Unattainability of Carnot efficiency in thermal motors: Coarse graining and entropy production of Feynman-Smoluchowski ratchets

Yohei Nakayama,[1,*] Kyogo Kawaguchi,[2,3,†] and Naoko Nakagawa[4,‡]
[1]*Department of Physics, Chuo University, Tokyo 112-8551, Japan*
[2]*Department of Systems Biology, Harvard Medical School, Boston, Massachusetts 02115, USA*
[3]*Universal Biology Institute, The University of Tokyo, Tokyo 113-0033, Japan*
[4]*Department of Physics, Ibaraki University, Mito 310-8512, Japan*

(Received 8 April 2018; published 3 August 2018)

We revisit and analyze the thermodynamic efficiency of the Feynman-Smoluchowski (FS) ratchet, a classical thought experiment describing an autonomous heat-work converter. Starting from the full kinetics of the FS ratchet and deriving the exact forms of the hidden dissipations resulting from coarse graining, we restate the historical controversy over its thermodynamic efficiency. The existence of hidden entropy productions implies that the standard framework of stochastic thermodynamics applied to the coarse-grained descriptions fails in capturing the dissipative feature of the system. In response to this problem, we explore an extended framework of stochastic thermodynamics to reconstruct the hidden entropy production from the coarse-grained dynamics. The approach serves as a key example of how we can systematically address the problem of thermodynamic efficiency in a multivariable fluctuating nonequilibrium system.

DOI: 10.1103/PhysRevE.98.022102

I. INTRODUCTION

The framework of stochastic thermodynamics has not only allowed experimental characterization of small thermodynamic systems [1], but has also established a unified scheme to address fundamental questions in thermodynamics. Identities and inequalities formulated for general stochastic dynamics have been given thermodynamical interpretations such as the second law [2–4], role of information feedback [1,5], bound on efficiencies of engines at finite time operations [6,7], and laws extended to nonequilibrium setups [8–10].

The crucial concept behind the developments in stochastic thermodynamics is the entropy production, which is typically introduced through local detailed balance using the logarithmic ratio of transition probabilities [11]. This quantity is equivalent, at least in several models, to the energy exchanged with the heat bath divided by the temperature of the bath [12], and satisfies the second lawlike inequality. Recent works, however, have clarified that fluctuating nonequilibrium systems can carry hidden entropy productions [13–24], and even under the properly controlled limit of coarse graining the coarse-grained model may not preserve the thermodynamic properties of the original system [14,16,20,25].

In this paper, we focus on the analysis of the Feynman-Smoluchowski (FS) ratchet [Fig. 1(a)] [26] as a model case to understand how the thermodynamic efficiency can seemingly change according to the different coarse-grained descriptions of the dynamics. The FS ratchet is one of the most celebrated thought experiments in thermodynamics, where there has been a controversy over its thermodynamic efficiency. The FS ratchet, due to its asymmetric design, may appear as if it can convert the thermal fluctuation of a single heat bath into work by unidirectional rotation, violating the second law of thermodynamics. After Smoluchowski showed that there is no rotation and extracted work if the FS ratchet is placed in an isothermal environment [27], in Ref. [26], Feynman considered whether it is possible for the ratchet to operate as a Carnot efficient engine. It was claimed, based on the analysis of a simplified discrete-stepping model, that the ratchet may attain Carnot efficiency at the stalled state between two heat baths with different temperatures.

Parrondo and Español, however, critisized Feynman's argument and pointed out that the momentum variable is non-negligible for the dissipation in the FS ratchet by using a different method of simplifying the model [28]. In addition, the unattainability of Carnot efficiency was established in another autonomous Brownian heat engine, the Büttiker-Landauer (BL) motor [29–31], which is a model that has been thought to be closely related to the FS ratchet. These studies have thus formed a consensus that the FS ratchet cannot attain Carnot efficiency [32–34].

Thermodynamic efficiencies in coarse-grained models are not only a theoretical concern but also important in the interpretation of experimental data, since the measurements are typically restricted to a small set of slow variables. The potential existence of hidden entropy productions (i.e., dissipation owing to the unobserved variables) will make it virtually impossible to draw any conclusions about thermodynamic efficiency in a nonequilibrium small system experiment. Therefore, it is of interest to extend the framework of stochastic thermodynamics to be able to reinterpret the coarse-grained data in order to estimate the original thermodynamic properties.

*nakayama@phys.chuo-u.ac.jp
†Kyogo_Kawaguchi@hms.harvard.edu
‡naoko.nakagawa.phys@vc.ibaraki.ac.jp

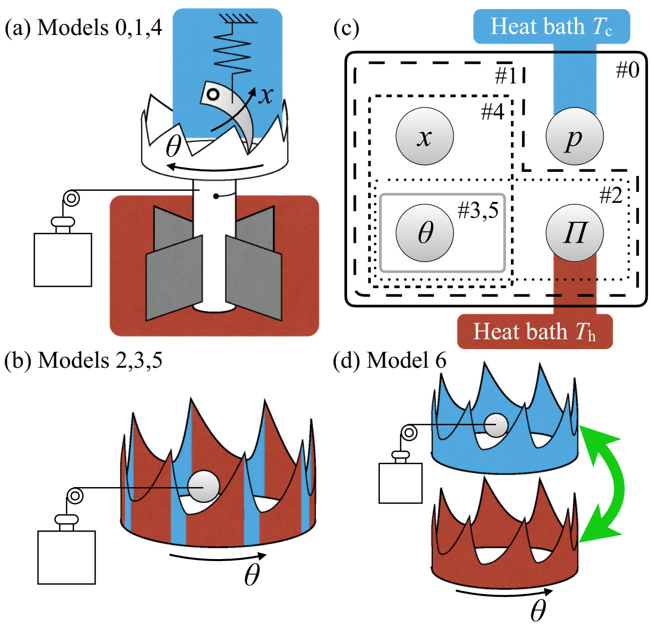

FIG. 1. Schematics of the FS ratchet and its coarse-grained descriptions. (a) In models 0, 1, and 4, the FS ratchet consists of a vane, a gear, and a pawl. A spring presses the pawl against the gear, and an external load applies torque to the axle. The vane and pawl are attached to different heat baths. (b) In models 2, 3, and 5, Langevin equations with the effective mechanical potential $U_{\rm eff}(\theta)$, inhomogeneous friction $\mathcal{G}(\theta)$, and temperature $T_{\rm eff}(\theta)$ describe the dynamics of the FS ratchet. (c) The scheme of coarse graining and the degrees of freedom in models 0–5. (d) Langevin model with stochastic switching between the two heat baths (model 6).

Here we aim to provide a unified understanding and a workaround to the FS ratchet problem, through a systematic procedure of coarse graining without any empirical simplification. We derive the coarse-grained descriptions of the original FS ratchet including the previously known models [35–39] together with additional models. We then ask how the entropy productions may differ in the series of models by obtaining the explicit expressions for the hidden entropy productions, and discuss their relations to the previous arguments on the controversial thermodynamic efficiencies. Finding that most of the coarse-grained dynamics do not preserve the thermodynamic property of the original FS ratchet, we further explore and find a way to quantify the hidden dissipation based on a limited number of coarse-grained observables.

This paper is organized as follows. In Sec. II we introduce the original FS ratchet model (model 0). In Sec. III the coarse-grained descriptions (models 2, 3, and 5) are derived by taking the time-scale separation limits. In Sec. IV we calculate the behavior of the dissipation through the framework of stochastic thermodynamics, in the limits where the coarse-grained descriptions are obtained. We derive the explicit forms of hidden entropy productions as the first main result of the paper. In addition, we clarify how Feynman's argument is equivalent to the stochastic thermodynamics of the coarse-grained description. In Sec. V we present the results of numerical simulations which clarify the impact of hidden entropy production on the thermodynamic efficiencies (Fig. 6). These results confirm that although the kinetics of the FS ratchet can be coarse grained systematically, most of the coarse-grained models do not reproduce the entropy production of the original system. In Sec. VI we describe our proposal of a workaround to the problem by demonstrating that even when using the coarse-grained variables, the fine-grained entropy production can be reconstructed by the decomposition of the Langevin dynamics (model 6). In Sec. VII we give concluding remarks. Some technical details are described in the appendixes.

II. SETUP

As shown in Fig. 1(a), the FS ratchet consists of a vane and a gear connected by a rigid axle, and a pawl meshing with the gear. A spring pushes the pawl against the gear. The vane and the pawl are in contact with different heat baths with temperatures T_h and T_c. An external load couples with the axle, and applies a constant torque f. By assuming that the interaction between the pawl and the gear is mechanical, the equations of motion for the angle θ of the coaxial vane and gear and the height x of the pawl reads

$$\begin{aligned}\dot{\theta} &= \frac{\Pi}{m},\\ \dot{\Pi} &= -\frac{\Gamma}{m}\Pi + f - \frac{\partial U(\theta,x)}{\partial\theta} + \sqrt{2\Gamma T_h}\xi,\\ \dot{x} &= \frac{p}{m_x},\\ \dot{p} &= -\frac{\gamma}{m_x}p - \frac{\partial U(\theta,x)}{\partial x} + \sqrt{2\gamma T_c}\zeta,\end{aligned} \tag{1}$$

which we refer to as model 0, where Π and p are the momentum conjugated to θ and x, respectively. Here, m is the corresponding moment of inertia, and m_x is the mass of the pawl. We take Langevin heat baths where Γ and γ are the viscous frictional coefficients. ξ and ζ are independent white Gaussian noises with zero means and unit variances. The Boltzmann constant is set to unity.

In a straightforward manner, we may obtain a coarse-grained description (model 1),

$$\begin{aligned}\dot{\theta} &= \frac{\Pi}{m},\\ \dot{\Pi} &= -\frac{\Gamma}{m}\Pi + f - \frac{\partial U(\theta,x)}{\partial\theta} + \sqrt{2\Gamma T_h}\xi,\\ \gamma\dot{x} &= -\frac{\partial U(\theta,x)}{\partial x} + \sqrt{2\gamma T_c}\tilde{\zeta},\end{aligned} \tag{2}$$

where the momentum degree of freedom p is eliminated by considering the overdamped limit for the pawl. The symbol $\tilde{\zeta}$ is an independent white Gaussian noise with zero mean and unit variance.

We assume the mechanical potential

$$U(\theta,x) = U_0(x) + U_I[x-\phi(\theta)], \tag{3}$$

where $U_0(x)$ is the elastic potential of the spring attached to the pawl and $U_I[x-\phi(\theta)]$ is the trapping potential between the tip of the pawl and the surface of the gear. $\phi(\theta)$ is a periodic function with the period L, which represents the shape of the gear.

III. COARSE GRAINING

We here explicitly derive the coarse-grained descriptions of model 1 by taking the limits where the time scales of the variables are separated. This is in contrast to the approaches taken for example in [26] where the discrete stepping model and the BL motor were introduced on the basis of phenomenological arguments. We start from model 1 and consider two limits, a "tightly confined limit" and an "overdamped limit," where x and Π are eliminated, respectively. Through this two-step coarse graining, we arrive at a closed equation of motion for θ. We note that the order of elimination of x and Π matters. As summarized in Fig. 2, we find that taking the tightly confined limit first will result in a different model from when the overdamped limit is taken first (model 3 vs model 5).

Hereafter, we denote the time scale of the relaxation in the trapping potential as $\tau_x := L_x^2\gamma/T_c$ with the length scale L_x of the trapping potential. The time scale for the relaxation of the momentum of the vane and the gear is $\tau_\Pi = m/\Gamma$. We assume the other time scales to be of the same order, represented by τ. In order to satisfy this assumption, we fix the functional forms of $\phi(\theta)$ and $U_0(x)$ and the ratios Γ/γ, T_c/T_h, and fL/T_h. We are interested in the cases where τ_x and τ_Π are separated from τ.

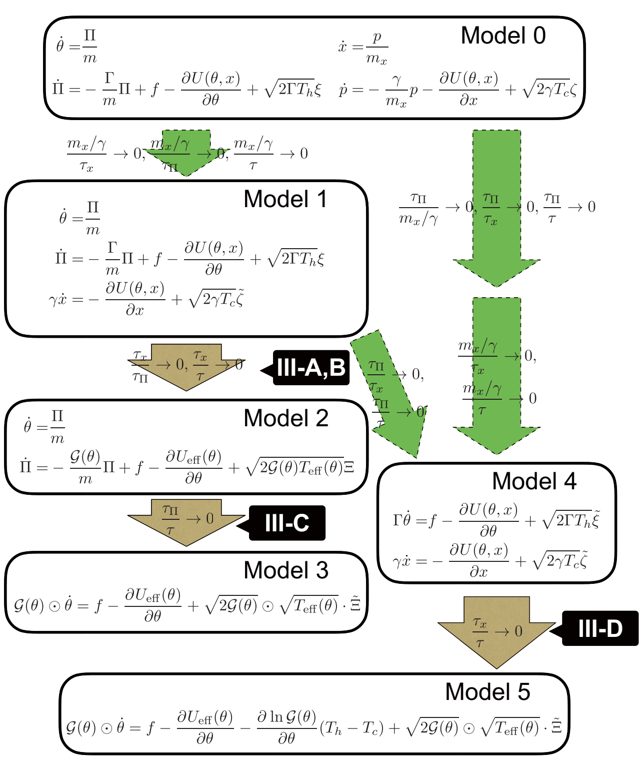

FIG. 2. A complete picture of the routes of coarse graining. The arrows represent the processes of coarse graining with specific limits. The arrows with solid outline (yellow) represent the coarse-graining processes derived in this paper, and the arrows with dashed outline (green) correspond to the trivial coarse graining eliminating the momentum variable attached to an isothermal environment. $\tau_x = L_x^2\gamma/T_c$ and $\tau_\Pi = m/\Gamma$ are the time scales of the pawl and momentum degree of freedom, respectively, and τ represents the other time scales of the system.

Here we note why we chose the interaction between the tip of the pawl and the surface of the gear $U_I[x-\phi(\theta)]$ as a trapping potential. In the original FS ratchet, this interaction was a hard-core repulsion, so the tight confinement of the pawl could only be realized by reinforcing the potential of the spring $[U_0(x)]$, since the length scale of the confinement is proportional to $T_h/[\partial U(\theta,x)/\partial x]$. However, the energy required to lift the pawl becomes larger than the thermal energy when increasing the spring force, which means that the FS ratchet will stop rotating in this limit. By introducing $U_I[x-\phi(\theta)]$ as a trapping potential, we may take the tightly confined limit by keeping the height of the potential barrier constant. This modification to the original dynamics is the key to the following coarse-graining procedures.

We mainly conduct coarse graining in terms of the master equation, i.e., partial differential equation satisfied by probability densities. The master equation for model 1 reads

$$\begin{aligned}\frac{\partial P(\theta,\Pi,x)}{\partial t} &= -\frac{\partial}{\partial\theta}\left[\frac{\Pi}{m}P(\theta,\Pi,x)\right] - \frac{\partial}{\partial\Pi}\left[\left(-\frac{\Gamma}{m}\Pi + f - \frac{\partial U(\theta,x)}{\partial\theta}\right)\right.\\ &\quad\left.\times P(\theta,\Pi,x) - \Gamma T_h\frac{\partial P(\theta,\Pi,x)}{\partial\Pi}\right]\\ &\quad - \frac{\partial}{\partial x}\left(-\frac{1}{\gamma}\frac{\partial U(\theta,x)}{\partial x}P(\theta,\Pi,x) - \frac{T_c}{\gamma}\frac{\partial P(\theta,\Pi,x)}{\partial x}\right). \end{aligned} \tag{4}$$

Here, $P(\theta,\Pi,x)$ represents the joint probability density of θ, Π, and x.

A. Coarse-grained description at tightly confined limit

We first consider the limit where the tip of the pawl is tightly confined to the surface of the gear. In this limit, τ_x is assumed to be separated from τ_Π and τ. Assuming that the ratio τ_Π/τ is fixed, we introduce a small parameter representing the separation of time scales, $\varepsilon := \tau_x/\tau_\Pi \sim \tau_x/\tau$. Here, we summarize the derivation of the coarse-grained description, and give the details in Appendix A 1.

In this tightly confined limit, the height of the pawl x is eliminated as the fast variable. Therefore, the coarse-grained dynamics is described by the master equation for the joint probability density, $P(\theta,\Pi) = \int dx\, P(\theta,\Pi,x)$. Throughout this paper, the integrals with respect to θ, Π, and x are taken over the domain of integrand. The time derivative of $P(\theta,\Pi)$ is obtained by integrating Eq. (4) with respect to x:

$$\begin{aligned}\frac{\partial P(\theta,\Pi)}{\partial t} &= -\frac{\partial}{\partial\theta}\left[\frac{\Pi}{m}P(\theta,\Pi)\right]\\ &\quad - \frac{\partial}{\partial\Pi}\left[\left(-\frac{\Gamma}{m}\Pi + f\right)P(\theta,\Pi) - \Gamma T_h\frac{\partial P(\theta,\Pi)}{\partial\Pi}\right]\\ &\quad - \frac{\partial}{\partial\Pi}\left[-\int dx\,\frac{\partial U(\theta,x)}{\partial\theta}P(\theta,\Pi,x)\right]. \end{aligned} \tag{5}$$

The closed equation for $P(\theta,\Pi)$ is obtained by evaluating the last line of Eq. (5) in the limit of $\varepsilon \to 0$. Employing the singular perturbation theory to avoid the divergence caused by the secular terms, we obtain

$$\frac{\partial P(\theta,\Pi)}{\partial t} = -\frac{\partial}{\partial\theta}\left[\frac{\Pi}{m}P(\theta,\Pi)\right] - \frac{\partial}{\partial\Pi}\left[-\frac{\mathcal{G}(\theta)}{m}\Pi - \frac{\partial U_{\rm eff}(\theta)}{\partial\theta} + f - \mathcal{G}(\theta)T_{\rm eff}(\theta)\frac{\partial}{\partial\Pi}\right]P(\theta,\Pi), \quad (6)$$

which is the Kramers equation. Here, $U_{\rm eff}(\theta)$, $\mathcal{G}(\theta)$, and $T_{\rm eff}(\theta)$ are the effective potential, the effective frictional coefficient, and the effective temperature:

$$U_{\rm eff}(\theta) = U_0[\phi(\theta)], \quad (7)$$

$$\mathcal{G}(\theta) = \Gamma + \gamma\phi'(\theta)^2, \quad (8)$$

$$T_{\rm eff}(\theta) = \frac{\Gamma T_h + \gamma\phi'(\theta)^2 T_c}{\Gamma + \gamma\phi'(\theta)^2}. \quad (9)$$

The Kramers equation (6) is equivalent to the Langevin equation (model 2)

$$\dot\theta = \frac{\Pi}{m},$$
$$\dot\Pi = -\frac{\mathcal{G}(\theta)}{m}\Pi + f - \frac{\partial U_{\rm eff}(\theta)}{\partial\theta} + \sqrt{2\mathcal{G}(\theta)T_{\rm eff}(\theta)}\,\Xi, \quad (10)$$

where Ξ is a white Gaussian noise with zero mean and unit variance. Model 2 describes the Brownian motion of a single degree of freedom under the effective potential, frictional coefficient, and temperature [Fig. 1(b)], which is known as the Büttiker-Landauer motor [35].

B. Quick derivation of model 2

We here give a quick derivation of model 2 in the special case where we set $U(\theta,x) = kx^2/2 + \lambda[x-\phi(\theta)]^2/2$, based on a temporal coarse-graining method [40,41]. The details will be given in Appendix A 2. Since the equation of motion,

$$\gamma\dot x = -(k+\lambda)x + \lambda\phi(\theta) + \sqrt{2\gamma T_c}\,\tilde\zeta, \quad (11)$$

is linear in x, we may formally solve x as the functional of $\phi(\theta)$ and $\tilde\zeta$, and eliminate x by substituting the formal solution of x into the equation of motion of Π:

$$\dot\Pi_t = -\frac{\Gamma}{m}\Pi_t + f + \lambda\phi'(\theta_t)\times\left[-\frac{k}{k+\lambda}\phi(\theta_t) - \frac{\gamma\lambda}{(k+\lambda)^2}\phi'(\theta_t)\frac{\Pi_t}{m}\right] + O\left(\frac{\varepsilon}{\tau_\Pi}\right) + \frac{\lambda\phi'(\theta_t)\sqrt{2\gamma T_c}}{\gamma}\int_{-\infty}^{t} dt' e^{-(k+\lambda)/\gamma(t-t')}\tilde\zeta_{t'} + \sqrt{2\Gamma T_h}\xi_t, \quad (12)$$

where we explicitly show the time dependence of the variables as the subscript. Note that Eq. (12) describes a non-Markovian dynamics with colored noise. The Markovian property is recovered in the limit of $\varepsilon = \tau_x/\tau_\Pi \to 0$. In this limit, we may introduce a time interval Δt, which is shorter than τ_Π and τ, but longer than $\tau_x = \gamma/(k+\lambda)$. By integrating Eq. (12) over the time interval Δt and taking the limit of $\varepsilon \to 0$, we obtain model 2.

C. Coarse-grained description at overdamped limit

Next, we discuss the overdamped limit. In this limit $\tau_\Pi = m/\Gamma$ is separated from the other time scales represented by τ. The underdamped Brownian motion corresponding to model 2 can be coarse grained to an overdamped Langevin equation (model 3) [42,43]:

$$\mathcal{G}(\theta)\odot\dot\theta = f - \frac{\partial U_{\rm eff}(\theta)}{\partial\theta} + \sqrt{2\mathcal{G}(\theta)}\odot\sqrt{T_{\rm eff}(\theta)}\cdot\tilde\Xi, \quad (13)$$

where $\tilde\Xi$ is a white Gaussian noise with zero mean and unit variance. The symbols $\cdot$ and $\odot$ indicate the product in the sense of Itô and anti-Itô, respectively, which specify the evaluation of the quantity on the left:

$$\sqrt{2\mathcal{G}(\theta)}\odot\sqrt{T_{\rm eff}(\theta)}\cdot\tilde\Xi = \lim_{\delta t\to 0}\sqrt{2\mathcal{G}(\theta_{t+\delta t})T_{\rm eff}(\theta_t)}\frac{1}{\delta t}\int_t^{t+\delta t}\tilde\Xi_s ds. \quad (14)$$

Model 3 is a generalized version of the Büttiker-Landauer motor to the case of θ-dependent friction. Solving Eq. (13), the net velocity of rotation of this motor is obtained as [35]

$$\langle\dot\theta\rangle = \frac{L[1-\exp(-L\Delta)]}{\int_0^L dy\exp[-\psi(y)]\int_y^{y+L}dy'\exp[\psi(y')]\mathcal{G}(y')/T_{\rm eff}(y')}, \quad (15)$$

where $\langle\cdot\rangle$ represents the steady-state ensemble average and

$$\psi(y) := \int^y dy'\frac{f - U'_{\rm eff}(y') - T'_{\rm eff}(y')}{T_{\rm eff}(y')}, \quad (16)$$

$$\Delta := \psi(y) - \psi(y+L). \quad (17)$$

Equation (15) indicates that there is unidirectional motion if $\Delta \neq 0$. The system works as a Brownian heat engine when the rotation is opposite to the direction of the constant torque: $\langle\dot\theta\rangle < 0$ and $\Delta < 0$ when $f > 0$. In Appendix A 3, we present the derivation of model 3 based on the singular perturbation theory [16].

D. Other routes of coarse graining

As shown in Fig. 2, there is another path to obtain the closed equation of motion for θ; we can take the overdamped limit before the tightly confined limit. In a straightforward manner, we can coarse grain model 1 and obtain model 4:

$$\Gamma\dot\theta = f - \frac{\partial U(\theta,x)}{\partial\theta} + \sqrt{2\Gamma T_h}\,\tilde\xi,$$
$$\gamma\dot x = -\frac{\partial U(\theta,x)}{\partial x} + \sqrt{2\gamma T_c}\,\tilde\zeta, \quad (18)$$

which is equivalent to the master equation:

$$\frac{\partial P(\theta,x)}{\partial t} = -\frac{\partial}{\partial\theta}\left[\frac{1}{\Gamma}\left(f - \frac{\partial U(\theta,x)}{\partial\theta}\right)P(\theta,x) - \frac{T_h}{\Gamma}\frac{\partial P(\theta,x)}{\partial\theta}\right] - \frac{\partial}{\partial x}\left[-\frac{1}{\gamma}\frac{\partial U(\theta,x)}{\partial x}P(\theta,x) - \frac{T_c}{\gamma}\frac{\partial P(\theta,x)}{\partial x}\right]. \quad (19)$$

Here, $\tilde{\xi}$ is an independent white Gaussian noise with zero mean and unit variance. Model 4 may also be obtained from model 0 by taking the overdamped limit for the vane and the gear first (Fig. 2). This type of model has already been analyzed in the context of the FS ratchet [37–39].

The time evolution of $P(\theta) = \int dx\, P(\theta,x)$ is obtained by taking the tightly confined limit in Eq. (19). In the form of Langevin equation, the coarse-grained description (model 5) is obtained as

$$\mathcal{G}(\theta) \odot \dot{\theta} = f - \frac{\partial U_{\rm eff}(\theta)}{\partial \theta} - \frac{\partial \ln \mathcal{G}(\theta)}{\partial \theta}(T_h - T_c) + \sqrt{2\mathcal{G}(\theta)} \odot \sqrt{T_{\rm eff}(\theta)} \cdot \hat{\Xi}, \tag{20}$$

where $\hat{\Xi}$ is a white Gaussian noise with zero mean and unit variance. Unidirectional motion is driven by the same mechanism as model 3. The details of the derivation are given in Appendix A 4.

Models 3 and 5 are similar but slightly different; there is an extra term in model 5 that vanishes when $T_h = T_c$ but affects the average velocity when $T_h \neq T_c$. Formally, this means that the two limits, $\tau_x \ll \tau_\Pi \ll \tau$ (model 3) and $\tau_\Pi \ll \tau_x \ll \tau$ (model 5), are different under a nonequilibrium setup.

We here remark on some of the previous works related to the coarse graining of models like the FS ratchet. In [44], the authors coarse grained model 4 based on phenomenlogical arguments, and correctly obtained the effective potential and temperature [Eqs. (7) and (9)]. However, they did not arrive at the inhomogeneous friction [Eq. (8)] and the force proportional to temperature difference [the third term of the right-hand side of Eq. (20)]. In [45], a similar attempt was made to obtain a phenomenological model by neglecting the temporal correlation of the fast variable, which resulted in an unphysical model that does not satisfy the fluctuation-dissipation relation. In [36], Millonas considered a nonequilibrium bath variable (x in our model) that couples to a motor, and essentially obtained Eqs. (7)–(9). The aspects of dissipation and thermodynamic efficiency, however, were not discussed.

IV. STOCHASTIC THERMODYNAMICS OF FEYNMAN-SMOLUCHOWSKI RATCHET

We here discuss the thermodynamic properties of the FS ratchet by applying the framework of stochastic thermodynamics. We first describe the entropy production for each model, and then take the coarse-graining limits in each case to see if there is a descrepancy (i.e., hidden entropy production) between the different scales of descriptions.

A. Entropy production rates

The standard prescription of stochastic thermodynamics [11] connects the entropy production rate in the heat baths with the transition probabilities of the model. The entropy production rates for the models are then

$$\begin{aligned}\sigma_1(\theta_{t'},\Pi_{t'},x_{t'}|\theta_t,\Pi_t,x_t) &:= \frac{1}{t'-t}\ln\frac{W_1(\theta_{t'},\Pi_{t'},x_{t'}|\theta_t,\Pi_t,x_t)}{W_1(\theta_t,-\Pi_t,x_t|\theta_{t'},-\Pi_{t'},x_{t'})}\\ &= \frac{Q_1^h}{T_h} + \frac{Q^c}{T_c},\end{aligned} \tag{21}$$

$$\begin{aligned}\sigma_2(\theta_{t'},\Pi_{t'}|\theta_t,\Pi_t) &:= \frac{1}{t'-t}\ln\frac{W_2(\theta_{t'},\Pi_{t'}|\theta_t,\Pi_t)}{W_2(\theta_t,-\Pi_t|\theta_{t'},-\Pi_{t'})}\\ &= \frac{1}{T_{\rm eff}(\theta)} \circ Q_2,\end{aligned} \tag{22}$$

$$\begin{aligned}\sigma_3(\theta_{t'}|\theta_t) &:= \frac{1}{t'-t}\ln\frac{W_3(\theta_{t'}|\theta_t)}{W_3(\theta_t|\theta_{t'})}\\ &= \frac{1}{T_{\rm eff}(\theta)} \circ \left(Q_3 - \frac{\partial T_{\rm eff}(\theta)}{\partial \theta} \circ \dot{\theta}\right),\end{aligned} \tag{23}$$

$$\begin{aligned}\sigma_4(\theta_{t'},x_{t'}|\theta_t,x_t) &:= \frac{1}{t'-t}\ln\frac{W_4(\theta_{t'},x_{t'}|\theta_t,x_t)}{W_4(\theta_t,x_t|\theta_{t'},x_{t'})}\\ &= \frac{Q_4^h}{T_h} + \frac{Q^c}{T_c},\end{aligned} \tag{24}$$

$$\begin{aligned}&\sigma_5(\theta_{t'}|\theta_t)\\ &:= \frac{1}{t'-t}\ln\frac{W_5(\theta_{t'}|\theta_t)}{W_5(\theta_t|\theta_{t'})}\\ &= \frac{1}{T_{\rm eff}(\theta)} \circ \left[Q_5 - \left(\frac{\partial T_{\rm eff}(\theta)}{\partial \theta} + \frac{\partial \mathcal{G}(\theta)}{\partial \theta}(T_h - T_c)\right) \circ \dot{\theta}\right],\end{aligned} \tag{25}$$

where W_i are the transition probabilities of each model ($i = 1,\ldots,5$) whose explicit expressions are given in Appendix B 1. The symbol $\circ$ represents the product in the sense of Stratonovich, and the time increment $t'-t$ is taken to be smaller than the time scales of each model. The heat flux Q_i^h, Q^c, and Q_i from the system to each heat bath are defined as

$$Q_1^h = -\left(\dot{\Pi} + \frac{\partial U(\theta,x)}{\partial \theta} - f\right) \circ \frac{\Pi}{m}, \tag{26}$$

$$Q^c = -\frac{\partial U(\theta,x)}{\partial x} \circ \dot{x}, \tag{27}$$

$$Q_2 = -\left(\dot{\Pi} + \frac{\partial U_{\rm eff}(\theta)}{\partial \theta} - f\right) \circ \frac{\Pi}{m}, \tag{28}$$

$$Q_3 = Q_5 = \left(-\frac{\partial U_{\rm eff}(\theta)}{\partial \theta} + f\right) \circ \dot{\theta}, \tag{29}$$

$$Q_4^h = \left(-\frac{\partial U(\theta,x)}{\partial \theta} + f\right) \circ \dot{\theta}. \tag{30}$$

The entropy production rates obtained from the transition probabilities are different from the heat flux divided by the effective temperature $T_{\rm eff}(\theta)$ in some cases [Eqs. (23) and (25)]. This is because the heat flux is defined based on the consistency with the energy balance [43]. An alternative definition of heat flux and its effect on the thermodynamic efficiency will be discussed in Appendix E.

We note that the elimination of the momentum degree of freedom attached to an isothermal heat bath does not involve any hidden entropy production [16]. Therefore, we here focus on the analysis of the entropy productions in models 1–5.

B. Hidden entropy production in coarse graining to model 3

In this subsection, we focus on the entropy production rates in the limit where model 3 is derived. Since we have the asymptotic behavior of the probability density function $P(\theta,\Pi,x)$ in the limit of tight confinement (cf. Appendix A 1), the ensemble average of σ_1 can be written as

$$\langle\sigma_1\rangle = \langle\sigma_2\rangle + \left\langle \frac{\Gamma(\mathcal{G}(\theta)-\Gamma)}{\mathcal{G}(\theta)T_{\rm eff}(\theta)}\left(\frac{1}{T_c}-\frac{1}{T_h}\right)(T_h-T_c)\frac{\Pi^2}{m^2}\right\rangle. \tag{31}$$

The derivation of Eq. (31) is given in Appendix B 2. Equation (31) states that there is a finite and positive difference between $\langle\sigma_1\rangle$ and $\langle\sigma_2\rangle$ for $T_h \neq T_c$, which is the hidden entropy production between models 1 and 2. This means that the dissipation is underestimated if we assume model 2 as the description of the FS ratchet.

Next, we evaluate the right-hand side of Eq. (31) at the overdamped limit. Taking the ensemble average of σ_2 with respect to the asymptotic form of $P(\theta,\Pi)$ in the overdamped limit, we obtain

$$\langle\sigma_2\rangle = \langle\sigma_3\rangle + \left\langle \frac{T_{\rm eff}(\theta)}{2\mathcal{G}(\theta)}\left(\frac{T'_{\rm eff}(\theta)}{T_{\rm eff}(\theta)}\right)^2\right\rangle. \tag{32}$$

The derivation of Eq. (32) is given in Appendix B 3. The hidden entropy production between models 2 and 3, $\langle\sigma_2\rangle - \langle\sigma_3\rangle$, is positive unless $T_{\rm eff}(\theta)$ is constant. The positivity of these hidden entropy productions is consistent with the general condition discussed in [20].

Since Π^2/m relaxes to $T_{\rm eff}(\theta)$ in the overdamped limit, we finally obtain

$$\langle\sigma_1\rangle = \langle\sigma_3\rangle + \left\langle \frac{T_{\rm eff}(\theta)}{2\mathcal{G}(\theta)}\left(\frac{T'_{\rm eff}(\theta)}{T_{\rm eff}(\theta)}\right)^2\right\rangle + \left\langle \frac{\Gamma(\mathcal{G}(\theta)-\Gamma)}{m\mathcal{G}(\theta)}\left(\frac{1}{T_c}-\frac{1}{T_h}\right)(T_h-T_c)\right\rangle. \tag{33}$$

We note that Eqs. (31)–(33) hold even in the non-steady states when σ's are substituted by the total entropy production [11].

Equation (33) indicates that the true entropy production rate $\langle\sigma_1\rangle$ is positive in the regime where the FS ratchet operates as a heat engine ($T_h \neq T_c$ and $\phi(\theta) \neq$ const), even when $\langle\sigma_3\rangle = 0$ holds. This is consistent with the previous studies of the Büttiker-Landauer motor system [16,29–31]. Indeed, the second term on the right-hand side of Eq. (32) may be considered as a generalization to the case where the frictional coefficient is state dependent.

Equation (33) states that the FS ratchet carries another hidden dissipation expressed as the last term on the right hand side, which also has a significant impact. Evaluating the order of each term in Eq. (33), we obtain

$$\langle\sigma_3\rangle = \int_0^L d\theta \frac{-U'_{\rm eff}(\theta)+f}{T_{\rm eff}(\theta)}\left\langle\frac{\dot\theta}{L}\right\rangle = \Delta\left\langle\frac{\dot\theta}{L}\right\rangle \sim \tau^{-1}, \tag{34}$$

$$\left\langle \frac{T_{\rm eff}(\theta)}{2\mathcal{G}(\theta)}\left(\frac{T'_{\rm eff}(\theta)}{T_{\rm eff}(\theta)}\right)^2\right\rangle \sim \left\langle\frac{T_{\rm eff}(\theta)}{\mathcal{G}(\theta)}\right\rangle\frac{1}{L^2} \sim \tau^{-1}, \tag{35}$$

$$\left\langle \frac{\Gamma(\mathcal{G}(\theta)-\Gamma)}{m\mathcal{G}(\theta)}\left(\frac{1}{T_c}-\frac{1}{T_h}\right)(T_h-T_c)\right\rangle \sim \frac{\Gamma}{m} = \tau_\Pi^{-1}, \tag{36}$$

where we use the fact that $T_{\rm eff}(\theta)/\mathcal{G}(\theta)$ is the effective diffusion coefficient in model 3. Since $\epsilon := \tau_\Pi/\tau$ is the small parameter which controls the overdamped limit, the ratio of Eq. (36) to Eqs. (34) and (35) diverges in the overdamped limit. Therefore, $\langle\sigma_1\rangle$ is dominated by the hidden entropy production [Eq. (36)] between models 1 and 2.

The effect of these hidden entropy productions on the thermodynamic efficiency is numerically investigated in the next section.

C. Hidden entropy production in coarse graining to model 5

We discuss the entropy production rate in the route of coarse graining to obtain model 5. We first have

$$\langle\sigma_1\rangle = \langle\sigma_4\rangle \tag{37}$$

in the overdamped limit, since the elimination of the momentum variable from isothermal dynamics does not involve hidden entropy production. In the tightly confined limit $\varepsilon' := \tau_x/\tau \to 0$,

$$\langle\sigma_4\rangle = \left\langle \frac{\phi'(\theta)^2}{\Gamma T_c + \gamma\phi'(\theta)^2 T_h}\left(\frac{1}{T_c}-\frac{1}{T_h}\right) \times (T_h-T_c)\left(\frac{\partial U_I(x-\phi(\theta))}{\partial x}\right)^2\right\rangle, \tag{38}$$

which is derived in Appendix B 5. The leading order of $\langle\sigma_4\rangle$ is estimated as

$$\langle\sigma_4\rangle \sim \left\langle \frac{\Gamma T_c + \gamma\phi'(\theta)^2 T_h}{\mathcal{G}(\theta)^2}\left(\frac{\mathcal{G}(\theta)}{\Gamma T_c + \gamma\phi'(\theta)^2 T_h}\frac{\partial U_I}{\partial x}\right)^2\right\rangle \lesssim \left\langle \frac{T_h}{\Gamma}\left(\frac{1}{T_c}\frac{\partial U_I}{\partial x}\right)^2\right\rangle \sim \frac{T_h}{\Gamma}\frac{1}{L_x^2} = \tau_x^{-1}. \tag{39}$$

Since $\langle\sigma_5\rangle = O(\tau^{-1})$, the leading order of $\langle\sigma_4\rangle$ does not include $\langle\sigma_5\rangle$. Thus, $\langle\sigma_1\rangle$ in the limit of model 5 is dominated by the hidden entropy production between models 4 and 5.

Let us compare $\langle\sigma_1\rangle$ in the two coarse-graining limits. As we saw in the previous subsection, we have

$$\langle\sigma_1\rangle \sim \left\langle \frac{\Gamma}{m}\frac{\mathcal{G}(\theta)-\Gamma}{\mathcal{G}(\theta)}\left(\frac{1}{T_c}-\frac{1}{T_h}\right)(T_h-T_c)\right\rangle, \tag{40}$$

as the leading order in the limit of obtaining model 3. Choosing $U_I[x-\phi(\theta)] = \lambda[x-\phi(\theta)]^2/2$, the coarse graining toward model 5 leads to

$$\langle\sigma_1\rangle \sim \left\langle \frac{\lambda}{\gamma}\frac{\mathcal{G}(\theta)-\Gamma}{\mathcal{G}(\theta)}\left(\frac{1}{T_c}-\frac{1}{T_h}\right)(T_h-T_c)\right\rangle, \tag{41}$$

since $x-\phi(\theta)$ follows the canonical distribution characterized by $T_s(\theta) = [\Gamma T_c + \gamma\phi'(\theta)^2 T_h]/[\Gamma + \gamma\phi'(\theta)^2]$ at the leading order of ε' (see Appendix A 4). We here find that the only difference between Eqs. (40) and (41) is which time scale is rate limiting ($\tau_\Pi^{-1} = \Gamma/m$ or $\tau_x^{-1} = \lambda/\gamma$).

D. Relation with Feynman's argument

We set $T_h > T_c$ and $f < 0$. In Feynman's argument, the forward (backward) stepping rotation, which produces positive (negative) work, is initiated by the absorption of heat from the hotter (colder) bath, and the excess energy is released to the colder (hotter) bath as the dissipated heat. According to these phenomenological assumptions, he estimated the heat absorbed from the hot and cold baths per forward step as

$Q^h = \Delta U - fL$ and $Q^c = \Delta U$, respectively, where the work per step is $-fL$ and the energy required to lift the pawl is ΔU. Then, the rate of forward and backward steps were considered as

$$R_f^F = \tau_s^{-1} \exp(-Q^h/T_h) \tag{42}$$

$$R_f^B = \tau_s^{-1} \exp(-Q^c/T_c), \tag{43}$$

where τ_s is the characteristic time scale of the steps. Taking into account the backward step, the thermodynamic efficiency when the FS ratchet operates as heat engine is written as

$$\eta_f = \frac{-fL}{Q^h} = 1 - \frac{Q^c}{Q^h} \leqslant \eta_C. \tag{44}$$

where $\eta_C := 1 - T_c/T_h$ is the Carnot efficiency. The equality is satisfied at the stalled condition $R^F = R^B$.

A similar discrete-stepping model may be obtained from models 3 and 5 in the limit of $\Delta U/T_{\rm eff}(\theta) \to \infty$ [36], where $\Delta U = \max_\theta U_{\rm eff}(\theta) - \min_\theta U_{\rm eff}(\theta)$ is the effective energy barrier. From Kramers theory, the forward and backward transition rates of model 3 in this limit are obtained as

$$R_{(3)}^F = \tau_s^{-1} \exp\left[-\int_{\theta_{\rm min}}^{\theta_{\rm max}^+} \frac{1}{T_{\rm eff}(\theta)}\left(\frac{\partial U_{\rm eff}(\theta)}{\partial\theta} - f\right)d\theta\right], \tag{45}$$

$$R_{(3)}^B = \tau_s^{-1} \exp\left[-\int_{\theta_{\rm min}}^{\theta_{\rm max}^-} \frac{1}{T_{\rm eff}(\theta)}\left(\frac{\partial U_{\rm eff}(\theta)}{\partial\theta} - f\right)d\theta\right]. \tag{46}$$

Here, $\theta_{\rm min}$ is the position of a local minimum of $U_{\rm eff}(\theta)$, and $\theta_{\rm max}^\pm$ are the position of the local maxima of $U_{\rm eff}(\theta)$ that are closest to $\theta_{\rm min}(\theta_{\rm max}^+ > \theta_{\rm min}, \theta_{\rm max}^- < \theta_{\rm min}$, and $\theta_{\rm max}^- + L = \theta_{\rm max}^+)$. The derivation of Eqs. (45) and (46) is given in Appendix C. Assuming a sawtooth shape for the gear,

$$\phi'(\theta_{\rm max}^- < \theta < \theta_{\rm min}) = \alpha_c > 0, \tag{47}$$

$$\phi'(\theta_{\rm min} < \theta < \theta_{\rm max}^+) = \alpha_h < 0, \tag{48}$$

the rates reduce to

$$R_{(3)}^F = \tau_s^{-1} \exp\left(-Q_3^h/T_{{\rm eff},h}\right), \tag{49}$$

$$R_{(3)}^B = \tau_s^{-1} \exp\left(-Q_3^c/T_{{\rm eff},c}\right), \tag{50}$$

where we can interpret that the heat

$$Q_3^h = \Delta U - f(\theta_{\rm max}^+ - \theta_{\rm min}), \tag{51}$$

$$Q_3^c = \Delta U - f(\theta_{\rm max}^- - \theta_{\rm min}) \tag{52}$$

are exchanged from the baths with effective temperatures

$$T_{{\rm eff},h} = \frac{\Gamma T_h + \gamma\alpha_h^2 T_c}{\Gamma + \gamma\alpha_h^2} \leqslant T_h, \tag{53}$$

$$T_{{\rm eff},c} = \frac{\Gamma T_h + \gamma\alpha_c^2 T_c}{\Gamma + \gamma\alpha_c^2} \geqslant T_c, \tag{54}$$

respectively. The efficiency then satisfies

$$\eta_3 := \frac{-fL}{Q_3^h} \leqslant 1 - \frac{T_{{\rm eff},c}}{T_{{\rm eff},h}} \leqslant \eta_C. \tag{55}$$

The first equality in Eq. (55) is met at the stalled state. For the second equality, however, we must take the limits $\alpha_c \to \infty$ and $\alpha_h \to 0$, which corresponds to an asymmetric sawtooth with $\theta_{\rm max}^- = \theta_{\rm min}$ and $\theta_{\rm max}^+ - \theta_{\rm min} \to \infty$. We also note that even under this asymmetric limit, where Feynman's rates and efficiency are reproduced, the real thermodynamic efficiency is much lower (effectively zero) since there is a large hidden entropy production between models 1 and 3. The same situation holds for model 5.

E. Relation with Parrondo and Español's model

Choosing $U(\theta,x) = \lambda(x-\theta)^2/2$ and $\Gamma = \gamma$, model 0 is written as

$$\dot\theta = \frac{\Pi}{m}, \qquad \dot\Pi = -\frac{\Gamma}{m}\Pi + \lambda(x-\theta) + \sqrt{2\Gamma T_h}\xi,$$
$$\dot x = \frac{p}{m_x}, \qquad \dot p = -\frac{\Gamma}{m_x}p - \lambda(x-\theta) + \sqrt{2\Gamma T_c}\zeta. \tag{56}$$

Note that there is no net rotation in this model due to the symmetry of the potential. Parrondo and Español identified the continuous heat flow in a case of $m = m_x$ as [28]

$$J_{PE} = \frac{\Gamma}{2m}\frac{\lambda m/\Gamma^2}{(1+\lambda m/\Gamma^2)}(T_h - T_c). \tag{57}$$

By taking the tightly confined limit, we obtain the coarse-grained description (model 2) for this model as

$$\dot\theta = \frac{\Pi}{m},$$
$$\dot\Pi = -\frac{2\Gamma}{m}\Pi + \sqrt{2\Gamma(T_h+T_c)}\Xi. \tag{58}$$

Correspondingly, the hidden entropy production rate becomes

$$\langle\sigma_1\rangle - \langle\sigma_2\rangle = \frac{\Gamma}{2m}\left(\frac{1}{T_c} - \frac{1}{T_h}\right)(T_h - T_c), \tag{59}$$

which means that the heat flow converges to $\Gamma(T_h - T_c)/(2m)$. As pointed out by Parrondo and Español [28], this nonvanishing heat flow prevents the FS ratchet to acquire Carnot efficiency. In their argument, the deviation from Gibbs-Boltzmann distribution played a crucial role. This deviation, however, seems to disappear when taking the tightly confined limit [Eqs. (58) and (A11)]. In fact, our result suggests that even an infinitesimal deviation from the Gibbs-Boltzmann distribution can contribute to a finite entropy production.

V. NUMERICAL SIMULATION

We performed numerical simulations of the FS ratchet with $U(\theta,x) = kx^2/2 + \lambda[x - \phi(\theta)]^2/2, \phi(\theta) = \sin(2\pi\theta) + 0.25\sin(4\pi\theta) + 1.1$, $\Gamma = 5.0$, $\gamma = 0.05$, $k = 1.0$, $T_h = 1.1$, $T_c = 0.9$. In this setting, $L_x = \sqrt{T_c/\lambda}$ and the shortest time scale included in τ is $\Gamma/T_h = 4.5$. To obtain the limit of model 3, we introduced λ_0 as $\lambda = \lambda_0/m$ and changed m and λ_0 to control the separation of time scales. The limit of $\lambda_0 \to \infty$ corresponds to the tight confinement of the pawl to the ratchet, $\varepsilon \simeq \gamma/\lambda\tau_\Pi \to 0$. The limit of $m \to 0$ realizes the overdamped limit, $\epsilon = mT_h/\Gamma^2 \to 0$, while keeping the ratio of γ/λ to τ_Π proportional to λ_0. The functional form of $\phi(\theta)$ is shown in Fig. 3. By choosing $\phi(\theta)$ to be asymmetric,

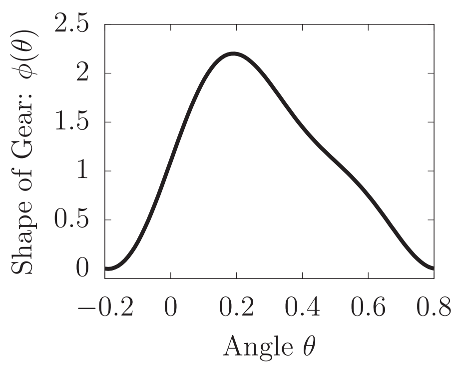

FIG. 3. The functional form of $\phi(\theta) = \sin(2\pi\theta) + 0.25\sin(4\pi\theta) + 1.1$ we used for the numerical simulation.

$U_{\text{eff}}(\theta)$ and $T_{\text{eff}}(\theta)$ obtained from Eqs. (7) and (9) become out of phase as shown in Fig. 4. The other details of the numerical simulations are given in Appendix D.

A. Entropy production rates

Numerical results of the steady-state entropy production rates are plotted in Fig. 5. The left figure of Fig. 5 shows that in the tightly confined regime $\varepsilon^{-1} \approx \tau_\Pi/(\gamma/\lambda) \gtrsim 40$, $\langle\sigma_1\rangle$ converges to the right-hand side of Eq. (31). As shown in the right figure of Fig. 5, we see the convergence of $\langle\sigma_2\rangle$ to the right-hand side of Eq. (32), in the overdamped limit $\epsilon \to 0$, where we fixed the parameters so as to $\tau_\Pi/(\gamma/\lambda) = 80$ to be at the tightly confined limit. Furthermore, we see that $\langle\sigma_1\rangle$ diverges with ϵ^{-1}, consistent with Eq. (33).

B. Thermodynamic efficiencies

To demonstrate the impact of hidden entropy production, we calculated the thermodynamic efficiencies defined in models 1–3. In model 1, the thermodynamic efficiency is $\eta_1 := 1 - \langle Q^c\rangle/\langle Q_1^h\rangle$. For models 2 and 3, the definition

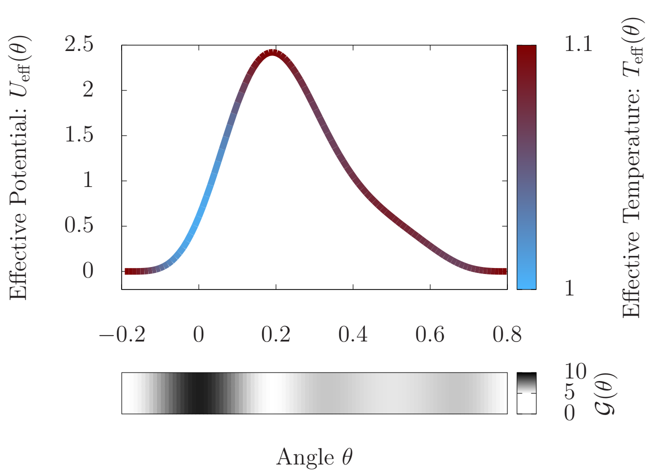

FIG. 4. The functional forms of $U_{\text{eff}}(\theta)$,$\mathcal{G}(\theta)$ and $T_{\text{eff}}(\theta)$ [Eqs. (7)–(9)] which follow from the setup in the numerical simulation, $U(\theta,x) = kx^2/2 + \lambda[x - \phi(\theta)]^2/2, \phi(\theta) = \sin(2\pi\theta) + 0.25\sin(4\pi\theta) + 1.1$. The temperature difference between the positions of positive and negative potential slopes causes a net flow.

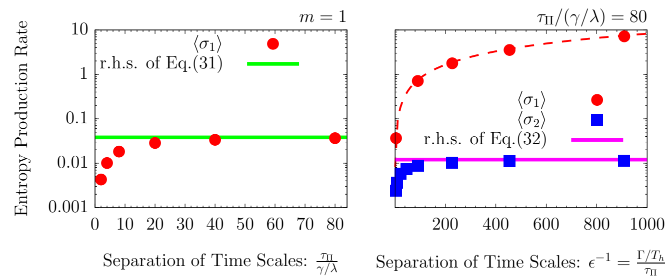

FIG. 5. Steady state entropy production rates. Filled circles and squares are the numerical results. Solid lines are obtained through Eqs. (31) and (32), respectively. The dashed line $\propto \epsilon^{-1}$ shows the asymptotic dependence of $\langle\sigma_1\rangle$ on ϵ.

of the efficiency is not trivial, since there is only a single heat bath with nonuniform continuous temperature. We here adopt a generalized definition of efficiency [46]. The average heat flux conditional on the effective temperature T is introduced by

$$\langle Q_{2,3}(T)\rangle := \langle Q_{2,3}\delta[T_{\text{eff}}(\theta) - T]\rangle. \tag{60}$$

The averaged heat release and absorption rates are then defined as

$$\begin{aligned} \langle Q_{2,3}^{\text{rel}}\rangle &= \int dT\,\langle Q_{2,3}(T)\rangle\Theta[\langle Q_{2,3}(T)\rangle], \\ \langle Q_{2,3}^{\text{abs}}\rangle &= -\int dT\,\langle Q_{2,3}(T)\rangle\Theta[-\langle Q_{2,3}(T)\rangle], \end{aligned} \tag{61}$$

where Θ represents the Heaviside step function. The generalized efficiencies η_2 and η_3 are

$$\eta_{2,3} = 1 - \frac{\langle Q_{2,3}^{\text{rel}}\rangle}{\langle Q_{2,3}^{\text{abs}}\rangle}. \tag{62}$$

For model 3, this definition agrees with the efficiency introduced as Eq. (55) for the special case of sawtooth potential in the discrete stepping limit.

The dependence of the efficiencies on f and ϵ are shown in Fig. 6. The behavior of η_1 and η_2 are different from

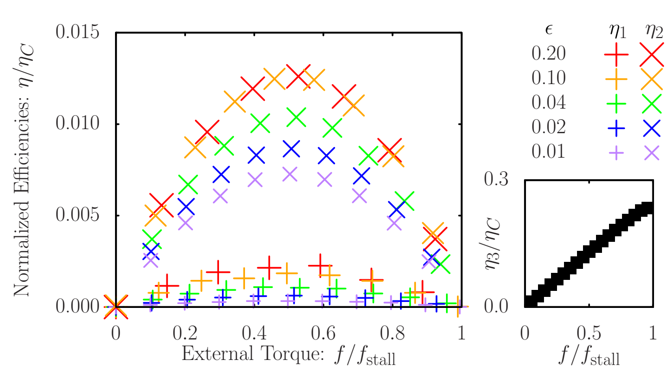

FIG. 6. Thermodynamic efficiencies as functions of the external torque f. In the left figure, η_1 and η_2 are parametrized by the separation of time scales ϵ. Smaller symbols correspond to cases of smaller epsilon. In the right figure, η_3 is plotted. The efficiencies are normalized by the Carnot efficiency $\eta_C = 1 - T_c/T_h$.

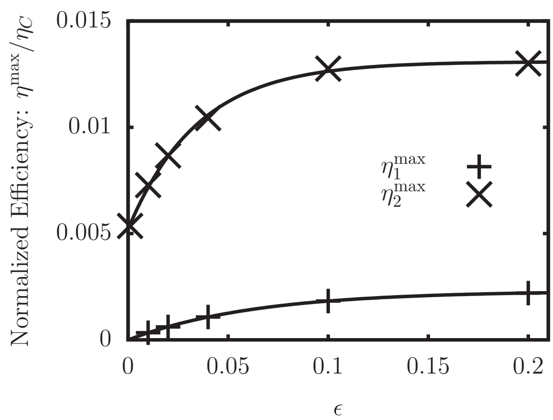

FIG. 7. Maximal efficiency, $\eta_i^{\max} = \max_f \eta_i$. Maximal value of η_1 and η_2 obtained from the fitting of Fig. 6 by parabolic functions are plotted against ϵ. Solid lines $A\exp(-\epsilon/\epsilon_0) + C$ are also plotted as guide for the eyes, where A, ϵ_0, and C are fitting parameters. The efficiencies are normalized by the Carnot efficiency of model 0, $\eta_C = 1 - T_c/T_h$.

that of η_3 at $\epsilon \to 0$. First, η_1 approaches to 0 regardless of f, since it follows from Eq. (15) that the steady-state power, $\dot{W} = -f\langle\dot{\theta}\rangle$, is of $O(\tau^{-1})$, while the heat flows at the rate with $O(\tau_\Pi^{-1})$. Second, η_2 does not converge to zero (Fig. 6). The finite efficiency means that $\langle Q_2^{abs}\rangle$ is of the same order as $\dot{W} = O(\tau^{-1})$. Here, $\langle Q_2^{abs}\rangle$ fails to capture the heat flow of $O(\tau_\Pi^{-1})$ which contributes to the hidden entropy production $\langle\sigma_1\rangle - \langle\sigma_2\rangle$. Nevertheless, η_2 vanishes at the stalled state, since $\langle Q_2^{abs}\rangle$ is kept finite while $\dot{W} \to 0$. Third, η_3 monotonically increases and reaches the maximal value at the stalled state. This corresponds to the seemingly reversible situation, $\langle\sigma_3\rangle = 0$. However, η_3 does not reach Carnot efficiency η_C, because $1.0 < T_{\rm eff}(\theta) < 1.1$ in this setup, which implies $\eta_3 < 1 - 1.0/1.1 < \eta_C$. In Fig. 7, we show the ϵ dependence of maximal efficiency, $\eta_i^{\max} := \max_f \eta_i$, obtained from the fitting of torque-efficiency curves. This result indicates that, in the limit of $\epsilon \to 0$, η_1 vanishes irrespective of f, and η_2 converges to a certain torque-dependent curve.

These results highlight the effects of coarse graining on the qualitative behaviors of thermodynamic efficiency; one may assume a significantly higher efficiency of an engine by neglecting the dissipative contributions of the fast variables.

VI. RECOVERY OF ENTROPY PRODUCTION BASED ON DECOMPOSITION OF COARSE-GRAINED LANGEVIN DYNAMICS

The exact expression of the entropy production rate in the tightly confined limit [Eq. (31)] inspires us to consider if it is possible to reconstruct the thermodynamic irreversibility defined at the fine-grained description from the observation at the coarse-grained scale. In a system where the time scales of variables are well separated, it is challenging to probe the dynamics of the fast variable, meaning that the hidden entropy production and the real thermodynamic efficiency are almost impossible to measure [24]. Although there is no general workaround to the problem of inaccessible fast variables, we here describe a way to evaluate $\langle\sigma_1\rangle$ from model 2 of the FS ratchet.

This is achieved by considering the dynamics as a mixture of two Langevin dynamics with different temperatures and frictions corresponding to the two heat baths [Fig. 1(d)], instead of a single set of effective temperature and friction [Eqs. (8) and (9)]. The dynamics we refer to as model 6 consists of two Langevin equations,

$$\begin{aligned}\dot{\theta} &= \frac{\Pi}{m},\\ \dot{\Pi} &= -\frac{\Gamma_b(\theta)}{m}\Pi - \frac{\partial U_{\rm eff}(\theta)}{\partial\theta} + f + \sqrt{2\Gamma_b(\theta)T_b}\xi,\end{aligned} \tag{63}$$

and stochastic switching of an auxiliary variable $b = h,c$, which controls which heat bath $[(\Gamma_h,T_h)$ or $(\Gamma_c,T_c)]$ the Langevin dynamics should be governed by. The stochastic process of θ, Π, and b is described by the master equation

$$\begin{aligned}\frac{\partial P(\theta,\Pi,b)}{\partial t} =& -\frac{\partial}{\partial\theta}\left(\frac{\Pi}{m}P(\theta,\Pi,b)\right)\\ &-\frac{\partial}{\partial\Pi}\bigg[\bigg(-\frac{\Gamma_b(\theta)}{m}\Pi - \frac{\partial U_{\rm eff}(\theta)}{\partial\theta}\\ &+f - \Gamma_b(\theta)T_b\frac{\partial}{\partial\Pi}\bigg)P(\theta,\Pi,b)\bigg]\\ &-\Lambda P(\theta,\Pi,b) + \Lambda P(\theta,\Pi,b'),\end{aligned} \tag{64}$$

where $b' = c,h$ for $b = h,c$, $P(\theta,\Pi,b)$ is the joint probability density of θ, Π, and b, and Λ is the rate of stochastic switching of the heat baths. According to the singular perturbation theory, in the limit where Λ^{-1} is separated from τ_Π and τ, model 6 will give effective dynamics that follows

$$\begin{aligned}\dot{\theta} &= \frac{\Pi}{m},\\ \dot{\Pi} &= -\frac{\Gamma_h(\theta)+\Gamma_c(\theta)}{2m}\Pi - \frac{\partial U_{\rm eff}(\theta)}{\partial\theta} + f\\ &\quad + \sqrt{[\Gamma_h(\theta)T_h + \Gamma_c(\theta)T_c]}\,\Xi.\end{aligned} \tag{65}$$

Therefore, by setting $\Gamma_h(\theta) = 2\Gamma$ and $\Gamma_c(\theta) = 2\gamma\phi'(\theta)^2$, Eq. (65) will reproduce the dynamics of model 2.

The entropy production of model 6 is

$$\sigma_6 = -\frac{1}{T_b}\left(\dot{\Pi} + \frac{\partial U_{\rm eff}(\theta)}{\partial\theta} - f\right)\circ\frac{\Pi}{m}. \tag{66}$$

In the limit of fast switching, $\langle\sigma_6\rangle$ converges to the weighted average of the contributions from the two dynamics,

$$\langle\sigma_6\rangle \xrightarrow{\{\Lambda\tau_\Pi,\Lambda\tau\}\to\infty} \frac{1}{2}\left\langle -\frac{1}{T_h}\left(\dot{\Pi}+\frac{\partial U_{\rm eff}(\theta)}{\partial\theta}-f\right)\circ\frac{\Pi}{m}\right\rangle_h + \frac{1}{2}\left\langle -\frac{1}{T_c}\left(\dot{\Pi}+\frac{\partial U_{\rm eff}(\theta)}{\partial\theta}-f\right)\circ\frac{\Pi}{m}\right\rangle_c, \tag{67}$$

where the subscripts h,c indicate which Langevin dynamics are used to calculate the ensemble average. By comparing Eq. (31) [or Eq. (B10)] with Eq. (67), we obtain

$$\lim_{\{\Lambda\tau_\Pi,\Lambda\tau\}\to\infty}\langle\sigma_6\rangle = \lim_{\varepsilon\to 0}\langle\sigma_1\rangle. \tag{68}$$

The details of the derivation of Eqs. (65) and (68) are given in Appendix F.

Equation (68) is useful when we know the original temperatures of the heat baths but can only observe the dynamics at the coarse-grained scale. Since $\mathcal{G}(\theta)$ and $T_{\rm eff}(\theta)$ can be measured at the coarse-grained scale, we may solve Eqs. (8) and (9) using T_h and T_c to obtain $\Gamma_b(\theta)$ in such a situation, which allows the evaluation of $\langle\sigma_6\rangle$. We note that the decomposition of model 2 into dynamics involving T_h and T_c is not unique if we are allowed to use general θ-dependent switching rates. Nevertheless, we may show that Eq. (68) always holds as far as model 2 is obtained in the fast switching limit (see Appendix F). The formulation of entropy production based on the decomposition of the stochastic transition has been previously discussed [18,28,43]. The approach here is a natural extension of these strategies to the case of a heat engine described by continuous variables.

VII. DISCUSSION AND CONCLUSION

We derived the coarse-grained descriptions of the FS ratchet starting from model 0 along the routes shown in Fig. 2. We obtained the exact expressions for the entropy production in each model and clarified the existence of the hidden entropy productions, which are the differences in the entropy production between the different descriptions. The impact of the hidden entropy production on the thermodynamic efficiency was investigated numerically, to track how the efficiency of models 2 and 3 significantly overestimate the true thermodynamic efficiency of model 0. Additionally, we proposed a way to reconstruct the entropy production for model 0 from the coarse-grained scale by introducing a pseudodynamics described by model 6.

Some of the coarse-grained descriptions obtained in this work have been studied as effective models of the FS ratchet, e.g., the BL motor (models 2 and 3) [30], the two-variable overdamped model (model 4) [37,44], the single-variable model (model 3) [29,47], and the discrete stepping model (Sec. IV D) [26,48–55]. This means that phenomenological descriptions in the previous works were correct when taking appropriate time-scale separation limits. However, the thermodynamic efficiency of the FS ratchet based on these models have been controversial and often misleading. For instance, the thermodynamic efficiency of the BL motor (models 2 or 3), if assuming it as an effective model of the FS ratchet, will always be an overestimation because of the hidden entropy production. Similarly, the efficiencies of single-variable models such as models 3 or 5 [29,48–55] are also overestimations. Although some studies [29,48–52] phenomenologically consider the dissipation corresponding to the hidden entropy production, the existence of two layers of hidden entropy production has not been discussed. In [37], the efficiency was correctly obtained as lower than Carnot, since the starting point corresponding to model 4 has no hidden entropy production compared with the original FS ratchet.

The problem of hidden entropy production is inevitable when analyzing the thermodynamic aspect of nonequilibrium system, since most of the models are constructed phenomenologically. In this sense, model 6 points at a promising solution for the hidden entropy production, since it enables us to evaluate the true entropy production without knowing the true fine-grained description (model 0). Therefore, it is important to develop such a framework, which may extract thermodynamic properties from the coarse-grained descriptions.

ACKNOWLEDGMENTS

We thank E. Muneyuki, S. Wang, Y. Murashita, and M. Itami for fruitful comments. This work was supported by the Grant-in-Aids for JSPS Research Fellow, Grants No. 24-3258 (Y.N.), No. 24-8031, and No. 28-908 (K.K.), Grant-in-Aid for Young Scientists (B), Grant No. 17K14373 (Y.N.), Grant-in-Aid for Scientific Research (C) No. 15K05196 (N.N.), and Grant-in-Aid for Scientific Research on Innovative Areas No. 25103002 (N.N.).

APPENDIX A: DERIVATION OF COARSE-GRAINED DYNAMICS

1. Derivation of model 2

In this subsection, we describe the details of the derivation of model 2. Our goal here is to obtain the time evolution equation of the joint probability density $P(\theta,\Pi)$ from Eq. (4). We evaluate the last term on the right-hand side of Eq. (5) in the limit of $\varepsilon := \tau_x/\tau_\Pi \to 0$. The heart of the singular perturbation theory is to decompose the time dependence of $P(\theta,\Pi,x)$ into the explicit part and the implicit part through $P(\theta,\Pi)$. In the limit of $\varepsilon\to 0$, the explicit part decays quickly and the right-hand side of Eq. (5) essentially turns into a functional of $P(\theta,\Pi)$.

A singular perturbation problem is mapped to an ordinary perturbation theory by introducing M which describes $P(\theta,\Pi,x)$ and Ω which describes the dynamics of $P(\theta,\Pi)$. For this purpose, we first switch the variables from (t,x) to the dimensionless time and distance,

$$\mathcal{T} := \frac{t}{\tau_x}, \quad s := \frac{x-\phi(\theta)}{L_x}, \tag{A1}$$

and rewrite Eq. (4) as

$$\tau_x^{-1}\frac{\partial P(\theta,\Pi,s)}{\partial\mathcal{T}} = (\mathcal{L}^\theta+\mathcal{L}^\Pi)P(\theta,\Pi,s) + \tau_x^{-1/2}\left[\sqrt{\frac{\gamma}{T_c}}\phi'(\theta)\left(\frac{\partial}{\partial s}\frac{\Pi}{m}-\frac{\partial}{\partial\Pi}\frac{\partial\tilde{U}_I(s)}{\partial s}\right)P(\theta,\Pi,s)\right]$$
$$-\tau_x^{-1}\frac{\partial}{\partial s}\left[-\frac{1}{T_c}\frac{\partial\tilde{U}_I(s)}{\partial s}P(\theta,\Pi,s)-\frac{\partial P(\theta,\Pi,s)}{\partial s}\right]-\tau_x^{-1/2}\frac{\partial}{\partial s}\left[-\sqrt{\frac{T_c}{\gamma}}\left[\frac{U_0'(\phi(\theta))}{T_c}+O\left(\frac{|U_0''|L_x}{|U_0'|}\right)\right]P(\theta,\Pi,s)\right], \tag{A2}$$

where

$$\mathcal{L}^\theta := -\frac{\partial}{\partial\theta}\frac{\Pi}{m},\quad \mathcal{L}^\Pi := -\frac{\partial}{\partial\Pi}\left[\left(-\frac{\Gamma}{m}\Pi+f\right)-\Gamma T_h\frac{\partial}{\partial\Pi}\right]. \tag{A3}$$

In terms of τ_x, τ_Π, and τ, we may estimate the order of the terms on the right-hand side of Eq. (A2) as $O(\tau^{-1/2}\tau_\Pi^{-1/2})$, $O(\tau_x^{-1/2}\tau_\Pi^{-1/2})$, $O(\tau_x^{-1})$, and $O(\tau_x^{-1/2}\tau^{-1/2})$, respectively. In addition, $O(|U_0''|L_x/|U_0'|)=O[(\tau_x/\tau)^{1/2}]$. Based on this order estimation, we may consider ε as the small parameter which controls the perturbative analysis.

The explicit and implicit dependence of $P(\theta,\Pi,s)$ on $\mathcal{T}$ is implemented by describing $P(\theta,\Pi,s)$ as output of a $\mathcal{T}$-dependent operator, M, that acts on $P(\theta,\Pi)$:

$$P(\theta,\Pi,s) = M[P(\theta',\Pi');\mathcal{T}](\theta,\Pi,s), \tag{A4}$$

where θ' and Π' are dummy variables placed only to indicate that M depends on the joint probability density of θ and Π. Furthermore, we represent the time evolution of $P(\theta,\Pi)$ by a $\mathcal{T}$-dependent operator Ω that acts on $P(\theta,\Pi)$:

$$\frac{\partial P(\theta,\Pi)}{\partial\mathcal{T}} = \Omega[P(\theta',\Pi');\mathcal{T}](\theta,\Pi) := \tau_x(\mathcal{L}^\theta+\mathcal{L}^\Pi)P(\theta,\Pi)$$
$$-\frac{\partial}{\partial\Pi}\left[\tau_x^{1/2}\sqrt{\frac{\gamma}{T_c}}\phi'(\theta)\int ds\frac{\partial\tilde{U}_I(s)}{\partial s}M[P(\theta',\Pi');\mathcal{T}](\theta,\Pi,s)\right], \tag{A5}$$

which is obtained by integrating Eq. (A2) with respect to s. Since M depends on $\mathcal{T}$ explicitly and implicitly [through $P(\theta,\Pi)$], the substitution of M into the left-hand side of Eq. (A2) gives

$$[\text{lhs of Eq. (A2)}] = \frac{\partial M[P(\theta',\Pi');\mathcal{T}](\theta,\Pi,s)}{\partial\mathcal{T}} + \int d\theta''d\Pi''\frac{\partial P(\theta'',\Pi'')}{\partial\mathcal{T}}\frac{\delta M[P(\theta'',\Pi'');\mathcal{T}](\theta,\Pi,s)}{\delta P(\theta'',\Pi'')}$$
$$= \frac{\partial M[P(\theta',\Pi');\mathcal{T}](\theta,\Pi,s)}{\partial\mathcal{T}} + \int d\theta''d\Pi''\Omega[P(\theta',\Pi');\mathcal{T}](\theta'',\Pi'')\frac{\delta M[P(\theta'',\Pi'');\mathcal{T}](\theta,\Pi,s)}{\delta P(\theta'',\Pi'')}, \tag{A6}$$

according to the chain rule. Applying Eq. (A4) also in the right-hand side of Eq. (A2), we obtain

$$\frac{\partial M[P(\theta',\Pi');\mathcal{T}](\theta,\Pi,s)}{\partial\mathcal{T}} + \int d\theta''d\Pi''\Omega[P(\theta',\Pi');\mathcal{T}](\theta'',\Pi'')\frac{\delta M[P(\theta'',\Pi'');\mathcal{T}](\theta,\Pi,s)}{\delta P(\theta'',\Pi'')}$$
$$= \tau_x(\mathcal{L}^\theta+\mathcal{L}^\Pi)M[P(\theta',\Pi');\mathcal{T}](\theta,\Pi,s) + \tau_x^{1/2}\sqrt{\frac{\gamma}{T_c}}\phi'(\theta)\left(\frac{\partial}{\partial s}\frac{\Pi}{m}-\frac{\partial}{\partial\Pi}\frac{\partial\tilde{U}_I(s)}{\partial s}\right)M[P(\theta',\Pi');\mathcal{T}](\theta,\Pi,s)$$
$$-\frac{\partial}{\partial s}\left[-\frac{1}{T_c}\frac{\partial\tilde{U}_I(s)}{\partial s}-\frac{\partial}{\partial s}\right]M[P(\theta',\Pi');\mathcal{T}](\theta,\Pi,s)$$
$$-\frac{\partial}{\partial s}\left[-\tau_x^{1/2}\sqrt{\frac{T_c}{\gamma}}\left(\frac{U_0'(\phi(\theta))}{T_c}+O(\varepsilon^{1/2})\right)M[P(\theta',\Pi');\mathcal{T}](\theta,\Pi,s)\right]. \tag{A7}$$

The remaining task is to apply the standard procedure of perturbation theory. We expand M and Ω into series of $\varepsilon^{1/2}$:

$$M[P(\theta',\Pi');\mathcal{T}](\theta,\Pi,s) = \sum_{n=0}\varepsilon^{n/2}M^{(n)}[P(\theta',\Pi');\mathcal{T}](\theta,\Pi,s), \tag{A8}$$

$$\Omega[P(\theta',\Pi');\mathcal{T}](\theta,\Pi) = \sum_{n=0}\varepsilon^{(n+1)/2}\Omega^{(n)}[P(\theta',\Pi');\mathcal{T}](\theta,\Pi). \tag{A9}$$

Here, the difference in the lowest order for M and Ω is due to Eq. (A5). The leading order of Eq. (A7) gives

$$\frac{\partial M^{(0)}[P(\theta',\Pi');\mathcal{T}](\theta,\Pi,s)}{\partial\mathcal{T}} = -\frac{\partial}{\partial s}\left[-\frac{1}{T_c}\frac{\partial\tilde{U}_I(s)}{\partial s}-\frac{\partial}{\partial s}\right]M^{(0)}[P(\theta',\Pi');\mathcal{T}](\theta,\Pi,s), \tag{A10}$$

from which we obtain

$$M^{(0)}[P(\theta',\Pi');\mathcal{T}](\theta,\Pi,s) = P(\theta,\Pi)\frac{\exp(-\tilde{U}_I(s)/T_c)}{Z} + \cdots, \tag{A11}$$

where $Z = \int ds \exp(-\tilde{U}_I(s)/T_c)$. The additional terms ... depend on $\mathcal{T}$ explicitly, and can be neglected since they decay exponentially with the time scale of $O(\tau_x)$. Under this assumption of the time scale, $\Omega^{(0)}[P(\theta',\Pi');\mathcal{T}](\theta,\Pi)$ vanishes, since the last term on the right-hand side of Eq. (A5) is zero in the leading order. The subleading order of Eq. (A7) is

$$\begin{aligned}\frac{\partial M^{(1)}[P(\theta',\Pi');\mathcal{T}](\theta,\Pi,s)}{\partial\mathcal{T}} =& \left(\frac{\tau_x}{\varepsilon}\right)^{1/2}\sqrt{\frac{\gamma}{T_c}}\phi'(\theta)\left(\frac{\partial}{\partial s}\frac{\Pi}{m} - \frac{\partial}{\partial\Pi}\frac{\partial\tilde{U}_I(s)}{\partial s}\right)M^{(0)}[P(\theta',\Pi');\mathcal{T}](\theta,\Pi,s)\\ &-\frac{\partial}{\partial s}\left[\left(-\frac{1}{T_c}\frac{\partial\tilde{U}_I(s)}{\partial s} - \frac{\partial}{\partial s}\right)M^{(1)}[P(\theta',\Pi');\mathcal{T}](\theta,\Pi,s)\right.\\ &\left.-\left(\frac{\tau_x}{\varepsilon}\right)^{1/2}\sqrt{\frac{T_c}{\gamma}}\frac{U_0'[\phi(\theta)]}{T_c}M^{(0)}[P(\theta',\Pi');\mathcal{T}](\theta,\Pi,s)\right],\end{aligned} \tag{A12}$$

which has a particular solution

$$\begin{aligned}M^{(1)}[P(\theta',\Pi');\mathcal{T}](\theta,\Pi,s) \propto s\frac{\exp(-\tilde{U}_I(s)/T_c)}{Z}\left[-\gamma\phi'(\theta)\left(\frac{\Pi}{m} + T_c\frac{\partial}{\partial\Pi}\right) - U_0'[\phi(\theta)]\right]P(\theta,\Pi)\\ +\,[\text{exponentially decaying terms}].\end{aligned} \tag{A13}$$

By substituting Eq. (A13) into Eq. (A5),

$$\begin{aligned}\Omega^{(1)}[P(\theta',\Pi');\mathcal{T}] =& -\frac{\tau_x}{\varepsilon}\left\{\frac{\partial}{\partial\theta}\left[\frac{\Pi}{m}P(\theta,\Pi)\right] + \frac{\partial}{\partial\Pi}\left[\left(-\frac{\Gamma}{m}\Pi + f\right) - \Gamma T_h\frac{\partial}{\partial\Pi}\right]P(\theta,\Pi)\right\}\\ &-\frac{\partial}{\partial\Pi}\left[\left(\frac{\tau_x}{\varepsilon}\right)^{1/2}\sqrt{\frac{\gamma}{T_c}}\phi'(\theta)\int ds\frac{\partial\tilde{U}_I(s)}{\partial s}M^{(1)}[P(\theta',\Pi');\mathcal{T}](\theta,\Pi,s)\right]\\ =& -\frac{\tau_x}{\varepsilon}\left\{\frac{\partial}{\partial\theta}\left[\frac{\Pi}{m}P(\theta,\Pi)\right] + \frac{\partial}{\partial\Pi}\left[\left(-\frac{\Gamma}{m}\Pi + f\right) - \Gamma T_h\frac{\partial}{\partial\Pi}\right]P(\theta,\Pi)\right.\\ &\left.+\frac{\partial}{\partial\Pi}\left[-\gamma\phi'(\theta)^2\left(\frac{\Pi}{m} + T_c\frac{\partial}{\partial\Pi}\right) - \phi'(\theta)U_0'(\phi(\theta))\right]P(\theta,\Pi)\right\}.\end{aligned} \tag{A14}$$

The Kramers equation [Eq. (6)] immediately follows from the relation $\partial P(\theta,\Pi)/\partial\mathcal{T} = \Omega[P(\theta',\Pi');\mathcal{T}]$ with Eqs. (7)–(9).

2. Quick derivation of model 2

In the procedure of temporal coarse graining, we first formally solve the equation of motion of eliminated variable x [Eq. (11)] as

$$\begin{aligned}x_t =& \frac{1}{\gamma}\int_{-\infty}^{t} dt' e^{-(k+\lambda)/\gamma(t-t')}[\lambda\phi(\theta_{t'}) + \sqrt{2\gamma T_c}\tilde{\zeta}_{t'}]\\ =& \frac{\lambda}{k+\lambda}\phi(\theta) - \frac{\gamma\lambda}{(k+\lambda)^2}\phi'(\theta_t)\frac{\Pi_t}{m}[1+o(\varepsilon)] + \frac{\sqrt{2\gamma T_c}}{\gamma}\int_{-\infty}^{t} dt' e^{-(k+\lambda)/\gamma(t-t')}\tilde{\zeta}_{t'}.\end{aligned} \tag{A15}$$

Here, we performed integration by part twice. The substitution of Eq. (A15) into the equation of motion of Π gives Eq. (12). As done in [41], the underdamped Langevin equation is obtained by integrating Eq. (12) over $t \in [t_0, t_0+\Delta t]$, using the identity

$$\int_{t_0}^{t_0+\Delta t} dt \int_{-\infty}^{t} dt' = \int_{-\infty}^{t_0} dt' \int_{t_0}^{t_0+\Delta t} dt + \int_{t_0}^{t_0+\Delta t} dt' \int_{t'}^{t_0+\Delta t} dt. \tag{A16}$$

By neglecting the $O(\sqrt{2\gamma T_c}\frac{\lambda}{k+\lambda}\int_{-\infty}^{t_0} dt' e^{-(k+\lambda)/\gamma(t_0-t')}\tilde{\zeta}_{t'})$ term, we obtain

$$\frac{\Pi_{t_0+\Delta t}-\Pi_{t_0}}{\Delta t} = -\frac{\Gamma+\frac{\gamma\lambda^2}{(k+\lambda)^2}\phi'(\theta_t)^2}{m}\Pi + f - \frac{k\lambda}{k+\lambda}\phi'(\theta)\phi(\theta) + \sqrt{2\Gamma T_h}\xi_t + \sqrt{2\gamma\phi'(\theta_t)^2 T_c}\frac{\lambda}{k+\lambda}\frac{1}{\Delta t}\int_{t_0}^{t_0+\Delta t} dt'\tilde{\zeta}_{t'}. \quad \text{(A17)}$$

Since $\sqrt{m/k}$ should be included in the set of slow time scales, $k/\lambda = O(\tau_x/\tau) = O(\varepsilon)$. Therefore, Eq. (A17) results in model 2 in the limit of $\varepsilon \to 0$.

3. Derivation of model 3

The coarse graining from model 2 to model 3 can also be formulated through the framework of Appendix A 1. By introducing the dimensionless time and momentum

$$\tilde{\mathcal{T}} = \frac{t}{\tau_\Pi}, \quad \varpi = \frac{\Pi}{\sqrt{mT_0}}, \quad \text{(A18)}$$

where T_0 is the reference point of temperature, the Kramers equation [Eq. (6)] corresponding to model 2 may be rewritten as

$$\frac{\Gamma}{m}\frac{\partial P(\theta,\varpi)}{\partial\tilde{\mathcal{T}}} = -\frac{\partial}{\partial\theta}\left(\sqrt{\frac{T_0}{m}}\varpi P(\theta,\varpi)\right) - \frac{\partial}{\partial\varpi}\left[\left(-\frac{1}{\sqrt{mT_0}}\frac{\partial U_{\rm eff}(\theta)}{\partial\theta} + \frac{1}{\sqrt{mT_0}}f\right)P(\theta,\varpi)\right.$$
$$\left. -\frac{\Gamma}{m}\frac{\mathcal{G}(\theta)}{\Gamma}\left(\varpi P(\theta,\varpi) - \frac{T_{\rm eff}(\theta)}{T_0}\frac{\partial P(\theta,\varpi)}{\partial\varpi}\right)\right]. \quad \text{(A19)}$$

The first line is $O(\tau^{-1/2}\tau_\Pi^{-1/2})$, and the second line is $O(\tau_\Pi^{-1})$, respectively. Following the procedure in Appendix A 1, we define

$$\tilde{M}[P(\theta');\tilde{\mathcal{T}}](\theta,\varpi) := P(\theta,\varpi) \quad \text{(A20)}$$

$$\tilde{\Omega}[P(\theta');\tilde{\mathcal{T}}](\theta) := -\tau_\Pi\int d\varpi\frac{\partial}{\partial\theta}\left(\sqrt{\frac{T_0}{m}}\varpi\tilde{M}[P(\theta');\tilde{\mathcal{T}}](\theta,\varpi)\right) \quad \text{(A21)}$$

$$= \int d\varpi\frac{\partial P(\theta,\varpi)}{\partial\tilde{\mathcal{T}}} = \frac{\partial P(\theta)}{\partial\tilde{\mathcal{T}}}, \quad \text{(A22)}$$

where $P(\theta) = \int d\varpi P(\theta,\varpi)$. In the standard perturbation theory of Eq. (A19) expressed in terms of $\tilde{M}$ and $\tilde{\Omega}$ with a small parameter $\epsilon = \tau_\Pi/\tau$, the leading order gives

$$\frac{\partial\tilde{M}^{(0)}[P(\theta');\tilde{\mathcal{T}}](\theta,\varpi)}{\partial\tilde{\mathcal{T}}} = -\frac{\partial}{\partial\varpi}\left[-\frac{\mathcal{G}(\theta)}{\Gamma}\varpi - \frac{\mathcal{G}(\theta)}{\Gamma}\frac{T_{\rm eff}(\theta)}{T_0}\frac{\partial}{\partial\varpi}\right]\tilde{M}^{(0)}[P(\theta');\tilde{\mathcal{T}}](\theta,\varpi), \quad \text{(A23)}$$

which has a solution

$$\tilde{M}^{(0)}[P(\theta');\tilde{\mathcal{T}}](\theta,\varpi) = P(\theta)\frac{\exp\left(-\frac{T_0\varpi^2}{2T(\theta)}\right)}{\sqrt{2\pi T(\theta)}} + [\text{exponentially decaying terms}]. \quad \text{(A24)}$$

Since $\tilde{\Omega}^{(0)}[P(\theta');\tilde{\mathcal{T}}](\theta)$ vanishes again, we proceed to the subleading order of Eq. (A19),

$$\frac{\partial\tilde{M}^{(1)}[P(\theta');\tilde{\mathcal{T}}](\theta,\varpi)}{\partial\tilde{\mathcal{T}}} = -\tau\frac{\partial}{\partial\theta}\left(\sqrt{\frac{T_0}{m}}\varpi\tilde{M}^{(0)}[P(\theta');\tilde{\mathcal{T}}](\theta,\varpi)\right) - \tau\frac{\partial}{\partial\varpi}\left[\frac{1}{\sqrt{mT_0}}\left(-\frac{\partial U_{\rm eff}(\theta)}{\partial\theta}+f\right)\tilde{M}^{(0)}[P(\theta');\tilde{\mathcal{T}}](\theta,\varpi)\right]$$
$$-\frac{\partial}{\partial\varpi}\left[-\frac{\mathcal{G}(\theta)}{\Gamma}\varpi - \frac{\mathcal{G}(\theta)}{\Gamma}\frac{T_{\rm eff}(\theta)}{T_0}\frac{\partial}{\partial\varpi}\right]\tilde{M}^{(1)}[P(\theta');\tilde{\mathcal{T}}](\theta,\varpi), \quad \text{(A25)}$$

which has a particular solution

$$\tilde{M}^{(1)}[P(\theta');\tilde{\mathcal{T}}](\theta,\varpi) = \left\{-\frac{\partial P(\theta)}{\partial\theta} - \left[\left(\frac{T_0\varpi^2}{6T_{\rm eff}(\theta)}+\frac{1}{2}\right)\frac{T'_{\rm eff}(\theta)}{T_{\rm eff}(\theta)} + \frac{1}{T_{\rm eff}(\theta)}\left(\frac{\partial U_{\rm eff}(\theta)}{\partial\theta} - f\right)\right]P(\theta)\right\}$$
$$\cdot\tau\sqrt{\frac{T_0}{m}}\frac{\Gamma}{\mathcal{G}(\theta)}\varpi\frac{\exp(-T_0\varpi^2/2T(\theta))}{\sqrt{2\pi T(\theta)}} + [\text{exponentially decaying terms}]. \quad \text{(A26)}$$

By substituting Eq. (A26) into Eq. (A21),

$$\tilde{\Omega}^{(1)}[P(\theta');\tilde{\mathcal{T}}](\theta) := -\tau\int d\varpi\,\frac{\partial}{\partial\theta}\left(\sqrt{\frac{T_0}{m}}\varpi\,\tilde{M}^{(1)}[P(\theta');\tilde{\mathcal{T}}](\theta,\varpi)\right)$$
$$= -\tau\frac{\partial}{\partial\theta}\left\{\frac{1}{\mathcal{G}(\theta)}\left[\left(-\frac{\partial U_{\rm eff}(\theta)}{\partial\theta}+f\right)P(\theta)-\frac{\partial}{\partial\theta}[T_{\rm eff}(\theta)P(\theta)]\right]\right\}. \quad \text{(A27)}$$

We finally reach

$$\frac{\partial P(\theta)}{\partial t} = -\frac{\partial}{\partial\theta}\left\{\frac{1}{\mathcal{G}(\theta)}\left[\left(-\frac{\partial U_{\rm eff}(\theta)}{\partial\theta}+f\right)P(\theta)-\frac{\partial}{\partial\theta}[T_{\rm eff}(\theta)P(\theta)]\right]\right\}. \quad \text{(A28)}$$

In order to obtain the overdamped Langevin equation corresponding to Eq. (A28), we first recall that [56]

$$\dot{X} = A(X)+C(X)\circ\tilde{\Xi} \quad \text{(A29)}$$

can be mapped to an additive Langevin equation

$$\dot{\bar{X}} = \bar{A}(\bar{X})+\tilde{\Xi}, \quad \text{(A30)}$$

where $\tilde{\Xi}$ is a white Gaussian noise with zero mean and unit variance, and $\bar{X}=\int^X dX/C(X)$, $\bar{A}(\bar{X})=A(X)/C(X)$. Since Eq. (A30) has a corresponding Fokker-Planck equation

$$\frac{\partial P(\bar{X})}{\partial t} = -\frac{\partial}{\partial\bar{X}}\left(\bar{A}(\bar{X})P(\bar{X})-\frac{1}{2}\frac{\partial P(\bar{X})}{\partial\bar{X}}\right), \quad \text{(A31)}$$

we obtain the Fokker-Planck equation for $P(X)$ through variable transformation [note that $P(\bar{X})=P(X)C(X)$]:

$$\frac{\partial P(X)}{\partial t} = -\frac{\partial}{\partial X}\left(A(X)P(X)-\frac{C(X)}{2}\frac{\partial}{\partial X}[C(X)P(X)]\right). \quad \text{(A32)}$$

Therefore, by rewriting Eq. (A28) in the form of Eq. (A32), the Langevin equation corresponding to Eq. (A28) is obtained as

$$\dot{\theta} = \frac{1}{\mathcal{G}(\theta)}\left(-\frac{\partial U_{\rm eff}(\theta)}{\partial\theta}+f\right)-\frac{1}{2\mathcal{G}(\theta)^2}\frac{\partial}{\partial\theta}[\mathcal{G}(\theta)T_{\rm eff}(\theta)]+\sqrt{2\frac{T_{\rm eff}(\theta)}{\mathcal{G}(\theta)}}\circ\tilde{\Xi}. \quad \text{(A33)}$$

By changing the product, we have

$$\dot{\theta} = \frac{1}{\mathcal{G}(\theta)}\left(-\frac{\partial U_{\rm eff}(\theta)}{\partial\theta}+f\right)+\sqrt{\frac{2}{\mathcal{G}(\theta)}}\odot\sqrt{T_{\rm eff}(\theta)}\cdot\tilde{\Xi}. \quad \text{(A34)}$$

Multiplying $\mathcal{G}(\theta)$ in the sense of anti-Itô to both sides, we obtain

$$\mathcal{G}(\theta)\odot\dot{\theta} = \left(-\frac{\partial U_{\rm eff}(\theta)}{\partial\theta}+f\right)+\sqrt{2\mathcal{G}(\theta)}\odot\sqrt{T_{\rm eff}(\theta)}\cdot\tilde{\Xi}. \quad \text{(A35)}$$

4. Derivation of model 5

We start by rescaling the variables in Eq. (19) by Eq. (A1):

$$\tau_x^{-1}\frac{\partial P(\theta,s)}{\partial\mathcal{T}} = -\frac{\partial}{\partial\theta}\left[\frac{f}{\Gamma}P(\theta,s)-\frac{T_h}{\Gamma}\frac{\partial}{\partial\theta}P(\theta,s)\right]+\tau_x^{-1/2}\sqrt{\frac{\gamma}{T_c}}\frac{\partial}{\partial s}\left[\phi'(\theta)\left(\frac{f}{\Gamma}P(\theta,s)-\frac{T_h}{\Gamma}\frac{\partial}{\partial\theta}P(\theta,s)\right)\right]$$
$$-\tau_x^{-1/2}\sqrt{\frac{\gamma}{T_c}}\frac{\partial}{\partial\theta}\left[\frac{\phi'(\theta)}{\Gamma}\frac{\partial U_I(s)}{\partial s}P(\theta,s)+\frac{\phi'(\theta)}{\Gamma}T_h\frac{\partial}{\partial s}P(\theta,s)\right]$$
$$+\tau_x^{-1/2}\sqrt{\frac{\gamma}{T_c}}\frac{\partial}{\partial s}\left\{\frac{1}{\gamma}\left[\frac{\partial U_0(\phi(\theta))}{\partial\phi(\theta)}+O\left(\frac{\tau_x}{\tau}\right)\right]P(\theta,s)\right\}$$
$$-\tau_x^{-1}\frac{1}{\Gamma T_c}\frac{\partial}{\partial s}\left[-(\Gamma+\gamma\phi'(\theta)^2)\frac{\partial U_I(s)}{\partial s}P(\theta,s)-(\Gamma T_c+\gamma\phi'(\theta)^2T_h)\frac{\partial}{\partial s}P(\theta,s)\right]. \quad \text{(A36)}$$

The first term on the right-hand side of Eq. (A36) is $O(\tau^{-1})$, the last line is $O(\tau_x^{-1})$, and the remaining terms are $O(\tau^{-1/2}\tau_x^{-1/2})$, respectively. Again, following the procedure in Appendix A 1, we define

$$\hat{M}[P(\theta');\mathcal{T}](\theta,s) := P(\theta,s), \tag{A37}$$

$$\hat{\Omega}[P(\theta');\mathcal{T}](\theta) := -\tau_x \frac{\partial}{\partial\theta}\left[\frac{f}{\Gamma}P(\theta) - \frac{T_h}{\Gamma}\frac{\partial}{\partial\theta}P(\theta)\right] - \tau_x^{1/2}\sqrt{\frac{\gamma}{T_c}}\frac{\partial}{\partial\theta}\left[\int ds \frac{\phi'(\theta)}{\Gamma}\frac{\partial U_I(s)}{\partial s}\hat{M}[P(\theta');\mathcal{T}](\theta,s)\right]. \tag{A38}$$

Now, we apply the standard perturbation theory to Eq. (A36) expressed in terms of $\hat{M}$ and $\hat{\Omega}$, with a small parameter $\varepsilon' = \tau_x/\tau$. The leading order of Eq. (A36) gives

$$\frac{\partial \hat{M}^{(0)}}{\partial \mathcal{T}} = -\frac{1}{\Gamma T_c}\frac{\partial}{\partial s}\left[-\mathcal{G}(\theta)\frac{\partial U_I(s)}{\partial s}\hat{M}^{(0)} - \mathcal{G}(\theta)T_s(\theta)\frac{\partial}{\partial s}\hat{M}^{(0)}\right], \tag{A39}$$

where

$$T_s(\theta) = \frac{T_h\gamma\phi'(\theta)^2 + T_c\Gamma}{\Gamma + \gamma\phi'(\theta)^2}. \tag{A40}$$

We obtain

$$\hat{M}^{(0)} = P(\theta)\frac{\exp\left(-\frac{U_I(s)}{T_s(\theta)}\right)}{z(\theta)} + [\text{exponentially decaying terms}], \tag{A41}$$

with

$$z(\theta) = \int ds \exp\left(-\frac{U_I(s)}{T_s(\theta)}\right). \tag{A42}$$

In the time scale of τ, the $O(\varepsilon'^{1/2})$ term of $\hat{\Omega}$ vanishes. The subleading order of Eq. (A36) becomes

$$\begin{aligned}\frac{\partial \hat{M}^{(1)}}{\partial \mathcal{T}} &= \tau^{-1/2}\sqrt{\frac{\gamma}{T_c}}\frac{\partial}{\partial s}\left[\phi'(\theta)\left(\frac{f}{\Gamma}\hat{M}^{(0)} - \frac{T_h}{\Gamma}\frac{\partial}{\partial\theta}\hat{M}^{(0)}\right)\right] - \tau^{-1/2}\sqrt{\frac{\gamma}{T_c}}\frac{\partial}{\partial\theta}\left[\frac{\phi'(\theta)}{\Gamma}\frac{\partial U_I(s)}{\partial s}\hat{M}^{(0)} + \frac{\phi'(\theta)}{\Gamma}T_h\frac{\partial}{\partial s}\hat{M}^{(0)}\right]\\ &+\tau^{-1/2}\sqrt{\frac{\gamma}{T_c}}\frac{\partial}{\partial s}\left[\frac{1}{\gamma}\frac{\partial U_0(\phi(\theta))}{\partial\phi(\theta)}\hat{M}^{(0)}\right] - \frac{1}{\Gamma T_c}\frac{\partial}{\partial s}\left[-\mathcal{G}(\theta)\frac{\partial U_I(s)}{\partial s}\hat{M}^{(1)} - \mathcal{G}(\theta)T_s(\theta)\frac{\partial}{\partial s}\hat{M}^{(1)}\right],\end{aligned} \tag{A43}$$

which gives

$$\begin{aligned}\hat{M}^{(1)} &\propto \frac{-s\hat{M}^{(0)}}{\mathcal{G}(\theta)T_s(\theta)}\left\{-\frac{\partial}{\partial\theta}\left[\frac{\phi'(\theta)}{\Gamma}(T_h - T_s(\theta))\right] + \frac{f}{\Gamma}\phi'(\theta) + \frac{1}{\gamma}\frac{\partial U_0(\phi(\theta))}{\partial\phi(\theta)}\right\} - \frac{\phi'(\theta)}{\Gamma}[2T_h - T_s(\theta)]\frac{1}{\mathcal{G}(\theta)T_s(\theta)}\\ &\times\left[\frac{\partial}{\partial\theta}(s\hat{M}^{(0)}) - \frac{1}{T_s(\theta)^2}\frac{\partial T_s(\theta)}{\partial\theta}\mathcal{I}(s)\hat{M}^{(0)}\right] + [\text{exponentially decaying terms}],\end{aligned} \tag{A44}$$

where

$$\mathcal{I}(s) = sU_I(s) - \int^s ds' U_I(s'). \tag{A45}$$

Substitution of Eq. (A44) into Eq. (A38) results in

$$\begin{aligned}\hat{\Omega}^{(1)}[P(\theta);\mathcal{T}] &:= -\tau\frac{\partial}{\partial\theta}\left[\frac{f}{\Gamma}P(\theta) - \frac{T_h}{\Gamma}\frac{\partial}{\partial\theta}P(\theta)\right] - \tau^{1/2}\sqrt{\frac{\gamma}{T_c}}\frac{\partial}{\partial\theta}\left[\int ds\frac{\phi'(\theta)}{\Gamma}\frac{\partial U_I(s)}{\partial s}\hat{M}^{(1)}[P(\theta');\mathcal{T}](\theta,s)\right]\\ &= -\tau\frac{\partial}{\partial\theta}\left[\frac{f}{\mathcal{G}(\theta)}P(\theta) + \frac{\partial}{\partial\theta}\left(\frac{\gamma\phi'(\theta)^2}{2\mathcal{G}(\theta)^2}\right)(T_h - T_c)P(\theta) - \frac{1}{\mathcal{G}(\theta)}\frac{\partial U_{\text{eff}}(\theta)}{\partial\theta}P(\theta) - \frac{T_{\text{eff}}(\theta)}{\mathcal{G}(\theta)}\frac{\partial}{\partial\theta}P(\theta)\right],\end{aligned} \tag{A46}$$

which gives the Fokker-Planck equation

$$\frac{\partial P(\theta)}{\partial t} = -\frac{\partial}{\partial\theta}\left[-\frac{1}{\mathcal{G}(\theta)}\frac{\partial \ln\mathcal{G}(\theta)}{\partial\theta}(T_h - T_c)P(\theta) + \frac{1}{\mathcal{G}(\theta)}\left(f - \frac{\partial U_{\text{eff}}(\theta)}{\partial\theta}\right)P(\theta) - \frac{1}{\mathcal{G}(\theta)}\frac{\partial}{\partial\theta}[T_{\text{eff}}(\theta)P(\theta)]\right], \tag{A47}$$

corresponding to model 5.

APPENDIX B: ASYMPTOTIC BEHAVIOR OF ENTROPY PRODUCTION RATES

1. Transition probabilities in Eqs. (21)–(25)

$$W_1(\theta_{t'},\Pi_{t'},x_{t'}|\theta_t,\Pi_t,x_t) = \delta\left(\theta_{t'}-\theta_t-\frac{\Pi_t+\Pi_{t'}}{2m}(t'-t)\right)W_{UD}\left(\Pi_{t'}\middle|\Pi_t;-\frac{\partial U(\bar\theta,\bar x)}{\partial\bar\theta}+f,\Gamma,T_h\right)\frac{1}{\sqrt{4\pi(t'-t)T_c/\gamma}} \times \exp\left(-\frac{[\gamma(x_{t'}-x_t)+\partial U(\bar\theta,\bar x)/\partial\bar x(t'-t)]^2}{4\gamma T_c(t'-t)}+\frac{\partial^2U(\bar\theta,\bar x)}{\partial\bar x^2}\frac{(t'-t)}{2\gamma}\right), \tag{B1}$$

$$W_2(\theta_{t'},\Pi_{t'}|\theta_t,\Pi_t) = \delta\left(\theta_{t'}-\theta_t-\frac{\Pi_t+\Pi_{t'}}{2m}(t'-t)\right)W_{UD}\left(\Pi_{t'}\middle|\Pi_t;-\frac{\partial U_{\rm eff}(\bar\theta)}{\partial\bar\theta}+f,\mathcal{G}(\bar\theta),T_{\rm eff}(\bar\theta)\right), \tag{B2}$$

$$W_3(\theta_{t'}|\theta_t) = \frac{1}{\sqrt{4\pi(t'-t)T_{\rm eff}(\bar\theta)/\mathcal{G}(\bar\theta)}}\exp\left\{-\frac{\{\mathcal{G}(\bar\theta)(\theta_{t'}-\theta_t)+[\partial(U_{\rm eff}(\bar\theta)+T_{\rm eff}(\bar\theta))/\partial\bar\theta-f](t'-t)\}^2}{4\mathcal{G}(\bar\theta)T_{\rm eff}(\bar\theta)(t'-t)} +\frac{1}{2}\frac{\partial}{\partial\bar\theta}\left[\frac{1}{\mathcal{G}(\bar\theta)}\left(\frac{\partial U_{\rm eff}(\bar\theta)}{\partial\bar\theta}-f\right)+\frac{T_{\rm eff}(\bar\theta)}{\mathcal{G}(\bar\theta)^2}\frac{\partial\mathcal{G}(\bar\theta)}{\partial\bar\theta}+\frac{1}{2}\frac{\partial}{\partial\bar\theta}\frac{T_{\rm eff}(\bar\theta)}{\mathcal{G}(\bar\theta)}\right](t'-t)\right\}, \tag{B3}$$

$$W_4(\theta_{t'},x_{t'}|\theta_t,x_t) = \frac{1}{\sqrt{4\pi(t'-t)T_h/\Gamma}}\exp\left(-\frac{[\Gamma(\theta_{t'}-\theta_t)+(\partial U(\bar\theta,\bar x)/\partial\bar\theta-f)(t'-t)]^2}{4\Gamma T_h(t'-t)}+\frac{\partial^2U(\bar\theta,\bar x)}{\partial\bar\theta^2}\frac{(t'-t)}{2\Gamma}\right) \times\frac{1}{\sqrt{4\pi(t'-t)T_c/\gamma}}\exp\left(-\frac{[\gamma(x_{t'}-x_t)+\partial U(\bar\theta,\bar x)/\partial\bar x(t'-t)]^2}{4\gamma T_c(t'-t)}+\frac{\partial^2U(\bar\theta,\bar x)}{\partial\bar x^2}\frac{(t'-t)}{2\gamma}\right), \tag{B4}$$

$$W_5(\theta_{t'}|\theta_t) = \frac{1}{\sqrt{4\pi(t'-t)T_{\rm eff}(\bar\theta)/\mathcal{G}(\bar\theta)}}\exp\left\{-\frac{\{\mathcal{G}(\bar\theta)(\theta_{t'}-\theta_t)+[\partial[U_{\rm eff}(\bar\theta)+T_{\rm eff}(\bar\theta)]+\ln\mathcal{G}(\bar\theta)(T_h-T_c)/\partial\bar\theta-f](t'-t)\}^2}{4\mathcal{G}(\bar\theta)T_{\rm eff}(\bar\theta)(t'-t)} +\frac{1}{2}\frac{\partial}{\partial\bar\theta}\left[\frac{1}{\mathcal{G}(\bar\theta)}\left(\frac{\partial U_{\rm eff}(\bar\theta)}{\partial\bar\theta}+\frac{\partial\ln\mathcal{G}(\bar\theta)}{\partial\bar\theta}(T_h-T_c)-f\right)+\frac{T_{\rm eff}(\bar\theta)}{\mathcal{G}(\bar\theta)^2}\frac{\partial\mathcal{G}(\bar\theta)}{\partial\bar\theta}+\frac{1}{2}\frac{\partial}{\partial\bar\theta}\frac{T_{\rm eff}(\bar\theta)}{\mathcal{G}(\bar\theta)}\right](t'-t)\right\}, \tag{B5}$$

where $\bar\theta=(\theta_t+\theta_{t'})/2,\bar x=(x_t+x_{t'})/2$, and

$$W_{UD}(\Pi_{t'}|\Pi_t;F,g,T) = \frac{1}{\sqrt{4\pi(t'-t)T/g}}\exp\left(-\frac{\left[\Pi_{t'}-\Pi_t+\left(\frac{g}{m}\frac{\Pi_t+\Pi_{t'}}{2}-F\right)(t'-t)\right]^2}{4gT(t'-t)}\right), \tag{B6}$$

is the transition probability of momentum degree of freedom following the underdamped Langevin equation.

2. Derivation of Eq. (31)

Based on the results of Appendix A, we evaluate the ensemble average of the entropy production rate $\langle\sigma_1\rangle$ in the limit of $\varepsilon\to 0$. Since $\langle Q_1^h\rangle$ may be rewritten as

$$\langle Q_1^h\rangle = \left\langle-\left(\dot\Pi+\frac{\partial U(\theta,x)}{\partial\theta}-f\right)\circ\frac{\Pi}{m}\right\rangle = \left\langle\frac{\Gamma}{m}\left(\frac{\Pi^2}{m}-T_h\right)\right\rangle, \tag{B7}$$

we consider the ensemble average of Q_1^c with respect to $M[P(\theta',\Pi');\mathcal{T}](\theta,\Pi,s)$, by replacing x in Q_1^c by $\phi(\theta)+L_xs$:

$$\begin{aligned} Q_1^c &= -\frac{\partial U(\theta,x)}{\partial x}\circ\dot x \\ &= -\left(\frac{1}{L_x}\frac{\partial U_I(s)}{\partial s}+\frac{\partial U_0(\phi(\theta))}{\partial\phi(\theta)}+O\left(\frac{L_x}{L^2}\right)\right)\circ[\phi'(\theta)\dot\theta+L_x\dot s] \\ &= \left(-\frac{1}{L_x}\frac{\partial U_I(s)}{\partial s}-\frac{\partial U_0(\phi(\theta))}{\partial\phi(\theta)}\right)\phi'(\theta)\frac{\Pi}{m}-\left(\frac{\partial U_I(s)}{\partial s}+\frac{\partial U_0(\phi(\theta))}{\partial\phi(\theta)}L_x+\frac{\partial^2U_0(\phi(\theta))}{\partial\phi(\theta)^2}L_x^2s+O(\varepsilon^{3/2})\right)\circ\dot s+O\left(\frac{\varepsilon}{\tau_\Pi}\right). \end{aligned} \tag{B8}$$

The ensemble average of the first term of Eq. (B8) is [up to $O(\tau_\Pi^{-1})$],

$$\left\langle-\frac{\phi'(\theta)}{L_x}\frac{\partial U_I(s)}{\partial s}\frac{\Pi}{m}\right\rangle \simeq \int d\theta d\Pi ds\left[-\frac{\phi'(\theta)}{L_x}\frac{\Pi}{m}\frac{\partial U_I(s)}{\partial s}M^{(0)}[P(\theta',\Pi');\mathcal{T}](\theta,\Pi,s) -\phi'(\theta)\frac{\Pi}{m}\frac{\partial U_I(s)}{\partial s}\sqrt{\frac{\gamma}{T_c}\frac{\varepsilon}{\tau_x}}M^{(1)}[P(\theta',\Pi');\mathcal{T}](\theta,\Pi,s)\right]$$

$$= -\int d\theta d\Pi ds \phi'(\theta)\frac{\Pi}{m}\frac{1}{T_c}\frac{\partial U_I(s)}{\partial s}s\frac{\exp[-U_I(s)/T_c]}{Z}\left[-\gamma\phi'(\theta)\left(\frac{\Pi}{m}+T_c\frac{\partial}{\partial \Pi}\right)-U_0'[\phi(\theta)]\right]P(\theta,\Pi)$$
$$= \left\langle \frac{\gamma\phi'(\theta)^2}{m}\left(\frac{\Pi^2}{m}-T_c\right)+\frac{\partial U_0[\phi(\theta)]}{\partial\theta}\frac{\Pi}{m}\right\rangle. \tag{B9}$$

Since $U_I(s)$ and T_c are fixed, the second term on the right hand side of Eq. (B8) vanishes. Putting these altogether, we obtain

$$\langle\sigma_1\rangle = \left\langle \frac{1}{T_h}\frac{\Gamma}{m}\left(\frac{\Pi^2}{m}-T_h\right)\right\rangle + \left\langle \frac{1}{T_c}\frac{\gamma\phi'(\theta)^2}{m}\left(\frac{\Pi^2}{m}-T_c\right)\right\rangle. \tag{B10}$$

By comparing this with

$$\langle\sigma_2\rangle = \left\langle \frac{1}{T_{\rm eff}(\theta)}\frac{\mathcal{G}(\theta)}{m}\left(\frac{\Pi^2}{m}-T_{\rm eff}(\theta)\right)\right\rangle, \tag{B11}$$

we obtain Eq. (31).

3. Derivation of Eq. (32)

Next, we evaluate the ensemble average of the entropy production rate $\langle\sigma_2\rangle$ in the limit of $\epsilon\to 0$. The entropy production rate may be rewritten as

$$\sigma_2 = \frac{1}{T_{\rm eff}(\theta)}\left(\dot\Pi+\frac{\partial U_0(\phi(\theta))}{\partial\theta}-f\right)\circ\frac{\Pi}{m} = -\frac{1}{T_{\rm eff}(\theta)}\left[\frac{d}{dt}\frac{\Pi^2}{2m}+\left(\frac{\partial U_0(\phi(\theta))}{\partial\theta}-f\right)\frac{\Pi}{m}\right]$$
$$= -\frac{d}{dt}\left(\frac{1}{T_{\rm eff}(\theta)}\frac{\Pi^2}{2m}\right)-\frac{\Pi^2}{2m}\frac{1}{T_{\rm eff}(\theta)^2}\frac{\Pi}{m}\frac{\partial}{\partial\theta}T_{\rm eff}(\theta)-\frac{1}{T_{\rm eff}(\theta)}\left(\frac{\partial U_0(\phi(\theta))}{\partial\theta}-f\right)\frac{\Pi}{m}, \tag{B12}$$

where we use that $T_{\rm eff}(\theta)$ does not depend on time explicitly. Since it immediately follows from the oddness of Eq. (B12) as the function of Π that the ensemble average with respect to $\tilde M^{(0)}[P(\theta');\tilde{\mathcal{T}}](\theta,\varpi)$ vanishes, we obtain a finite contribution from that with respect to $\tilde M^{(1)}[P(\theta');\tilde{\mathcal{T}}](\theta,\varpi)$ as

$$\langle\sigma_2\rangle = \int d\theta\frac{1}{\mathcal{G}(\theta)T_{\rm eff}(\theta)}\left\{\left[-T_{\rm eff}(\theta)\frac{\partial P(\theta)}{\partial\theta}+\left(-\frac{\partial U_0(\phi(\theta))}{\partial\theta}-T_{\rm eff}'(\theta)+f\right)P(\theta)\right]\right.$$
$$\left.\times\left(-\frac{\partial U_0(\phi(\theta))}{\partial\theta}-\frac{3}{2}T_{\rm eff}'(\theta)+f\right)+\frac{1}{2}T_{\rm eff}'(\theta)^2P(\theta)\right\}$$
$$= \left\langle\frac{1}{T_{\rm eff}(\theta)}\left(-\frac{\partial U_0(\phi(\theta))}{\partial\theta}-\frac{3}{2}T_{\rm eff}'(\theta)+f\right)\circ\dot\theta+\frac{T_{\rm eff}(\theta)}{2\mathcal{G}(\theta)}\left(\frac{T_{\rm eff}'(\theta)}{T_{\rm eff}(\theta)}\right)^2\right\rangle$$
$$= \langle\sigma_3\rangle+\left\langle\frac{T_{\rm eff}(\theta)}{2\mathcal{G}(\theta)}\left(\frac{T_{\rm eff}'(\theta)}{T_{\rm eff}(\theta)}\right)^2\right\rangle. \tag{B13}$$

The third line is obtained by using the overdamped Langevin equation of model 3.

4. Derivation of Eq. (31) based on coarse graining in Sec. III B

Here, we present a different coarse-graining method based on temporal coarse graining, which does not involve the ensemble average. We first substitute Eq. (A15) into the expression of σ_1,

$$\sigma_1 = -\frac{1}{T_h}(\dot\Pi-\lambda\phi'(\theta)[x-\phi(\theta)])\circ\frac{\Pi}{m}-\frac{1}{T_c}(\lambda+k)\left(x-\frac{\lambda}{\lambda+k}\phi(\theta)\right)\circ\dot x$$
$$= -\frac{1}{T_h}\left(\dot\Pi-k\frac{\lambda}{k+\lambda}\phi'(\theta)\phi(\theta)\right)\circ\frac{\Pi}{m}+\frac{1}{T_h}\lambda\phi'(\theta)\left(-\frac{\gamma\lambda}{(k+\lambda)^2}\phi'(\theta)\frac{\Pi}{m}+\frac{\sqrt{2\gamma T_c}}{k+\lambda}\tilde\zeta^c\right)\circ\frac{\Pi}{m}$$
$$-\frac{1}{T_c}(\lambda+k)\left(-\frac{\gamma\lambda}{(k+\lambda)^2}\phi'(\theta)\frac{\Pi}{m}+\frac{\sqrt{2\gamma T_c}}{k+\lambda}\tilde\zeta^c\right)\circ\left[\frac{\lambda}{k+\lambda}\phi'(\theta)\frac{\Pi}{m}+\frac{\sqrt{2\gamma T_c}}{\gamma}(\tilde\zeta-\tilde\zeta^c)\right], \tag{B14}$$

where $\tilde{\zeta}^c = \int_{-\infty}^{t} e^{-(k+\lambda)/\gamma(t-t')}\tilde{\zeta}_{t'}$. Integrating this over $t \in [t_0, t_0+\Delta t]$ and neglecting the higher order terms in the limit of $\varepsilon \to 0$ as done in Appendix A 2, we obtain

$$\int_{t_0}^{t_0+\Delta t} dt\sigma_1 = \int_{t_0}^{t_0+\Delta t} dt\Bigg\{-\frac{1}{T_c}\phi'(\theta)\Big(-\gamma\phi'(\theta)\frac{\Pi}{m} + \sqrt{2\gamma T_c}\,\Xi\Big) - \frac{1}{T_h}\Big[\dot{\Pi} - k\phi'(\theta)\phi(\theta) + \phi'(\theta)\Big(\gamma\phi'(\theta)\frac{\Pi}{m} - \sqrt{2\gamma T_c}\,\Xi\Big)\Big]\Bigg\} \circ \frac{\Pi}{m}. \tag{B15}$$

Here, we use $\int_{t_0}^{t_0+\Delta t} dt\tilde{\zeta}^c(\tilde{\zeta} - \tilde{\zeta}^c) \to 0$ in the sense of the convergence in mean square. The ensemble average of Eq. (B15) is the same as the right-hand side of Eq. (31).

5. Derivation of Eq. (38)

We here show that $\langle\sigma_4\rangle$ is dominated by the $O(\tau_x^{-1})$ terms in the tightly confined limit. We rewrite $\langle\sigma_4\rangle$ in terms of s,

$$\begin{aligned}
\langle\sigma_4\rangle &= \frac{1}{T_h}\left\langle\left(f - \frac{\partial U(\theta,x)}{\partial\theta}\right)\circ\dot{\theta}\right\rangle - \frac{1}{T_c}\left\langle\frac{\partial U(\theta,x)}{\partial x}\circ\dot{x}\right\rangle \\
&= \frac{1}{\Gamma T_h}\left\langle-\left(f - \frac{\partial U(\theta,x)}{\partial\theta}\right)\frac{\partial U(\theta,x)}{\partial\theta} - T_h\frac{\partial^2 U(\theta,x)}{\partial\theta^2}\right\rangle - \frac{1}{\gamma T_c}\left\langle-\left(\frac{\partial U(\theta,x)}{\partial x}\right)^2 + T_c\frac{\partial^2 U(\theta,x)}{\partial x^2}\right\rangle \\
&= \left\langle\frac{\mathcal{G}(\theta)T_{\rm eff}(\theta)}{\Gamma\gamma T_h T_c}\left(\frac{1}{L_x}\frac{\partial U_I(s)}{\partial s}\right)^2 - \frac{\mathcal{G}(\theta)}{\Gamma\gamma}\frac{1}{L_x^2}\frac{\partial^2 U_I(s)}{\partial s^2}\right\rangle \\
&= \left\langle\frac{\mathcal{G}(\theta)}{\Gamma\gamma}\left(\frac{T_{\rm eff}(\theta)}{T_h T_c} - \frac{1}{T_s(\theta)}\right)\left(\frac{1}{L_x}\frac{\partial U_I(s)}{\partial s}\right)^2\right\rangle + \left\langle\frac{\mathcal{G}(\theta)}{\Gamma\gamma}\left[\frac{1}{T_s(\theta)}\left(\frac{1}{L_x}\frac{\partial U_I(s)}{\partial s}\right)^2 - \frac{1}{L_x^2}\frac{\partial^2 U_I(s)}{\partial s^2}\right]\right\rangle + O(\tau^{-1}).
\end{aligned} \tag{B16}$$

The second ensemble average in the last line (with respect to $\hat{M}^{(0)} + \varepsilon'\hat{M}^{(1)}$) is smaller than $O(\tau_x^{-1})$. Therefore, we obtain Eq. (38) as the leading term.

APPENDIX C: DERIVATION OF EQS. (45) AND (46)

We calculate the forward and backward transition rates of model 3 in the limit of $\Delta U_{\rm eff}/T_{\rm eff}(\theta) \to \infty$. In the case of model 3, we may obtain an additive Langevin equation from Eq. (A33) by transforming the variable from θ to q as

$$\dot{q} = -\frac{\partial\psi(q)}{\partial q} + \frac{1}{2}\frac{\partial}{\partial q}\ln\left(\frac{T_{\rm eff}(q)}{\mathcal{G}(q)}\right) + \sqrt{2}\tilde{\Xi}, \tag{C1}$$

where

$$q := \int^{\theta}\sqrt{\frac{\mathcal{G}(\theta')}{T_{\rm eff}(\theta')}}d\theta', \tag{C2}$$

and ψ [defined in Eq. (16)], $T_{\rm eff}$, and $\mathcal{G}$ are regarded as functions of q. By applying Kramers theory [57] to Eq. (C1), the forward and backward transition rates are given as

$$\begin{aligned}
R_{(3)}^{F,B} &= \frac{1}{2\pi}\sqrt{-\left.\frac{\partial^2\tilde{\psi}(q)}{\partial q^2}\right|_{q=q_{\rm min}}\left.\frac{\partial^2\tilde{\psi}(q)}{\partial q^2}\right|_{q=q_{\rm max}}}\exp[-\tilde{\psi}(q_{\rm max}) + \tilde{\psi}(q_{\rm min})] \\
&= \frac{1}{2\pi}\sqrt{-\left.\frac{\partial^2\tilde{\psi}(q)}{\partial q^2}\right|_{q=q_{\rm min}}\left.\frac{\partial^2\tilde{\psi}(q)}{\partial q^2}\right|_{q=q_{\rm max}}}\sqrt{\frac{\mathcal{G}(q_{\rm min})T_{\rm eff}(q_{\rm max})}{T_{\rm eff}(q_{\rm min})\mathcal{G}(q_{\rm max})}}\exp\left[-\psi(q_{\rm max}) + \psi(q_{\rm min})\right] \propto \exp\left[-\psi(q_{\rm max}) + \psi(q_{\rm min})\right],
\end{aligned} \tag{C3}$$

where $\tilde{\psi}(q) = \psi(q) + \frac{1}{2}[\ln\mathcal{G}(q) - \ln T_{\rm eff}(q)]$, $q_{\rm min}$ is a local minimum of $\tilde{\psi}(q)$, and $q_{\rm max}$ is the nearest local maximum of $\tilde{\psi}(q)$ so that $q_{\rm max} > q_{\rm min}$ for $R_{(3)}^F$ and $q_{\rm max} < q_{\rm min}$ for $R_{(3)}^B$. Since, in the limit of $\Delta U_{\rm eff}/T_{\rm eff}(\theta) \to \infty$, the local maxima and minima of $\tilde{\psi}(q)$ agree with those of $U_{\rm eff}[\theta(q)]$, $\theta_{\rm max}$ and $\theta_{\rm min}$,

$$\begin{aligned}
R_{(3)}^{F,B} &\propto \exp\left[-\int_{\theta_{\rm min}}^{\theta_{\rm max}}\frac{1}{T_{\rm eff}(\theta)}\left(\frac{\partial U_{\rm eff}(\theta)}{\partial\theta} + \frac{\partial T_{\rm eff}(\theta)}{\partial\theta} - f\right)d\theta\right] \\
&= \frac{T_{\rm eff}(\theta_{\rm min})}{T_{\rm eff}(\theta_{\rm max})}\exp\left[-\int_{\theta_{\rm min}}^{\theta_{\rm max}}\frac{1}{T_{\rm eff}(\theta)}\left(\frac{\partial U_{\rm eff}(\theta)}{\partial\theta} - f\right)d\theta\right].
\end{aligned} \tag{C4}$$

By defining the common prefactor as τ_s^{-1}, we obtain

$$R_{(3)}^{F,B} = \tau_s^{-1} \exp\left[-\int_{\theta_{\min}}^{\theta_{\max}} \frac{1}{T_{\text{eff}}(\theta)}\left(\frac{\partial U_{\text{eff}}(\theta)}{\partial\theta} - f\right)d\theta\right]. \tag{C5}$$

We may estimate the transition rates of model 5 in the same manner.

APPENDIX D: DETAILS OF NUMERICAL SIMULATION

The numerical simulations are mainly carried out based on the Langevin equation of model 1. In the numerical integration of Langevin equation, we employ the velocity Verlet method for the underdamped part and the Euler method for the overdamped part. The time step is set to 2×10^{-3} and the total length of simulations is set to 2^{12}. The ensemble averages of the entropy production are calculated from 2^{12} independent runs, and the average entropy production rates are obtained from linear fitting.

In the numerical investigation of efficiency (Figs. 6 and 7), we use the numerical integration of the Kramers equation of model 2 together with the Langevin equation of model 1. The phase space with a cutoff of momentum at $\Pi = \pm 8$ is discretized into $2^8 \times (2^7 + 1)$ elements along the position and momentum axes, respectively. The derivatives with respect to θ or Π are approximated by the central difference. The time step is set to 0.056×10^{-5} and the total length of simulations is set to 2^3.

APPENDIX E: DEFINITION OF HEAT AND ITS EFFECT ON THERMODYNAMIC EFFICIENCY

In Sec. IV, we introduced heat flux by respecting the energy balance, and used them to discuss the thermodynamic efficiencies of the FS ratchet at the coarse-grained scales. We here remark on how these results will be affected when adopting a different definition for heat flux $\tilde{Q}_{3,5}$ which satisfies

$$\sigma_3(\theta_{t'}|\theta_t) = \frac{1}{T_{\text{eff}}(\theta)} \circ \tilde{Q}_3, \tag{E1}$$

$$\sigma_5(\theta_{t'}|\theta_t) = \frac{1}{T_{\text{eff}}(\theta)} \circ \tilde{Q}_5, \tag{E2}$$

while keeping the definitions of $\sigma_{3,5}$.

The difference between Q_3 [Eq. (29)] and $\tilde{Q}_3$ is

$$Q_3 - \tilde{Q}_3 = \frac{\partial T_{\text{eff}}(\theta)}{\partial\theta} \circ \dot{\theta}. \tag{E3}$$

By multiplying $\delta[T_{\text{eff}}(\theta) - T]$ in the sense of Stratonovich and taking the ensemble average, we obtain

$$\begin{aligned}\langle Q_3(T)\rangle - \langle\tilde{Q}_3(T)\rangle &= \left\langle \frac{\partial T_{\text{eff}}(\theta)}{\partial\theta}\delta[T_{\text{eff}}(\theta) - T] \circ \dot{\theta}\right\rangle \\ &= \int \frac{\partial T_{\text{eff}}(\theta)}{\partial\theta}\delta[T_{\text{eff}}(\theta) - T]J\,d\theta \\ &= \sum_j \frac{T'_{\text{eff}}(\theta_j)}{|T'_{\text{eff}}(\theta_j)|}J = 0,\end{aligned} \tag{E4}$$

where $\langle\tilde{Q}_3(T)\rangle := \langle\delta[T_{\text{eff}}(\theta) - T] \circ \tilde{Q}_3\rangle$, J is the probability current at the steady state, and θ_j are the angles satisfying $T_{\text{eff}}(\theta_j) = T$. [The stochastic product for $\langle Q_3(T)\rangle$ was not specified in [46].] Equation (E4) suggests that the average heat flux under the condition of $T_{\text{eff}}(\theta) = T$ is the same between the two definitions of heat flux, which means that the thermodynamic efficiency η_3 [Eq. (62)] is unaffected by the change from Q_3 to $\tilde{Q}_3$. The generalization to multidimensional cases is straightforward.

By the same argument, η_5 is independent on which heat flux (Q_5 or $\tilde{Q}_5$) is used.

APPENDIX F: DERIVATION OF EQS. (65) AND (68)

We derive Eqs. (65) and (68) based on the singular perturbation theory starting from a generalized version of model 6:

$$\begin{aligned}\frac{\partial P(\theta,\Pi,b)}{\partial t} &= -\frac{\partial}{\partial\theta}\left(\frac{\Pi}{m}P(\theta,\Pi,b)\right) \\ &\quad - \frac{\partial}{\partial\Pi}\left[\left(-\frac{\Gamma_b(\theta)}{m}\Pi - \frac{\partial U_{\text{eff}}(\theta)}{\partial\theta} + f - \Gamma_b(\theta)T_b\frac{\partial}{\partial\Pi}\right)P(\theta,\Pi,b)\right] \\ &\quad - \Lambda_{b\to b'}(\theta)P(\theta,\Pi,b) + \Lambda_{b'\to b}(\theta)P(\theta,\Pi,b').\end{aligned} \tag{F1}$$

The limit of fast switching is when $\tau_\Lambda = \max_\theta[\Lambda_{h\to c}(\theta) + \Lambda_{c\to h}(\theta)]^{-1}$ is separated from τ_Π and τ while the ratio τ_Π/τ is kept fixed. Under this condition, the first and second lines of the right-hand side of Eq. (F1) are $O(\tau^{-1/2}\tau_\Pi^{-1/2}) + O(\tau_\Pi^{-1})$, and the third line is $O(\tau_\Lambda^{-1})$. By introducing

$$\begin{aligned}\check{M}[P(\theta',\Pi');\mathscr{T}](\theta,\Pi,b) &:= P(\theta,\Pi,b)\\ \check{\Omega}[P(\theta',\Pi');\mathscr{T}](\theta,\Pi) &:= -\frac{\partial}{\partial\theta}\left(\frac{\Pi}{m}P(\theta,\Pi)\right) - \frac{\partial}{\partial\Pi}\left[\left(-\frac{\partial U_{\rm eff}(\theta)}{\partial\theta} + f\right)P(\theta,\Pi)\right]\\ &\quad - \frac{\partial}{\partial\Pi}\left[\sum_b\left(-\frac{\Gamma_b(\theta)}{m}\Pi - \Gamma_b(\theta)T_b\frac{\partial}{\partial\Pi}\right)\check{M}[P(\theta',\Pi');\mathscr{T}](\theta,\Pi,b)\right],\end{aligned} \tag{F2}$$

with $\mathscr{T} := t/\tau_\Lambda$, and expanding $\check{M}$ and $\check{\Omega}$ into series of $\delta = \tau_\Lambda/\tau_\Pi \sim \tau_\Lambda/\tau$, we obtain the leading order of Eq. (F1) expressed in terms of $\check{M}$ and $\check{\Omega}$,

$$\frac{\partial \check{M}^{(0)}[P(\theta',\Pi');\mathscr{T}](\theta,\Pi,b)}{\partial\mathscr{T}} = -\Lambda_{b\to b'}(\theta)\check{M}^{(0)}[P(\theta',\Pi');\mathscr{T}](\theta,\Pi,b) + \Lambda_{b'\to b}(\theta)\check{M}^{(0)}[P(\theta',\Pi');\mathscr{T}](\theta,\Pi,b'), \tag{F3}$$

and a solution

$$\check{M}^{(0)}[P(\theta',\Pi');\mathscr{T}](\theta,\Pi,b) = P(\theta,\Pi)\frac{\Lambda_{b'\to b}(\theta)}{\Lambda_{h\to c}(\theta)+\Lambda_{c\to h}(\theta)} + [\text{exponentially decaying terms}], \tag{F4}$$

which gives

$$\begin{aligned}\check{\Omega}^{(0)}[P(\theta',\Pi');\mathscr{T}](\theta,\Pi) &:= -\frac{\partial}{\partial\theta}\left(\frac{\Pi}{m}P(\theta,\Pi)\right) - \frac{\partial}{\partial\Pi}\left[\left(-\frac{\partial U_{\rm eff}(\theta)}{\partial\theta} + f\right)P(\theta,\Pi)\right]\\ &\quad - \frac{\partial}{\partial\Pi}\left[\sum_b\left(-\frac{\Gamma_b(\theta)}{m}\Pi - \Gamma_b(\theta)T_b\frac{\partial}{\partial\Pi}\right)\frac{\Lambda_{b'\to b}(\theta)}{\Lambda_{h\to c}(\theta)+\Lambda_{c\to h}(\theta)}P(\theta,\Pi)\right].\end{aligned} \tag{F5}$$

Therefore, Eq. (F1) will be equivalent to model 2 if $\Gamma_b(\theta)$ and the transition rates satisfy

$$\sum_b \Gamma_b(\theta)\frac{\Lambda_{b'\to b}(\theta)}{\Lambda_{h\to c}(\theta)+\Lambda_{c\to h}(\theta)} = \mathcal{G}(\theta), \tag{F6}$$

$$\sum_b \Gamma_b(\theta)T_b\frac{\Lambda_{b'\to b}(\theta)}{\Lambda_{h\to c}(\theta)+\Lambda_{c\to h}(\theta)} = \mathcal{G}(\theta)T_{\rm eff}(\theta). \tag{F7}$$

Model 6 satisfies this condition since $\Gamma_h(\theta) = 2\Gamma$, $\Gamma_c(\theta) = 2\gamma\phi'(\theta)^2$, and $\Lambda_{h\to c}(\theta) = \Lambda_{c\to h}(\theta) = \Lambda$.

The entropy production rate defined through the transition probability of Eq. (F1) is given as the sum of Eq. (66) and the logarithmic ratio of the rates of stochastic switching. The ensemble average of the latter contribution with respect to $\check{M}^{(0)}$ vanishes, and the average with respect to $\check{M}^{(1)}$ does not contribute to $\langle\sigma_6\rangle$ at the steady state. Therefore, in the limit of fast switching, we obtain

$$\langle\sigma_6\rangle = \left\langle -\frac{1}{T_b}\left(\dot{\Pi} + \frac{\partial U_{\rm eff}(\theta)}{\partial\theta} - f\right)\circ\frac{\Pi}{m}\right\rangle = \left\langle\frac{\Gamma_b(\theta)}{m}\left(\frac{\Pi^2}{mT_b}-1\right)\right\rangle = \left\langle\sum_b\frac{\Lambda_{b'\to b}(\theta)}{\Lambda_{h\to c}(\theta)+\Lambda_{c\to h}(\theta)}\frac{\Gamma_b(\theta)}{m}\left(\frac{\Pi^2}{mT_b}-1\right)\right\rangle \tag{F8}$$

$$= \left\langle\frac{\Gamma}{m}\left(\frac{\Pi^2}{mT_h}-1\right) + \frac{\gamma\phi'(\theta)^2}{m}\left(\frac{\Pi^2}{mT_c}-1\right)\right\rangle = \langle\sigma_1\rangle. \tag{F9}$$

To obtain the second line, we used the solutions of Eqs. (F6) and (F7).

[1] S. Toyabe, T. Sagawa, M. Ueda, E. Muneyuki, and M. Sano, Experimental demonstration of information-to-energy conversion and validation of the generalized Jarzynski equality, Nat. Phys. **6**, 988 (2010).

[2] D. J. Evans, E. G. D. Cohen, and G. P. Morriss, Probability of Second Law Violations in Shearing Steady States, Phys. Rev. Lett. **71**, 2401 (1993).

[3] C. Jarzynski, Nonequilibrium Equality for Free Energy Differences, Phys. Rev. Lett. **78**, 2690 (1997).

[4] G. M. Wang, E. M. Sevick, E. Mittag, D. J. Searles, and D. J. Evans, Experimental Demonstration of Violations of the Second Law of Thermodynamics for Small Systems and Short Time Scales, Phys. Rev. Lett. **89**, 050601 (2002).

[5] T. Sagawa and M. Ueda, Second Law of Thermodynamics with Discrete Quantum Feedback Control, Phys. Rev. Lett. **100**, 080403 (2008).

[6] M. Esposito, K. Lindenberg, and C. Van den Broeck, Universality of Efficiency at Maximum Power, Phys. Rev. Lett. **102**, 130602 (2009).

[7] N. Shiraishi, K. Saito, and H. Tasaki, Universal Trade-Off Relation Between Power and Efficiency for Heat Engines, Phys. Rev. Lett. **117**, 190601 (2016).

[8] T. Hatano and S.-i. Sasa, Steady-State Thermodynamics of Langevin Systems, Phys. Rev. Lett. **86**, 3463 (2001).

[9] T. S. Komatsu, N. Nakagawa, S.-i. Sasa, and H. Tasaki, Steady-State Thermodynamics for Heat Conduction: Microscopic Derivation, Phys. Rev. Lett. **100**, 230602 (2008).

[10] C. Maes and K. Netočný, A Nonequilibrium Extension of the Clausius Heat Theorem, J. Stat. Phys. **154**, 188 (2014).

[11] U. Seifert, Stochastic thermodynamics, fluctuation theorems and molecular machines, Rep. Prog. Phys. **75**, 126001 (2012).

[12] G. E. Crooks, Entropy production fluctuation theorem and the nonequilibrium work relation for free energy differences, Phys. Rev. E **60**, 2721 (1999).

[13] A. Gomez-Marin, J. M. R. Parrondo, and C. Van den Broeck, Lower bounds on dissipation upon coarse graining, Phys. Rev. E **78**, 011107 (2008).

[14] A. Puglisi, S. Pigolotti, L. Rondoni, and A. Vulpiani, Entropy production and coarse graining in Markov processes, J. Stat. Mech.: Theory Exp. (2010) P05015.

[15] E. Roldán and J. M. R. Parrondo, Estimating Dissipation from Single Stationary Trajectories, Phys. Rev. Lett. **105**, 150607 (2010).

[16] A. Celani, S. Bo, R. Eichhorn, and E. Aurell, Anomalous Thermodynamics at the Microscale, Phys. Rev. Lett. **109**, 260603 (2012).

[17] A. Crisanti, A. Puglisi, and D. Villamaina, Nonequilibrium and information: The role of cross correlations, Phys. Rev. E **85**, 061127 (2012).

[18] M. Esposito, Stochastic thermodynamics under coarse graining, Phys. Rev. E **85**, 041125 (2012).

[19] J. Mehl, B. Lander, C. Bechinger, V. Blickle, and U. Seifert, Role of Hidden Slow Degrees of Freedom in the Fluctuation Theorem, Phys. Rev. Lett. **108**, 220601 (2012).

[20] K. Kawaguchi and Y. Nakayama, Fluctuation theorem for hidden entropy production, Phys. Rev. E **88**, 022147 (2013).

[21] P. Strasberg, G. Schaller, T. Brandes, and M. Esposito, Thermodynamics of a Physical Model Implementing a Maxwell Demon, Phys. Rev. Lett. **110**, 040601 (2013).

[22] T. Munakata and M. L. Rosinberg, Entropy Production and Fluctuation Theorems for Langevin Processes under Continuous Non-Markovian Feedback Control, Phys. Rev. Lett. **112**, 180601 (2014).

[23] E. Zimmermann and U. Seifert, Effective rates from thermodynamically consistent coarse-graining of models for molecular motors with probe particles, Phys. Rev. E **91**, 022709 (2015).

[24] S.-W. Wang, K. Kawaguchi, S.-i. Sasa, and L.-H. Tang, Entropy Production of Nanosystems with Time Scale Separation, Phys. Rev. Lett. **117**, 070601 (2016).

[25] S. Bo and A. Celani, Multiple-scale stochastic processes: Decimation, averaging and beyond, Phys. Rep. **670**, 1 (2017).

[26] R. P. Feynman, R. B. Leighton, and M. Sands, *The Feynman Lectures on Physics* (Basic Books, New York, 2010).

[27] M. Smoluchowski, Experimentell nachweisbare, der üblichen Thermodynamik widersprechende Molekularphänomene, Phys. Z. **13**, 1069 (1912).

[28] J. M. R. Parrondo and P. Español, Criticism of Feynman's analysis of the ratchet as an engine, Am. J. Phys. **64**, 1125 (1996).

[29] I. Derényi and R. D. Astumian, Efficiency of Brownian heat engines, Phys. Rev. E **59**, R6219(R) (1999).

[30] T. Hondou and K. Sekimoto, Unattainability of Carnot efficiency in the Brownian heat engine, Phys. Rev. E **62**, 6021 (2000).

[31] R. Benjamin and R. Kawai, Inertial effects in Büttiker-Landauer motor and refrigerator at the overdamped limit, Phys. Rev. E **77**, 051132 (2008).

[32] N. Shiraishi, Attainability of Carnot efficiency with autonomous engines, Phys. Rev. E **92**, 050101 (2015).

[33] E. Açıkkalp, Analysis of a Brownian heat engine with ecological criteria, Eur. Phys. J. Plus **131**, 426 (2016).

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

[35] M. Büttiker, Transport as a consequence of state-dependent diffusion, Z. Phys. B: Condens. Matter **68**, 161 (1987).

[36] M. M. Millonas, Self-Consistent Microscopic Theory of Fluctuation-Induced Transport, Phys. Rev. Lett. **74**, 10 (1995).

[37] K. Sekimoto, Kinetic characterization of heat bath and the energetics of thermal ratchet models, J. Phys. Soc. Jpn. **66**, 1234 (1997).

[38] T. Hondou and F. Takagi, Irreversible operation in a stalled state of feynman's ratchet, J. Phys. Soc. Jpn. **67**, 2974 (1998).

[39] M. W. Jack and C. Tumlin, Intrinsic irreversibility limits the efficiency of multidimensional molecular motors, Phys. Rev. E **93**, 052109 (2016).

[40] R. Zwanzig, Nonlinear generalized Langevin equations, J. Stat. Phys. **9**, 215 (1973).

[41] K. Sekimoto, Temporal coarse graining for systems of brownian particles with non-constant temperature, J. Phys. Soc. Jpn. **68**, 1448 (1999).

[42] A. M. Jayannavar and M. C. Mahato, Macroscopic equation of motion in inhomogeneous media: A microscopic treatment, Pramana **45**, 369 (1995).

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

[44] M. O. Magnasco and G. Stolovitzky, Feynman's Ratchet and Pawl, J. Stat. Phys. **93**, 615 (1998).

[45] A. Gomez-Marin and J. M. Sancho, Ratchet, pawl and spring Brownian motor, Physica D **216**, 214 (2006).

[46] S. Bo and A. Celani, Entropic anomaly and maximal efficiency of microscopic heat engines, Phys. Rev. E **87**, 050102 (2013).

[47] M. Asfaw and M. Bekele, Exploring the operation of a tiny heat engine, Physica A **384**, 346 (2007).

[48] S. Velasco, J. M. M. Roco, A. Medina, and A. C. Hernández, Feynman's ratchet optimization: maximum power and maximum efficiency regimes, J. Phys. D: Appl. Phys. **34**, 1000 (2001).

[49] N. S. Salas and A. C. Hernández, Unified working regime of nonlinear systems rectifying thermal fluctuations, Europhys. Lett. **61**, 287 (2003).

[50] Z. C. Tu, Efficiency at maximum power of Feynman's ratchet as a heat engine, J. Phys. A: Math. Theor. **41**, 312003 (2008).

[51] B. Lin and J. Chen, Performance characteristics and parametric optimum criteria of a Brownian micro-refrigerator in a spatially periodic temperature field, J. Phys. A: Math. Theor. **42**, 075006 (2009).

[52] L. Chen, Z. Ding, and F. Sun, Optimum performance analysis of Feynman's engine as cold and hot ratchets, J. Non-Equilib. Thermodyn. **36**, 155 (2011).

[53] X. G. Luo, N. Liu, and J. Z. He, Optimum analysis of a Brownian refrigerator, Phys. Rev. E **87**, 022139 (2013).

[54] Y. Apertet, H. Ouerdane, C. Goupil, and Ph. Lecoeur, Revisiting Feynman's ratchet with thermoelectric transport theory, Phys. Rev. E **90**, 012113 (2014).

[55] S. Shi-Qi, Y. Pan, and T. Zhan-Chun, Coefficient of performance at maximum χ-criterion for feynman ratchet as a refrigerator, Commun. Theor. Phys. **62**, 589 (2014).

[56] N. G. van Kampen, *Stochastic Processes in Physics and Chemistry*, North-Holland personal library (North-Holland, Amsterdam,1992).

[57] H. A. Kramers, Brownian motion in a field of force and the diffusion model of chemical reactions, Physica **7**, 284 (1940).