the variance

$$\begin{aligned}\langle x^2\rangle &= \frac{D}{\kappa}\frac{1}{1-\kappa t_\mathrm{d}} = \frac{(k_\mathrm{B}T/\gamma)}{(k/\gamma)}\frac{1}{1-(k/\gamma)t_\mathrm{d}}\\ &= \frac{k_\mathrm{B}T}{k}\frac{1}{(1-K)},\end{aligned} \tag{E6}$$

which increases with t_d as in Eq. (E3). In analogy to $\bar{T} = T$ in Eq. (E4), the result in Eq. (E6) leads to the increasing function

$$f(k) = \frac{k\langle x^2\rangle}{k_\mathrm{B}\bar{T}} = \frac{1}{1-K} \equiv \hat{f}(K). \tag{E7}$$

APPENDIX F: GENERALIZED LANGEVIN EQUATION I

Let $F(h(t)) \equiv \tilde{h}(\omega) = \int_{-\infty}^{\infty} h(t)e^{-i\omega t}dt$ be the Fourier transform of the function $h(t)$ and $h(t) = (2\pi)^{-1}\int_{-\infty}^{\infty}\tilde{h}(\omega)e^{i\omega t}d\omega$ be its inverse transform. Some useful examples are collected in Table I. Therein $[^x_0]$ stands for $\{^x_0\ ^{\text{if } t \geqslant 0}_{\text{if } t<0}$ and all t, $1/a$, and the number of one 1_T have the dimension of time. Here, 1_T is introduced to make all terms in the equations consistent in dimension for a dimensional analysis. For $h(t) = e^{-|t|/\tau}/(2\tau)$ (see Table within the Supplemental Material [49]),

$$\begin{aligned}F(h(t)) &= \int_{-\infty}^{\infty}\frac{1}{2\tau}e^{-|t|/\tau}e^{-i\omega t}dt\\ &= \frac{1}{2\tau}\frac{2/\tau}{(1/\tau)^2+\omega^2} = \frac{1}{1+\omega^2\tau^2}.\end{aligned}$$

At $\tau \to 0$, $e^{-|t|/\tau}/(2\tau)$ in the integral converges to a delta function $\delta(t)/1_\mathrm{T}$ and $F(e^{-|t|/\tau}/(2\tau)) = (1+\omega^2\tau^2)^{-1} \to 1$. For the normalized sinc function $h(t) = \mathrm{sinc}(at)/1_\mathrm{T}$ in Table I, $F(h(t)) = \mathrm{rect}(\omega/(2\pi a))/(1_\mathrm{T}|a|)$ is an even function, whose support (domain of nonzero values) extending from $\omega = -\pi a$ to πa has a constant value $1/|a|$. For $h(t) = [^{\delta(t)/1_\mathrm{T}}_{0}]$, $F(h(t))$ should be understood as

$$\begin{aligned}&\int_{-\infty}^{\infty} h(t)e^{-i\omega t}dt\\ &\quad = \int_{-\infty}^{0^-} 0 \times e^{-i\omega t}dt + \int_{0^-}^{\infty}\frac{\delta(t)}{1_\mathrm{T}}e^{-i\omega t}dt = 1,\end{aligned}$$

with a number 0^- slightly smaller than 0 to let $\int_{0^-}^{\infty}$ fully cover the "support" of the $\delta(t)$ peak. For $h(t) = [^{e^{-|t|/\tau}/\tau}_{0}]$ (see Table within the Supplemental Material [49]),

$$\begin{aligned}F(h(t)) &= \frac{1}{\tau}\int_0^{\infty} e^{\frac{-t}{\tau}}e^{-i\omega t}dt = \frac{1}{\tau}\int_0^{\infty} e^{-(\frac{1}{\tau}+i\omega)t}dt\\ &= -\frac{1}{\tau}\left(\frac{1}{\tau}+i\omega\right)^{-1} e^{-(\frac{1}{\tau}+i\omega)t}\Bigg|_0^{\infty} = \frac{1}{1+i\omega\tau}.\end{aligned}$$

Next, consider the generalized Langevin equation

$$\gamma\dot{x}(t) + k\int_{-\infty}^{t} K_\mathrm{M}(t-t')x(t')dt' = \xi(t), \tag{F1}$$

with γ and k the same as those in Eq. (B1). If, due to causality, $K_\mathrm{M}(t-t')$ is nonzero only for $t' \leqslant t$, Eq. (F1) is equivalent to $\gamma\dot{x}(t) + k\int_{-\infty}^{\infty} K_\mathrm{M}(t-t')x(t')\,dt' = \xi(t)$. Its Fourier transform becomes

$$i\omega\gamma\tilde{x}(\omega) + k\tilde{K}_\mathrm{M}(\omega)\tilde{x}(\omega) = \tilde{\xi}(\omega), \tag{F2}$$

which implies

$$\begin{aligned}\tilde{x}(\omega) &= \frac{\tilde{\xi}(\omega)}{i\omega\gamma + k\tilde{K}_\mathrm{M}(\omega)}\\ \langle\tilde{x}(\omega)\tilde{x}(\omega')\rangle &= \frac{\langle\tilde{\xi}(\omega)\tilde{\xi}(\omega')\rangle}{[i\omega\gamma + k\tilde{K}_\mathrm{M}(\omega)][i\omega'\gamma + k\tilde{K}_\mathrm{M}(\omega')]}.\end{aligned} \tag{F3}$$

If $\xi(t)$ is a stationary stochastic process, then $\langle\xi(t+t')\xi(t')\rangle = \langle\xi(t)\xi(0)\rangle$ and thus

$$\begin{aligned}\langle\tilde{\xi}(\omega)\tilde{\xi}(\omega')\rangle &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\langle\xi(t)\xi(t')\rangle e^{-i(\omega t+\omega' t')}dtdt'\\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\langle\xi(t+t')\xi(t')\rangle e^{-i(\omega t+\omega t'+\omega' t')}dtdt'\\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\langle\xi(t+t')\xi(t')\rangle e^{-i\omega t}dt e^{-i(\omega+\omega')t'}dt'\\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\langle\xi(t)\xi(0)\rangle e^{-i\omega t}dt e^{-i(\omega+\omega')t'}dt'\\ &= \int_{-\infty}^{\infty} F(\langle\xi(t)\xi(0)\rangle)e^{-i(\omega+\omega')t'}dt'\\ &= S_\xi(\omega)\int_{-\infty}^{\infty} e^{-i(\omega+\omega')t'}dt'\\ &= 2\pi\delta(\omega+\omega')1_\mathrm{T}S_\xi(\omega).\end{aligned} \tag{F4}$$

Here, $F(\langle\xi(t)\xi(0)\rangle)$ is a function of ω, rather than of t or t'. It is the power spectrum $S_\xi(\omega)$ of the noise $\xi(t)$, which leads to $\langle\xi^2\rangle = (2\pi)^{-1}\int_{-\infty}^{\infty} S_\xi(\omega)d\omega$. Using Eqs. (F3) and (F4), we obtain

$$\begin{aligned}\langle x^2\rangle &= \frac{1}{4\pi^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\langle\tilde{x}(\omega)\tilde{x}(\omega')\rangle e^{i(\omega+\omega')t}d\omega d\omega'\\ &= \frac{1}{4\pi^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{\langle\tilde{\xi}(\omega)\tilde{\xi}(\omega')\rangle e^{i(\omega+\omega')t}d\omega d\omega'}{[i\omega\gamma + k\tilde{K}_\mathrm{M}(\omega)][i\omega'\gamma + k\tilde{K}_\mathrm{M}(\omega')]}\\ &= \frac{1}{4\pi^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{2\pi\delta(\omega+\omega')1_\mathrm{T}S_\xi(\omega)e^{i(\omega+\omega')t}d\omega d\omega'}{[i\omega\gamma + k\tilde{K}_\mathrm{M}(\omega)][i\omega'\gamma + k\tilde{K}_\mathrm{M}(\omega')]}\\ &= \frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{S_\xi(\omega)1_\mathrm{T}}{[i\omega\gamma + k\tilde{K}_\mathrm{M}(\omega)][-i\omega\gamma + k\tilde{K}_\mathrm{M}(-\omega)]1_\mathrm{T}}d\omega\\ &= \frac{1}{2\pi}\int_{-\infty}^{\infty}S_x(\omega)d\omega,\end{aligned} \tag{F5}$$

with the power spectrum of x,

$$S_x(\omega) \equiv \frac{S_\xi(\omega)}{\omega^2\gamma^2 + k^2\tilde{K}_\mathrm{M}(\omega)\tilde{K}_\mathrm{M}(-\omega) + i\omega\gamma k\Delta_\mathrm{M}}, \tag{F6}$$

where $\Delta_\mathrm{M} \equiv \tilde{K}_\mathrm{M}(-\omega) - \tilde{K}_\mathrm{M}(\omega)$. The dimensions of some terms in the above calculations are as follows: $K_\mathrm{M}(t-t')$ (T^{-1}) in Eq. (F1), $\tilde{x}(\omega)$ (LT), $\tilde{K}_\mathrm{M}(\omega)$ (dimensionless), and $\tilde{\xi}(\omega)$ (MLT^{-1}) in Eq. (F2), $i\omega\gamma + k\tilde{K}_\mathrm{M}(\omega)$ (MT^{-2}) in Eq. (F3), $\langle\tilde{\xi}(\omega)\tilde{\xi}(\omega')\rangle$ ($\mathrm{M^2L^2T^{-2}}$) and $S_\xi(\omega)$ ($\mathrm{M^2L^2T^{-3}}$) in Eq. (F4), and $S_x(\omega)$ ($\mathrm{L^2T}$) in Eq. (F6).

Let $K_\mathrm{M}(t) = [^{e^{-|t|/\tau_\mathrm{p}}/\tau_\mathrm{p}}_{0}]$ and $\langle\xi(t)\xi(t')\rangle = q\delta(t-t')/1_\mathrm{T}$ with $q = 2\gamma k_\mathrm{B}\bar{T}$, where $\bar{T}$ is defined as before Eq. (A1). Then, according to Table I, $\tilde{K}_\mathrm{M}(\omega) = (1+i\omega\tau_\mathrm{p})^{-1}$ and $S_\xi(\omega) = F(\langle\xi(t)\xi(0)\rangle) = q$. Therefore,

$$\begin{aligned}\Delta_\mathrm{M} &= \tilde{K}_\mathrm{M}(-\omega) - \tilde{K}_\mathrm{M}(\omega)\\ &= \frac{1}{1-i\omega\tau_\mathrm{p}} - \frac{1}{1+i\omega\tau_\mathrm{p}} = \frac{2i\omega\tau_\mathrm{p}}{1+\omega^2{\tau_\mathrm{p}}^2}\end{aligned}$$

and subsequently the denominator of Eq. (F6) is

$$\begin{aligned}&\omega^2\gamma^2 + k^2\tilde{K}_\mathrm{M}(\omega)\tilde{K}_\mathrm{M}(-\omega) + i\omega\gamma k\Delta_\mathrm{M}\\ &= \omega^2\gamma^2 + \frac{k^2}{1+\omega^2{\tau_\mathrm{p}}^2} + i\omega\gamma k\left(\frac{2i\omega\tau_\mathrm{p}}{1+\omega^2{\tau_\mathrm{p}}^2}\right)\\ &= \omega^2\gamma^2 + \frac{k^2 - 2\gamma k\tau_\mathrm{p}\omega^2}{1+\omega^2{\tau_\mathrm{p}}^2}\\ &= \frac{\omega^4\gamma^2{\tau_\mathrm{p}}^2 + (\gamma^2 - 2\gamma k\tau_\mathrm{p})\omega^2 + k^2}{1+\omega^2{\tau_\mathrm{p}}^2}\\ &= \frac{\gamma^2}{{\tau_\mathrm{p}}^2}\frac{\omega^4{\tau_\mathrm{p}}^4 + (1-2k\tau_\mathrm{p}/\gamma)\omega^2{\tau_\mathrm{p}}^2 + k^2{\tau_\mathrm{p}}^2/\gamma^2}{\omega^2{\tau_\mathrm{p}}^2+1}\\ &= \left[\frac{\tau_\mathrm{p}}{\gamma^2}\frac{w^2+1}{w^4+(1-2K)w^2+K^2}\tau_\mathrm{p}\right]^{-1},\end{aligned} \tag{F7}$$

with $w \equiv \omega\tau_\mathrm{p}$ and $K \equiv k\tau_\mathrm{p}/\gamma$. As a result (see F8 within the Supplemental Material [49]),

$$\begin{aligned}\langle x^2\rangle &= \frac{1}{2\pi}\int_{-\infty}^{\infty}S_x(\omega)d\omega\\ &= \frac{q}{2\pi}\frac{\tau_\mathrm{p}}{\gamma^2}\int_{-\infty}^{\infty}\frac{w^2+1}{w^4+(1-2K)w^2+K^2}dw\\ &= \frac{q\tau_\mathrm{p}}{2\pi\gamma^2}\frac{\pi(K+1)}{K} = \frac{k_\mathrm{B}\bar{T}}{k}(K+1),\end{aligned} \tag{F8}$$

which yields a simple rising function

$$f(k) = \frac{k\langle x^2\rangle}{k_\mathrm{B}\bar{T}} = K+1 \equiv \hat{f}(K). \tag{F9}$$

A comparison shows that Eq. (E1) is a special case of Eq. (F1) when $K_\mathrm{M}(t) = \delta(t-t_\mathrm{d})$.

APPENDIX G: GENERALIZED LANGEVIN EQUATION II

Another well-known generalized Langevin equation is

$$\gamma\int_{-\infty}^{t}K_\mathrm{D}(t-t')\dot{x}(t')dt' + kx(t) = \xi(t). \tag{G1}$$

It can be written as $\gamma\int_{-\infty}^{\infty}dt'K_\mathrm{D}(t-t')\dot{x}(t') + kx(t) = \xi(t)$ if, due to causality, $K_\mathrm{D}(t-t') = 0$ for $t' > t$. Its Fourier transform is

$$i\omega\gamma\tilde{x}(\omega)\tilde{K}_\mathrm{D}(\omega) + k\tilde{x}(\omega) = \tilde{\xi}(\omega), \tag{G2}$$

which yields

$$\begin{aligned}\tilde{x}(\omega) &= \frac{\tilde{\xi}(\omega)}{k + i\omega\gamma\tilde{K}_\mathrm{D}(\omega)}\\ \langle\tilde{x}(\omega)\tilde{x}(\omega')\rangle &= \frac{\langle\tilde{\xi}(\omega)\tilde{\xi}(\omega')\rangle}{[k+i\omega\gamma\tilde{K}_\mathrm{D}(\omega)][k+i\omega'\gamma\tilde{K}_\mathrm{D}(\omega')]}.\end{aligned} \tag{G3}$$

The dimension of $K_\mathrm{D}(t-t')$ in Eq. (G1) is T^{-1} and $\tilde{K}_\mathrm{D}(\omega)$ in Eq. (G2) is dimensionless. In analogy to Eq. (F5),

$$\begin{aligned}\langle x^2\rangle &= \left(\frac{1}{2\pi}\right)^2\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\langle\tilde{x}(\omega)\tilde{x}(\omega')\rangle e^{i(\omega+\omega')t}d\omega d\omega'\\ &= \frac{1}{2\pi}\int_{-\infty}^{\infty}S_x(\omega)d\omega,\end{aligned} \tag{G4}$$

with the power spectrum

$$S_x(\omega) \equiv \frac{S_\xi(\omega)}{k^2 + \omega^2\gamma^2\tilde{K}_\mathrm{D}(\omega)\tilde{K}_\mathrm{D}(-\omega) + i\omega\gamma k\Delta_\mathrm{D}}, \tag{G5}$$

where $\Delta_\mathrm{D} \equiv \tilde{K}_\mathrm{D}(\omega) - \tilde{K}_\mathrm{D}(-\omega)$. For $K_\mathrm{D}(t) = [^{e^{-|t|/\tau_\mathrm{v}}/\tau_\mathrm{v}}_{0}]$, we have $\tilde{K}_\mathrm{D}(\omega) = F(K_\mathrm{D}(t)) = (1+i\omega\tau_\mathrm{v})^{-1}$ known from Table I and thus

$$\begin{aligned}\Delta_\mathrm{D} &= \tilde{K}_\mathrm{D}(\omega) - \tilde{K}_\mathrm{D}(-\omega)\\ &= \frac{1}{1+i\omega\tau_\mathrm{v}} - \frac{1}{1-i\omega\tau_\mathrm{v}} = \frac{-2i\omega\tau_\mathrm{v}}{1+\omega^2{\tau_\mathrm{v}}^2}.\end{aligned}$$

It yields

$$\begin{aligned} &k^2+\gamma^2\omega^2\tilde{K}_D(\omega)\tilde{K}_D(-\omega)+i\omega\gamma k\Delta_D\\ &=k^2+\frac{\gamma^2\omega^2}{1+\omega^2\tau_v{}^2}+i\omega\gamma k\left(\frac{-2i\omega\tau_v}{1+\omega^2\tau_v{}^2}\right)\\ &=k^2+\frac{\gamma^2}{\tau_v{}^2}\frac{\omega^2\tau_v{}^2}{1+\omega^2\tau_v{}^2}+\frac{2\gamma k}{\tau_v}\frac{\omega^2\tau_v{}^2}{1+\omega^2\tau_v{}^2}\\ &=k^2+\left(\frac{\gamma^2}{\tau_v{}^2}+\frac{2\gamma k}{\tau_v}\right)\frac{\omega^2\tau_v{}^2}{1+\omega^2\tau_v{}^2}\\ &\sim \text{const of}\omega \quad \text{at } \omega\to\infty. \end{aligned} \tag{G6}$$

Therefore, $S_\xi(\omega)$ in Eq. (G5) must decay sufficiently fast with ω to make $S_x(\omega)$ in Eq. (G4) integrable.

Case 1. For $\langle\xi(t)\xi(t')\rangle=qe^{-|t-t'|/\tau}/(2\tau)$ with $q=2\gamma k_B\bar{T}$ as that below Eq. (F6), one has $S_\xi(\omega)=F(\langle\xi(t)\xi(0)\rangle)=q(1+\omega^2\tau^2)^{-1}$ from Table I. Together with Eq. (G6), Eq. (G5) becomes

$$\begin{aligned} S_x(\omega)&=\frac{\left(\frac{q}{1+\omega^2\tau^2}\right)}{k^2+\left(\frac{\gamma^2}{\tau_v{}^2}+\frac{2\gamma k}{\tau_v}\right)\frac{\omega^2\tau_v{}^2}{1+\omega^2\tau_v{}^2}}\\ &=\frac{q\left(1+\omega^2\tau_v{}^2\right)}{(1+\omega^2\tau^2)\left[k^2(1+\omega^2\tau_v{}^2)+\left(\frac{\gamma^2}{\tau_v{}^2}+\frac{2\gamma k}{\tau_v}\right)\omega^2\tau_v{}^2\right]}. \end{aligned} \tag{G7}$$

Since

$$\begin{aligned} [\cdots]&=k^2+\left(k^2\tau_v{}^2+\gamma^2+2\gamma k\tau_v\right)\omega^2\\ &=k^2+(k\tau_v+\gamma)^2\omega^2=(k\tau_v+\gamma)^2\left[\omega^2+\frac{k^2}{(k\tau_v+\gamma)^2}\right], \end{aligned}$$

Eq. (G7) has the form

$$\begin{aligned} S_x(\omega)&=\frac{q\tau_v{}^2\left(\omega^2+\frac{1}{\tau_v{}^2}\right)}{\tau^2(k\tau_v+\gamma)^2\left(\omega^2+\frac{1}{\tau^2}\right)\left[\omega^2+\frac{k^2}{(k\tau_v+\gamma)^2}\right]}\\ &=\frac{D(\omega^2+A^2)}{(\omega^2+B^2)(\omega^2+C^2)}, \end{aligned} \tag{G8}$$

with $A\equiv\tau_v{}^{-1}$, $B\equiv\tau^{-1}$, $C\equiv k/(k\tau_v+\gamma)=K/[(K+1)\tau_v]$, where $K\equiv k\tau_v/\gamma$, and

$$\begin{aligned} D&\equiv\frac{q\tau_v{}^2}{\tau^2(k\tau_v+\gamma)^2}=\frac{2\gamma k_B\bar{T}k\tau_v{}^2/\gamma^2}{k\tau^2(k\tau_v/\gamma+1)^2}\\ &=\frac{k_B\bar{T}}{k\tau}\frac{\tau_v}{\tau}\frac{2k\tau_v/\gamma}{(K+1)^2}=\frac{k_B\bar{T}}{k\tau}\left(\frac{c}{1-c}\right)\frac{2K}{(K+1)^2}. \end{aligned} \tag{G9}$$

Here, $c\equiv\tau_v/(\tau_v+\tau)$, such that $0\leqslant c\leqslant 1$ and $\tau_v/\tau=c/(1-c)$. With Eq. (G8), we obtain

$$\begin{aligned} \langle x^2\rangle&=\frac{1}{2\pi}\int_{-\infty}^{\infty}S_x(\omega)d\omega\\ &=\frac{D}{2\pi}\int_{-\infty}^{\infty}\frac{(\omega^2+A^2)}{(\omega^2+B^2)(\omega^2+C^2)}d\omega\\ &=\frac{D}{2\pi}\frac{\pi(B^2C-A^2C+A^2B-C^2B)}{(B^2-C^2)BC}\\ &=\frac{D}{2}\left[\frac{\frac{A^2}{BC}+1}{C+B}\right], \end{aligned} \tag{G10}$$

where

$$\begin{aligned} [\cdots]&=\frac{\frac{1}{\tau_v{}^2}\tau\frac{(K+1)\tau_v}{K}+1}{\frac{K}{(K+1)\tau_v}+\frac{1}{\tau}}=\frac{[(K+1)\tau+K\tau_v](K+1)\tau_v\tau}{[K\tau+(K+1)\tau_v]K\tau_v}\\ &=\frac{\left(K+\frac{\tau}{\tau+\tau_v}\right)(K+1)\tau}{\left(K+\frac{\tau_v}{\tau+\tau_v}\right)K}=\frac{(K+1-c)(K+1)\tau}{(K+c)K}. \end{aligned} \tag{G11}$$

Inserting Eqs. (G9) and (G11) into Eq. (G10) gives

$$\begin{aligned} \langle x^2\rangle&=\frac{1}{2}\frac{k_B\bar{T}}{k\tau}\frac{c}{1-c}\frac{2K}{(K+1)^2}\frac{(K+1-c)(K+1)\tau}{(K+c)K}\\ &=\frac{k_B\bar{T}}{k}\frac{c}{1-c}\frac{(K+1-c)}{(K+1)(K+c)}, \end{aligned} \tag{G12}$$

which has been confirmed by another calculation (see G12 within the Supplemental Material [49]). Equation (G12) implies

$$f(k)=\frac{k\langle x^2\rangle}{k_B\bar{T}}=\frac{c(K+1-c)}{(1-c)(K+1)(K+c)}=\hat{f}(K). \tag{G13}$$

At $\tau_v\to 0$,

$$\begin{aligned} f(k)&=\frac{\tau_v}{\tau}\frac{\left(\frac{k\tau_v}{\gamma}+\frac{\tau}{\tau_v+\tau}\right)}{\left(\frac{k\tau_v}{\gamma}+1\right)\left(\frac{k\tau_v}{\gamma}+\frac{\tau_v}{\tau_v+\tau}\right)}\\ &=\frac{\left(\frac{k\tau_v}{\gamma}+\frac{\tau}{\tau_v+\tau}\right)}{\left(\frac{k\tau_v}{\gamma}+1\right)\left(\frac{k\tau}{\gamma}+\frac{\tau}{\tau_v+\tau}\right)}\to\frac{1}{\frac{k\tau}{\gamma}+1} \end{aligned} \tag{G14}$$

reduces to Eq. (B10).

Whether $\hat{f}(K)$ in Eq. (G13) increases with K depends on the sign of

$$\begin{aligned} \frac{\partial\hat{f}(K)}{\partial K}&=\frac{c}{(1-c)}\frac{-K^2+(2c-2)K+\left(c^2+c-1\right)}{(K+1)^2(K+c)^2}\\ &\equiv L_1(K)L_2(K), \end{aligned} \tag{G15}$$

or merely on the sign of $L_2(K)\equiv -K^2+(2c-2)K+(c^2+c-1)$, as $L_1(K)\equiv c[(1-c)(K+1)^2(K+c)^2]^{-1}$ is always positive. At $K=0$, due to $0\leqslant c\leqslant 1$, we have

$$\left.\frac{\partial\hat{f}(K)}{\partial K}\right|_{K=0}=\frac{c^2+c-1}{(1-c)c}\begin{cases}\geqslant 0 & \text{if } c\geqslant c_0,\\ <0 & \text{if } c<c_0,\end{cases} \tag{G16}$$

where $c_0\equiv(-1+\sqrt{5})/2$. On the other hand, at $K\gg 1$,

$$\left.\frac{\partial\hat{f}(K)}{\partial K}\right|_{K\gg 1}\approx-\frac{c}{(1-c)K^2}<0. \tag{G17}$$

That means, for $c>c_0$, $\hat{f}(K)$ will first rise and then fall. The maximum of $\hat{f}(K)$, if exists, will be a solution of $L_2(K)=0$, which is located at

$$\begin{aligned} K&=\frac{-(2c-2)\pm\sqrt{(2c-2)^2-4(-1)(c^2+c-1)}}{2(-1)}\\ &=c-1\mp\sqrt{c(2c-1)}. \end{aligned} \tag{G18}$$

Here, (K) must be a positive number, which requires $c(2c-1)>0$ and $K=c-1+\sqrt{c(2c-1)}>0$. The former gives rise to $c>1/2$, while the latter implies $\sqrt{c(2c-1)}>$

$1-c$ and thus $2c^2-c>1-2c+c^2$. It implies $c^2+c-1>0$ and subsequently $c>c_0$, in consistent with that in Eq. (G16). Thus, a positive slope of $\hat{f}(K)$ can exist in the regime $K\in[0,c-1+\sqrt{c(2c-1)}]$ or equivalently $k\in(\gamma/\tau_v)\times[0,c-1+\sqrt{c(2c-1)}]$ when $c>c_0\approx0.618$, or equivalently $\tau_v/\tau=c/(1-c)>c_0/(1-c_0)=[(-1+\sqrt{5})/2]/[1-(-1+\sqrt{5})/2]=(-1+\sqrt{5})/(3-\sqrt{5})\approx1.618$. That is, the memory in the friction term will enhance the efficiency, while that in the noise term will suppress the engine efficiency. The correlation time of the former τ_v must be larger than that of the latter τ to some extent to see the outperformance of an active engine over its passive counterpart.

For $\gamma=10^{-8}$ kg/s, $\tau_v=4\times10^{-4}$ s, and $\tau=10^{-4}$ s, we have $c=\tau_v/(\tau_v+\tau)=0.8>c_0\approx0.618$. Hence, $\hat{f}(K)$ will be a rising function when $k\in(\gamma/\tau_v)\times[0,(c-1+\sqrt{c(2c-1)})]=(10^{-8}/(4\times10^{-4}))\times[0,\ (0.8-1+\sqrt{0.8(1.6-1)})]\approx[0,\ 1.23\times10^{-5}]$ N/m, which is an experimentally accessible range of k and corresponds to $K=k\tau_v/\gamma\in[0,\ 0.49]$. For all $c=0.90$, 0.85, and 0.70, the $\hat{f}(K)$ in Eq. (G13) has a rising branch, as the blue, violet, and red lines, respectively, shown in Fig. 4(e). The rising trend for $c=0.70$ is weaker because it has been very close to the threshold 0.618.

Case 2. For $\langle\xi(t)\xi(t')\rangle=q\,a\,\mathrm{sinc}(a(t-t'))$ with $q=2\gamma k_B\bar{T}$ and $a>0$, one has $S_\xi(\omega)=F(\langle\xi(t)\xi(0)\rangle)=q\,\mathrm{rect}(\omega/(2\pi a))$, as in Table I. Recall that the dimension of $a\,\mathrm{sinc}(at)$ is T^{-1} and $\mathrm{rect}(\omega/(2\pi a))$ is dimensionless. Together with Eqs. (G4)~(G6), one obtains

$$\begin{aligned}\langle x^2\rangle&=\frac{1}{2\pi}\int_{-\infty}^{\infty}S_x(\omega)d\omega\\&=\frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{q\,\mathrm{rect}\left(\frac{\omega}{2\pi a}\right)}{k^2+\left(\frac{\gamma^2}{\tau_v^2}+\frac{2k\gamma}{\tau_v}\right)\frac{\omega^2\tau_v^2}{1+\omega^2\tau_v^2}}d\omega\\&=\int_{-\pi a}^{\pi a}L(\omega)d\omega,\end{aligned}\tag{G19}$$

with

$$\begin{aligned}L(\omega)&\equiv\frac{q}{2\pi}\frac{\left(1+\omega^2\tau_v^2\right)}{\left[k^2(1+\omega^2\tau_v^2)+\left(\frac{\gamma^2}{\tau_v^2}+\frac{2\gamma k}{\tau_v}\right)\omega^2\tau_v^2\right]}\\&=\frac{q\tau_v^2\left(\omega^2+\frac{1}{\tau_v^2}\right)}{2\pi(k\tau_v+\gamma)^2\left[\omega^2+\frac{k^2}{(k\tau_v+\gamma)^2}\right]}\\&=\frac{D'(\omega^2+A^2)}{(\omega^2+C^2)},\end{aligned}\tag{G20}$$

where the equality before Eq. (G8) has been used. Here, $A\equiv\tau_v^{-1}$ and

$$C\equiv\frac{k}{k\tau_v+\gamma}=\frac{1}{\tau_v}\frac{\frac{k\tau_v}{\gamma}}{\frac{k\tau_v}{\gamma}+1}=\frac{1}{\tau_v}\frac{K_v}{K_v+1}=\frac{1}{\tau_v}K,$$

where $K_v\equiv k\tau_v/\gamma$ and $K\equiv(1+K_v^{-1})^{-1}=k\tau_v/(k\tau_v+\gamma)$, such that $K_v=(-1+K^{-1})^{-1}$ and $K^2/K_v=K(1-K)$. Note that both K_v and K increase with k. With that

$$\begin{aligned}D'&\equiv\frac{q\tau_v^2}{2\pi(k\tau_v+\gamma)^2}=\frac{2\gamma k_B\bar{T}}{2\pi k^2}\frac{(k\tau_v)^2}{(k\tau_v+\gamma)^2}\\&=\frac{k_B\bar{T}}{k}\frac{\tau_v}{\pi}\frac{\gamma}{k\tau_v}\left(\frac{k\tau_v}{k\tau_v+\gamma}\right)^2=\frac{k_B\bar{T}}{k}\frac{\tau_v}{\pi}\frac{1}{K_v}K^2\\&=\frac{k_B\bar{T}}{k}\frac{\tau_v}{\pi}K(1-K).\end{aligned}\tag{G21}$$

Using $\int_b^d(\omega^2+A^2)(\omega^2+C^2)^{-1}d\omega=(d-b)+(A^2/C-C)[\tan^{-1}(d/C)-\tan^{-1}(b/C)]$ to calculate Eq. (G19) and considering the above values of A, C, and D', it yields

$$\begin{aligned}\langle x^2\rangle&=D'\int_{-\pi a}^{\pi a}\frac{(\omega^2+A^2)}{(\omega^2+C^2)}d\omega\\&=D'\left\{2\pi a+\left(\frac{A^2}{C}-C\right)\left[2\tan^{-1}\left(\frac{\pi a}{C}\right)\right]\right\}\\&=\frac{k_B\bar{T}}{k}\frac{\tau_v}{\pi}K(1-K)\\&\quad\times\left[2\pi a+2\left(\frac{1}{K\tau_v}-\frac{K}{\tau_v}\right)\tan^{-1}\left(\frac{\pi a\tau_v}{K}\right)\right]\\&=\frac{k_B\bar{T}}{k}\frac{2}{\pi}K(1-K)\left[\pi a\tau_v+\left(\frac{1}{K}-K\right)\tan^{-1}\left(\frac{\pi a\tau_v}{K}\right)\right]\\&=\frac{k_B\bar{T}}{k}\frac{2(1-K)}{\pi}\left[cK+(1-K^2)\tan^{-1}\left(\frac{c}{K}\right)\right],\end{aligned}\tag{G22}$$

where $c\equiv\pi a\tau_v$. Subsequently, $\hat{f}(K)=k\langle x^2\rangle/(k_B\bar{T})$ implies

$$\hat{f}(K)=\frac{2(1-K)}{\pi}\left[cK+(1-K^2)\tan^{-1}\left(\frac{c}{K}\right)\right],\tag{G23}$$

which has been confirmed by another calculation (see G23 within the Supplemental Material [49]). At $k\propto K\to0^+$, a number slightly larger than 0, one obtains $\hat{f}(K)\to(2/\pi)[\tan^{-1}(c/0^+)]=1$.

The slope of $\hat{f}(K)$ is (see G23 within the Supplemental Material [49])

$$\frac{d\hat{f}(K)}{dK}=\frac{2Q(K)}{\pi(K^2+c^2)},\tag{G24}$$

with

$$\begin{aligned}Q(K)\equiv&-3cK^3+2cK^2+(c-2c^3)K+c^3-c\\&+[3K^4-2K^3+(3c^2-1)K^2-2c^2K-c^2]\\&\times\tan^{-1}(c/K).\end{aligned}\tag{G25}$$

At $K\approx0^+$, $\tan^{-1}(c/K)\approx\pi/2$ and all other terms containing K are close to zero. Hence, $Q(K)\approx c^3-c+(-c^2)(\pi/2)=(c/2)(2c^2-\pi c-2)$, which is positive when $c>c_0\equiv(\pi+\sqrt{\pi^2+16})/4\approx2.057$. However, when c/K decreases from ∞ to 2π, the value of $\tan^{-1}(c/K)$ has already reduced to 90% of $\tan^{-1}(\infty)\approx\pi/2$. For a larger K, such that $c/K<2\pi$ or equivalently $K>c/(2\pi)$, it is better to solve the range of c for a rising $Q(K)$ numerically, instead of using the above estimation $\tan^{-1}(c/K)\approx\pi/2$.

For $\gamma=10^{-8}$ kg/s, $\tau_v=2.5\times10^{-4}$ s, and $k\in[0,\ 4]\times10^{-5}$ N/m, it yields $K=k\tau_v/(k\tau_v+\gamma)\in[0,1/2]$. If $a=10^4$ s^{-1}, one has $c=\pi a\tau_v=2.5\pi$. The corresponding

$c/K \in [5\pi, \infty]$ is large enough to consider the estimation $\tan^{-1}(c/K) \approx \pi/2$ for $Q(K)$ in Eq. (G24). This $Q(K)$ should be positive because $c = 2.5\pi > c_0 \equiv 2.057$. For $a = 1.2 \times 10^4$, 8×10^3, and 4×10^3 s^{-1}, one obtains $c = 3\pi, 2\pi$, and π, respectively, for which the rising trends of $\hat{f}(K)$ in Eq. (G23) are as shown in the blue, violet, and red lines in Fig. 4(f).

[1] J. Casas-Vázquez and D. Jou, Temperature in non-equilibrium states: A review of open problems and current proposals, Rep. Prog. Phys. **66**, 1937 (2003).

[2] L. F. Cugliandolo, The effective temperature, J. Phys. A: Math. Theor. **44**, 483001 (2011).

[3] A. Puglisi, A. Sarracino, and A. Vulpiani, Temperature in and out of equilibrium: A review of concepts, tools and attempts, Phys. Rep. **709-710**, 1 (2017).

[4] S. Krishnamurthy, S. Ghosh, D. Chatterji, R. Ganapathy, and A. K. Sood, A micrometre-sized heat engine operating between bacterial reservoirs, Nat. Phys. **12**, 1134 (2016).

[5] V. Blickle and C. Bechinger, Realization of a micrometre-sized stochastic heat engine, Nat. Phys. **8**, 143 (2012).

[6] I. A. Martínez, É. Roldán, L. Dinis, D. Petrov, J. M. R. Parrondo, and R. A. Rica, Brownian Carnot engine, Nat. Phys. **12**, 67 (2016).

[7] J. A. C. Albay, Z.-Y. Zhou, C.-H. Chang, H.-Y. Chen, J. S. Lee, C.-H. Chang, and Y. Jun, Colloidal heat engine driven by engineered active noise, Phys. Rev. Res. **5**, 033208 (2023).

[8] R. Zakine, A. Solon, T. Gingrich, and F. van Wijland, Stochastic Stirling engine operating in contact with active baths, Entropy **19**, 193 (2017).

[9] V. Holubec, S. Steffenoni, G. Falasco, and K. Kroy, Active Brownian heat engines, Phys. Rev. Res. **2**, 043262 (2020).

[10] J. S. Lee and H. Park, Effects of the non-Markovianity and non-Gaussianity of active environmental noises on engine performance, Phys. Rev. E **105**, 024130 (2022).

[11] X. L. Wu and A. Libchaber, Particle diffusion in a quasi-two-dimensional bacterial bath, Phys. Rev. Lett. **84**, 3017 (2000).

[12] C. Maggi, M. Paoluzzi, N. Pellicciotta, A. Lepore, L. Angelani, and R. Di Leonardo, Generalized energy equipartition in harmonic oscillators driven by active baths, Phys. Rev. Lett. **113**, 238303 (2014).

[13] H. H. Rugh, Dynamical approach to temperature, Phys. Rev. Lett. **78**, 772 (1997).

[14] J. Tailleur and M. E. Cates, Sedimentation, trapping, and rectification of dilute bacteria, Europhys. Lett. **86**, 60002 (2009).

[15] K. Sekimoto, *Stochastic Energetics, Lecture Notes in Physics 799* (Springer, Berlin, 2010).

[16] K. Kanazawa, T. G. Sano, A. Cairoli, and A. Baule, Loopy Lévy flights enhance tracer diffusion in active suspensions, Nature (London) **579**, 364 (2020).

[17] J. A. C. Albay, G. Paneru, H. K. Pak, and Y. Jun, Optical tweezers as a mathematically driven spatio-temporal potential generator, Opt. Express **26**, 29906 (2018).

[18] Y. Jun and J. Bechhoefer, Virtual potentials for feedback traps, Phys. Rev. E **86**, 061106 (2012).

[19] J. A. C. Albay, Z. Y. Zhou, C.-H. Chang, and Y. Jun, Shift a laser beam back and forth to exchange heat and work in thermodynamics, Sci. Rep. **11**, 4394 (2021).

[20] D. Geiss, K. Kroy, and V. Holubec, Corrigendum: Brownian molecules formed by delayed harmonic interactions (2019 New J. Phys. 21 093014), New J. Phys. **21**, 119503 (2019).

[21] W. Guo, C.-J. Wang, L.-C. Du, and D.-C. Mei, Effects of time delay on transport processes in an active Brownian particle, Physica A **392**, 4210 (2013).

[22] U. Khadka, V. Holubec, H. Yang, and F. Cichos, Active particles bound by information flows, Nat. Commun. **9**, 3864 (2018).

[23] F. A. Rihan, *Delay Differential Equations and Applications to Biology* (Springer, Singapore, 2022).

[24] R. C. Kumaran, T. G. Venkatesh, and K. S. Swarup, Stochastic delay differential equations: Analysis and simulation studies, Chaos, Solitons Fractals **165**, 112819 (2022).

[25] S. Guillouzic, I. L'Heureux, and A. Longtin, Small delay approximation of stochastic delay differential equations, Phys. Rev. E **59**, 3970 (1999).

[26] U. Küchler and B. Mensch, Langevins stochastic differential equation extended by a time-delayed term, Stochastics Stochastic Rep. **40**, 23 (1992).

[27] T. Frank and P. Beek, Stationary solutions of linear stochastic delay differential equations: Applications to biological systems, Phys. Rev. E **64**, 021917 (2001).

[28] A. M. Wazwaz, Volterra integro-differential equations, in *Linear and Nonlinear Integral Equations* (Springer, Berlin, 2011).

[29] H. Mori, Transport, collective motion, and Brownian motion, Prog. Theor. Phys. **33**, 423 (1965).

[30] R. Kubo, The fluctuation-dissipation theorem, Rep. Prog. Phys. **29**, 255 (1966).

[31] D. Loi, S. Mossa, and L. F. Cugliandolo, Non-conservative forces and effective temperatures in active polymers, Soft Matter **7**, 10193 (2011).

[32] S. Henkes, Y. Fily, and M. C. Marchetti, Active jamming: Self-propelled soft particles at high density, Phys. Rev. E **84**, 040301(R) (2011).

[33] J. R. Gomez-Solano, Work extraction and performance of colloidal heat engines in viscoelastic baths, Front. Phys. **9**, 643333 (2021).

[34] C. A. Guevara-Valadez, R. Marathe, and J. R. Gomez-Solano, A Brownian cyclic engine operating in a viscoelastic active suspension, Physica A **609**, 128342 (2023).

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

[36] A. V. Dass, T. Georgelin, F. Westall, F. Foucher, P. De Los Rios, D. M. Busiello, S. Liang, and F. Piazza, Equilibrium and non-equilibrium furanose selection in the ribose isomerisation network, Nat. Commun. **12**, 2749 (2021).

[37] D. M. Busiello, S. Liang, F. Piazza, and P. De Los Rios, Dissipation-driven selection of states in non-equilibrium chemical networks, Commun. Chem. **4**, 16 (2021).

[38] R. Filliger and P. Reimann, Brownian gyrator: A minimal heat engine on the nanoscale Phys. Rev. Lett. **99**, 230602 (2007).

[39] M. C. Marchetti, J. F. Joanny, S. Ramaswamy, T. B. Liverpool, J. Prost, M. Rao, and R. Aditi Simha, Hydrodynamics of soft active matter, Rev. Mod. Phys. **85**, 1143 (2013).

[40] J. Mabillard, C. A. Weber, and F. Jülicher, Heat fluctuations in chemically active systems, Phys. Rev. E **107**, 014118 (2023).

[41] P. Song, H. Gao. Z. Gao, J. Liu, R. Zhang, B. Kang, J.-J. Xu, and H.-Y. Chen, Heat transfer and thermoregulation within single cells revealed by transient plasmonic imaging, Chem **7**, 1569 (2021).

[42] S. Sotoma, C. Zhong, J. C. Y. Kah, H. Yamashita, T. Plakhotnik, Y. Harada, and M. Suziki, *In situ* measurements of intracellular thermal conductivity using heater-thermometer hybrid diamond nanosensors, Sci. Adv. **7**, eabd7888 (2021).

[43] D. Chrétien, P. Bénit, H.-H. Ha, S. Keipert, R. El-Khoury, Y.-T. Chang, M. Jastroch, H. T. Jacobs, P. Rustin, and M. Rak, Mitochondria are physiologically maintained at close to 50 $^{\circ}$C, PLoS Biol. **16**, e2003992 (2018).

[44] A. W. Fritsch, A. F. Diaz-Delgadillo, O. Adame-Arana, C. Hoege, M. Mittasch, M. Kreysing, M. Leaver, A. A. Hyman, F. Jülicher, and C. A. Weber, Local thermodynamics govern formation and dissolution of *Caenorhabditis elegans* P granule condensates, Proc. Natl. Acad. Sci. USA **118**, e2102772 (2021).

[45] J. A. Smorodinsky and P. Ziescher, *Was ist temperatur? Begriff, Geschichte, Labro und Kosmos* (Harri Deuitsch, Frankfurt am Main, 2000).

[46] H. Chang, *Inventing Temperature* (Oxford University Press, Oxford, 2004).

[47] C. Bechinger, R. Di Leonardo, H. Löwen, C. Reichhardt, G. Volpe, and G. Volpe, Active Brownian particles in complex and crowded environments, Rev. Mod. Phys. **88**, 045006 (2016).

[48] I. A. Martínez, É. Roldán, L. Dinis, and R. A. Rica, Colloidal heat engines: A review, Soft Matter **13**, 22 (2017).

[49] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevResearch.5.043085 for calculations and verifications of some equations by using Maple codes.

[50] M. Esposito, R. Kawai, K. Lindenberg, and C. Van den Broeck, Efficiency at maximum power of low-dissipation Carnot engines, Phys. Rev. Lett. **105**, 150603 (2010).

[51] C. J. Pierce, E. Osborne, E. Mumper, B. H. Lower, S. K. Lower, and R. Sooryakumar, Thrust and power output of the bacterial flagellar motor: A micromagnetic tweezers approach, Biophys. J. **117**, 1250 (2019).