Heat fluctuations and initial ensembles

Kwangmoo Kim,[1,2] Chulan Kwon,[3] and Hyunggyu Park[2]
[1]*Research Institute of Advanced Materials, Seoul National University, Seoul 151-742, Korea*
[2]*School of Physics, Korea Institute for Advanced Study, Seoul 130-722, Korea*
[3]*Department of Physics, Myongji University, Yongin, Gyeonggi-do 449-728, Korea*
(Received 30 June 2014; published 16 September 2014)

Time-integrated quantities such as work and heat increase incessantly in time during nonequilibrium processes near steady states. In the long-time limit, the average values of work and heat become asymptotically equivalent to each other, since they only differ by a finite energy change in average. However, the fluctuation theorem (FT) for the heat is found not to hold with the equilibrium initial ensemble, while the FT for the work holds. This reveals an intriguing effect of everlasting initial memory stored in rare events. We revisit the problem of a Brownian particle in a harmonic potential dragged with a constant velocity, which is in contact with a thermal reservoir. The heat and work fluctuations are investigated with initial Boltzmann ensembles at temperatures generally different from the reservoir temperature. We find that, in the infinite-time limit, the FT for the work is fully recovered for arbitrary initial temperatures, while the heat fluctuations significantly deviate from the FT characteristics except for the infinite initial-temperature limit (a uniform initial ensemble). Furthermore, we succeed in calculating finite-time corrections to the heat and work distributions analytically, using the modified saddle point integral method recently developed by us. Interestingly, we find noncommutativity between the infinite-time limit and the infinite-initial-temperature limit for the probability distribution function (PDF) of the heat.

DOI: 10.1103/PhysRevE.90.032117 PACS number(s): 05.70.Ln, 02.50.−r, 05.40.−a

I. INTRODUCTION

The fluctuation theorem (FT) has been regarded as a fundamental principle in nonequilibrium statistical mechanics. It concerns time-integrated quantities such as heat and work in nonequilibrium processes. The FT provides a rigorous rule for the thermal fluctuations of such quantities, independent of any detailed dynamics. The first form of the FT was found for entropy production piled in a heat bath for a deterministic thermostated system [1–3], given as $P(s)/P(-s) \to e^{\tau s}$ as $\tau \to \infty$, where τ is the measuring time and s is the entropy production rate in the unit of the Boltzmann constant k_B. It is termed as a steady-state detailed FT in literature.

Later, the FT was found to hold in stochastic systems [4–6]. In particular, Crooks showed that the (transient) FT rigorously holds at all times (finite τ) for the work W produced in nonequilibrium systems starting from equilibrium distributions. Furthermore, it can be expressed in a more detailed form, termed as a detailed FT, regarding the probability distribution function (PDF) of work fluctuations, given as

$$\frac{P_F(W)}{P_R(-W)} = e^{\beta(W-\Delta F)}, \tag{1}$$

where β is the inverse temperature of the heat bath and ΔF the free energy difference between the initial and final times due to the change in a time-dependent protocol such as a volume, an external field, or a potential shape. P_F denotes the PDF for the forward (F) process, while P_R denotes that for the reverse (R) process where the protocol varies in time reversely to the forward process. The symmetry of the PDF such as in Eq. (1) is known, in general, as the Gallavotti-Cohen symmetry [3]. The Jarzynski equality $\langle e^{-\beta W}\rangle = e^{-\beta\Delta F}$ [7] is nothing but the integral FT corresponding to the Crooks detailed FT. The discovery of the FT, which is expected to be valid in general stochastic systems, has resulted in extensive studies on unprecedented nonequilibrium phenomena. Many experimental evidences have also been reported [8–12].

The choice of an initial ensemble is critical for the validity of the FT. For example, the transient detailed FT's for any finite τ hold and so do the integral FT's, only with the equilibrium Boltzmann distribution as the initial ensemble for the work or with the uniform (infinite-temperature) distributions for the heat [13–15] (see also Sec. III C). The total entropy production satisfies the transient integral FT with an arbitrary initial ensemble [16], though its detailed FT is valid only in the steady state. In fact, the detailed FT guarantees the integral FT, but the converse is true only if the initial distributions for the forward and reverse paths are *involutory* to each other [17].

A natural question arises: "What happens to the FT when other types of initial ensembles are taken?" It is clear that the transient FT for a time-integrated quantity does not hold without a proper initial ensemble, but not clear whether or not it holds in the $\tau \to \infty$ limit. If not, how can the initial memory persist in the long-time limit? In order to answer these questions, we investigate the effect of initial ensembles on the detailed FT for the heat and work, in particular for large τ.

It is a formidable task to calculate the PDF exactly for finite τ, so we restrict ourselves only to its large deviation function and corrections in the large τ limit. The PDF for a time-integrated quantity A for a long period of time τ can be written in a scaling form

$$P(A) \sim e^{-\tau h(A/\tau)} \quad \text{for large } \tau, \tag{2}$$

where A is usually scaled dimensionless and $h(A/\tau)$ is called a large deviation function (LDF) [18,19]. Many interesting properties on the LDF were found on such as the current fluctuations [20,21], the (non-Gaussian) exponential tail [22], the everlasting initial memory threshold [14,23], and so on. If

 ©2014 American Physical Society

the detailed FT holds, the Gallavotti-Cohen (GC) symmetry is expressed in terms of the LDF's as

$$-h_F(A/\tau) + h_R(-A/\tau) = A/\tau, \tag{3}$$

where h_F (h_R) is the LDF for the forward (reverse) process. When $A = \beta W$, Eq. (1) leads to the above symmetry in the large τ limit [24].

The thermodynamic first law reads $\Delta E = W - Q$ for the energy change ΔE. We define Q as the heat transferred into the heat bath. In nonequilibrium close to the steady state, $\langle W\rangle$ and $\langle Q\rangle$ grow linearly in τ, but $\langle \Delta E\rangle$ remains finite. Thus, one might expect that both quantities approximately have the identical PDF's for large τ as the difference ΔE may become negligible. Starting with the equilibrium Boltzmann ensemble, the detailed FT for W is satisfied even at finite τ, and thus is expected to be valid also for Q at least in the infinite τ limit. However, the reality is against the expectation. The detailed FT for the heat was examined analytically for the motion of a particle in a harmonic potential dragged with a constant velocity, which is one of the experimental prototypes [8,9,25,26]. It was found in the infinite τ limit that $\ln P_F(Q)/P_R(-Q) \simeq \beta Q$ only in the central region around $Q = 0$ [27], while it approaches a plateau for large $|Q|$, which is the origin of the extended FT [28,29]. Recently, the modification of the detailed FT for the heat has been proposed for general systems in terms of correlations between ΔE and Q [15].

The violation of the FT is due to a rare but non-negligible chance of ΔE having an extremely large value, which causes the FT modified in the tail region of the PDF [14]. The probability to find the initial system with an extremely large energy is exponentially small, but it will almost always dissipate the most of its energy into the reservoir in the long-time limit. Thus, this event becomes relevant to the tail part of the heat PDF which also decays exponentially for large $|Q|$. Even for very large τ and $|Q|$, there is always an exponentially small probability to find the event with the corresponding large energy in the initial Boltzmann ensemble. Therefore, this effect cannot go away even in the long-time limit. This so-called "boundary effect" recognized in many references [30–33] is observed for an unbounded energy distribution in the initial ensemble, but obviously not observed when the initial energy is bounded.

As the initial ensemble plays a crucial role in the FT violation, we study its effect on the work and heat fluctuations more systematically in this paper. As an initial ensemble, we take the Boltzmann distribution at a temperature generally different from that of the heat reservoir. In this case, the FT's for both W and Q do not hold for finite τ. However, it is not obvious whether the FT will hold or not in the large τ limit. The validity may depend on the quantity of interest and also on the temperature of the initial ensemble. In fact, it is already reported that the injected and dissipated PDF's of heat in an equilibration process show phase transitions at two different finite initial temperatures, respectively, below which the LDF is not affected, while above which the LDF is significantly modified by the boundary term [14,30].

In this paper, we revisit the problem of a Brownian particle in a harmonic potential dragged with a constant velocity, which is in contact with the thermal reservoir. We then investigate the PDF's of the work and heat for a long period of τ. For the heat PDF, we find the singularities due to the boundary terms, which vary with the temperature of the initial ensemble and break the GC symmetry of the PDF. As the initial temperature approaches the infinity in the infinite-τ limit, the GC symmetry is restored. We also calculated a finite-τ correction for the heat PDF, where the singularity structure becomes more complicated. Using the modified saddle point integral method recently developed by us [23], we exactly obtained the LDF of the heat up to $\mathcal{O}(\tau^{-1})$ and thus the FT violation is measured up to the same order. Interestingly, the finite-τ correction of the FT violation does not vanish in the infinite initial-temperature limit, which implies the noncommutativity between the two limits of the infinite τ and the infinite initial temperature. However, we can show that the transient FT is satisfied for any τ if one takes a proper infinite initial-temperature limit before taking the infinite-τ limit.

In contrast, the work PDF turns out to be free of any singularity even at any initial temperature. Furthermore, the work PDF can be calculated exactly at any finite τ and any initial temperature. We can show that the transient FT does not hold except when the initial temperature is identical to the temperature of the reservoir. However, in the infinite-τ limit, the FT is fully restored, regardless of the initial temperature. The difference between the FT violations for the heat and work comes from the presence of ΔE, which induces everlasting initial memory in the heat PDF.

The remainder of this paper is organized as follows. In Sec. II, we introduce a model and theoretical formalism to obtain the PDF's of the heat and work. The generating functions for the heat and work PDF's are derived. In Sec. III, we present the LDF and the FT violation for the heat fluctuations in the long-time limit and their finite-time corrections. The restoration of the FT for the heat in the infinite initial-temperature limit is also discussed. In Sec. IV, we repeat the calculations for the work fluctuations. Finally, in Sec. V, we summarize our study and discuss the physical origin of the everlasting initial memory in the time-accumulated quantities. In the Appendix, the exact generating functions for the heat and work are given at finite τ.

II. MODEL AND GENERATING FUNCTIONS

A. Model

The Brownian motion of a particle in a moving harmonic potential with a constant velocity $\mathbf{v}^*$ [28] is described by an overdamped Langevin equation as

$$\dot{\mathbf{x}}_t = -\tau_r^{-1}(\mathbf{x}_t - \mathbf{x}_t^*) + \alpha^{-1}\boldsymbol{\zeta}_t, \tag{4}$$

where $\mathbf{x}_t$ is the position of the particle at time t, τ_r is the relaxation time, $\mathbf{x}_t^* = \mathbf{v}^* t$ is the moving center of the harmonic potential, and α is the Stokes friction of the particle in a fluid. The relaxation time is given by $\tau_r = \alpha/k$, where k is the force constant of the harmonic potential. $\boldsymbol{\zeta}_t$ is a fluctuating white noise given as

$$\langle \zeta_t^a\rangle = 0, \quad \langle \zeta_t^a \zeta_{t'}^{a'}\rangle = 2\alpha k_B T \delta_{aa'}\delta(t-t'), \tag{5}$$

where the superscripts a and a' denote component indices $(a,a' = 1,\ldots,d)$ for a d-dimensional motion and T is the temperature of the heat bath. The particle and the center of the harmonic potential are initially positioned at the origin: $\mathbf{x}_0 = \mathbf{x}_0^* = 0$.

For convenience, we first find out the *deterministic* part $\mathbf{y}_t^*$ of the solution to Eq. (4) as

$$\mathbf{y}_t^* = \mathbf{v}^* t - \mathbf{v}^* \tau_r (1 - e^{-t/\tau_r}), \tag{6}$$

satisfying the deterministic equation $\dot{\mathbf{y}}_t^* = -\tau_r^{-1}(\mathbf{y}_t^* - \mathbf{x}_t^*)$ with an initial condition $\mathbf{y}_0^* = 0$. If we look at the relative motion of the particle as

$$\mathbf{X}_t = \mathbf{x}_t - \mathbf{y}_t^*, \tag{7}$$

then it satisfies a simpler equation of motion as

$$\dot{\mathbf{X}}_t = -\tau_r^{-1}\mathbf{X}_t + \alpha^{-1}\boldsymbol{\zeta}_t. \tag{8}$$

The harmonic potential energy $U_t = \frac{k}{2}(\mathbf{x}_t - \mathbf{x}_t^*)^2$ has an explicit time dependence. As recognized by Jarzynski [7], the work is transferred into the system by the rate of $\partial U_t/\partial t$. It is performed by an external agent (experimental device) to change the protocol $\mathbf{x}_t^*$. Then, the work W_τ delivered into the system can be expressed along a given trajectory $[\mathbf{x}_t]_0^\tau$ for $0 \leqslant t \leqslant \tau$ as

$$\begin{aligned} W_\tau &= -k\int_0^\tau dt(\mathbf{x}_t - \mathbf{x}_t^*)\cdot\mathbf{v}^* \\ &= -k\int_0^\tau dt[\mathbf{v}^*\cdot\mathbf{X}_t + \mathbf{v}^*\cdot(\mathbf{y}_t^* - \mathbf{x}_t^*)]. \end{aligned} \tag{9}$$

The heat Q_τ going into the fluid along the same trajectory $[\mathbf{x}_t]_0^\tau$ is given by

$$Q_\tau = W_\tau - \Delta U_\tau, \tag{10}$$

where $\Delta U_\tau = U_\tau - U_0$ is the potential energy change. Note that only the potential energy change is considered in the overdamped limit.

The PDF for the work or heat can be obtained by considering all the possible trajectories. For convenience, we scale the heat and work by the temperature of the heat bath to get dimensionless quantities as βQ and βW with $\beta = 1/(k_B T)$. The finite-τ PDF for a quantity A_τ ($=\beta Q_\tau$ or βW_τ) can be written as

$$\begin{aligned} P_\tau(A) &= \langle\delta(A - A_\tau)\rangle \\ &= \int_{-i\infty}^{+i\infty}\frac{d\lambda}{2\pi i}\, e^{\lambda A}\langle e^{-\lambda A_\tau}\rangle, \end{aligned} \tag{11}$$

where A_τ is the trajectory-dependent fluctuating quantity and $\langle\cdots\rangle$ denotes an average over all the possible trajectories and the initial distribution.

It is convenient to introduce a generating function defined as

$$g_A(\lambda) \equiv \langle e^{-\lambda A_\tau}\rangle = \int_{-\infty}^{+\infty} dA\; P_\tau(A)e^{-\lambda A}. \tag{12}$$

Then, the PDF is simply a Fourier transform of the generating function as in Eq. (11). The GS symmetry in terms of the generating function can be obtained, using Eq. (3), as

$$g_A(\lambda) = g_A(1-\lambda), \tag{13}$$

where the process indices, F and R, are dropped because the generating functions for the forward and reverse processes are equivalent to each other in our constantly moving harmonic potential. Any energetic quantity like heat or work does not depend on the sign of the velocity $\mathbf{v}^*$ of the moving harmonic potential. Furthermore, the free energy difference in Eq. (1) is always zero ($\Delta F = 0$), since the shape of the potential energy does not change except for a translation.

In order to study the influence of an initial condition, we assume that the particle initially has an equilibrium distribution at the initial inverse temperature β' as

$$\rho_{\rm in}(\mathbf{X}_0) = \left(\frac{\beta' k}{2\pi}\right)^{d/2} e^{-(\beta' k/2)\mathbf{X}_0^2}. \tag{14}$$

B. Generating function for heat

The generating function $g_Q(\lambda)$ for the heat is written as

$$\begin{aligned} g_Q(\lambda) &= \langle e^{-\lambda\beta Q_\tau}\rangle = \langle e^{-\lambda\beta(W_\tau - \Delta U_\tau)}\rangle \\ &= \int d\mathbf{X}_\tau e^{(\lambda\beta k/2)(\mathbf{X}_\tau + \Delta\mathbf{y}_\tau^*)^2}\int d\mathbf{X}_0\rho_{\rm in}(\mathbf{X}_0)e^{-(\lambda\beta k/2)\mathbf{X}_0^2} \\ &\quad\times\int D[\mathbf{X}_t]e^{\int_0^\tau dt\mathcal{L}}e^{\lambda\beta k\int_0^\tau dt(\mathbf{v}^*\cdot\mathbf{X}_t + \mathbf{v}^*\cdot\Delta\mathbf{y}_t^*)}, \end{aligned} \tag{15}$$

where $\Delta\mathbf{y}_t^* = \mathbf{y}_t^* - \mathbf{x}_t^* = -\mathbf{v}^*\tau_r(1 - e^{-t/\tau_r})$ from Eq. (6) and $\int D[\mathbf{X}_t]$ denotes the path integral over all the trajectories connecting $\mathbf{X}_0$ and $\mathbf{X}_\tau$, with proper normalizations. The Lagrangian $\mathcal{L}$ is given in a prepoint (Ito) representation for the time discretization [34] as

$$\mathcal{L} = -\frac{1}{4D}\left(\dot{\mathbf{X}}_t + \tau_r^{-1}\mathbf{X}_t\right)^2 \tag{16}$$

for $D = (\beta\alpha)^{-1}$.

Noting that $\mathcal{L}$ is quadratic in $\mathbf{X}_t$, the generating function is in fact a succession of a multivariate Gaussian integral over $\mathbf{X}_j$ at discretized times $t_j = j\tau/N$ $(j = 0,1,\ldots,N)$ with a large N. We can compute the integral in the $N \to \infty$ limit by using the method in our previous work [35]. It is convenient to rewrite the generating function in terms of normalized Gaussian integrations over $\{\mathbf{X}_j\}$ $(0 \leqslant j \leqslant N)$ as

$$g_Q(\lambda) = c\mathcal{N}\langle e^{\lambda\beta k(\Delta\mathbf{y}_\tau^*\cdot\mathbf{X}_\tau + \mathbf{v}^*\cdot\int_0^\tau dt\mathbf{X}_t)}\rangle_{\{\mathbf{X}_j\}}, \tag{17}$$

where the average is defined as

$$\begin{aligned} \langle\mathcal{O}\rangle_{\{\mathbf{X}_j\}} &= \frac{1}{\mathcal{N}}\int d\mathbf{X}_N e^{(\lambda\beta k/2)\mathbf{X}_N^2}\int d\mathbf{X}_0\rho_{\rm in}(\mathbf{X}_0)e^{-(\lambda\beta k/2)\mathbf{X}_0^2} \\ &\quad\times\int\prod_j d\mathbf{X}_j e^{\int_0^\tau dt\mathcal{L}}\,\mathcal{O}, \end{aligned} \tag{18}$$

with the normalization constant $\mathcal{N}$ obtained from $\langle 1\rangle_{\{\mathbf{X}_j\}} = 1$. The nonfluctuating deterministic part yields

$$c = e^{(\lambda\beta k/2)(\Delta\mathbf{y}_\tau^*)^2 + \lambda\beta k\mathbf{v}^*\cdot\int_0^\tau dt\Delta\mathbf{y}_t^*}. \tag{19}$$

As the distribution in the above average is a simple Gaussian, it is sufficient to consider the cumulants only up to the second order. It is then straightforward to find

$$\begin{aligned} g_Q(\lambda) &= c\mathcal{N}e^{[(\lambda\beta k)^2/2]|\mathbf{v}^*|^2\int_0^\tau dt\int_0^\tau dt' C(t,t')} \\ &\quad\times e^{[(\lambda\beta k)^2/2][|\Delta\mathbf{y}_\tau^*|^2 C(\tau,\tau) + 2\Delta\mathbf{y}_\tau^*\cdot\mathbf{v}^*\int_0^\tau dt C(\tau,t)]}, \end{aligned} \tag{20}$$

where $C(t,t')$ is a correlation function given by

$$\left\langle X_t^a X_{t'}^{a'} \right\rangle_{\{\mathbf{X}_i\}} = \delta_{aa'} C(t,t'). \tag{21}$$

The integrations at the initial and final points in Eq. (18) include the extra boundary factors $e^{-\lambda\beta k \mathbf{X}_0^2/2}$ and $e^{\lambda\beta k \mathbf{X}_N^2/2}$, respectively, which modify $\mathcal{N}$ and $C(t,t')$ significantly. After some algebra with the initial Boltzmann condition with the inverse temperature β' in Eq. (14), we find

$$\mathcal{N} = \left(\frac{\beta'}{\beta' + \lambda\beta}\right)^{d/2} \left(\frac{1}{1 - \lambda\beta k A_\tau^{-1}}\right)^{d/2} \tag{22}$$

and

$$C(t,t') = e^{-(t-t')/\tau_r} A_{t'}^{-1} \frac{1 - \lambda\beta k a_{\tau-t}^{-1}}{1 - \lambda\beta k A_\tau^{-1}} \quad \text{for} \quad t \geqslant t', \tag{23}$$

where

$$\begin{aligned} A_t^{-1} &= (\beta k)^{-1}(1 - e^{-2t/\tau_r}) + (\beta' k + \lambda\beta k)^{-1} e^{-2t/\tau_r}, \\ a_{\tau-t}^{-1} &= (\beta k)^{-1}(1 - e^{-2(\tau-t)/\tau_r}). \end{aligned} \tag{24}$$

Note that A_t is the Gaussian kernel at time t during the path integral. Without any extra term, $\mathcal{N} = 1$ and $C(t,t') = e^{-(t-t')/\tau_r} A_{t'}^{-1}$ with $A_{t'}^{-1} = (\beta k)^{-1}$.

For simplicity, we adopt the same parameter values and notations in Ref. [29] as

$$\beta = 1, \quad k = 1, \quad \tau_r = 1, \quad \text{and} \quad |\mathbf{v}^*|^2 = w. \tag{25}$$

In these units, w is equal to the average work rate in the steady state: $w = \lim_{\tau\to\infty}\langle W_\tau\rangle/\tau$. Putting all together into Eq. (20), we find the exact expression for $g_Q(\lambda)$, which is quite complicated and is shown in the Appendix. In the following, we will evaluate the LDF up to the order of τ^{-1}, so here we ignore all the exponentially decaying terms like $e^{-\tau}$ in $g_Q(\lambda)$. Then, we get a rather simple form as

$$g_Q(\lambda) = \frac{\beta'^{d/2} e^{-w\{\tau\lambda(1-\lambda) - (3/2)\lambda + (\lambda^2/2)[4 - 1/(\beta'+\lambda)]\}}}{[(\beta' + \lambda)(1 - \lambda)]^{d/2}}. \tag{26}$$

Note that the GC symmetry in Eq. (13) seems to be preserved at the level of the large deviation function (exponent) in the $\tau \to \infty$ limit. However, the singular property of the prefactor coming from the boundary terms does not uphold the GC symmetry, which causes a significant violation of the GC symmetry in the heat PDF, even in the $\tau \to \infty$ limit.

C. Generating function for work

The generating function $g_W(\lambda)$ for the work is given as

$$\begin{aligned} g_W(\lambda) &= \langle e^{-\beta\lambda W_\tau} \rangle \\ &= \int d\mathbf{X}_\tau \int d\mathbf{X}_0 \rho_{\text{in}}(\mathbf{X}_0) \int D[\mathbf{X}_t]\, e^{\int_0^\tau dt \mathcal{L}} \\ &\quad \times e^{\lambda\beta k \int_0^\tau dt (\mathbf{v}^*\cdot\mathbf{X}_t + \mathbf{v}^*\cdot\Delta\mathbf{y}_t^*)}. \end{aligned} \tag{27}$$

Similarly, we get

$$g_W(\lambda) = e^{\lambda\beta k \mathbf{v}^*\cdot\int_0^\tau dt\, \Delta\mathbf{y}_t^*} e^{[(\lambda\beta k)^2/2]|\mathbf{v}^*|^2} e^{\int_0^\tau dt \int_0^\tau dt' C(t,t')}, \tag{28}$$

with the correlation function

$$C(t,t') = e^{-(t-t')/\tau_r} A_{t'}^{-1} \quad \text{for} \quad t > t', \tag{29}$$

where

$$A_t^{-1} = (\beta k)^{-1}(1 - e^{-2t/\tau_r}) + (\beta' k)^{-1} e^{-2t/\tau_r}. \tag{30}$$

Using the same convention ($\beta = 1$, $k = 1$, $\tau_r = 1$, and $|\mathbf{v}^*|^2 = w$) and neglecting the exponentially decaying terms like $e^{-\tau}$ (see the full solution in the Appendix), we find

$$g_W(\lambda) = e^{-w[\tau\lambda(1-\lambda) - \lambda + (\lambda^2/2)(3 - 1/\beta')]}. \tag{31}$$

The GC symmetry is satisfied only in the $\tau = \infty$ limit, but for an arbitrary β'. At $\beta' = 1$, it holds for any finite τ as expected, even when the exponentially decaying terms are included in $g_W(\lambda)$ (see the Appendix).

III. LDF AND FT FOR HEAT

A. Long-time limit

As $\tau \to \infty$ in Eq. (26), the generating function $g_Q(\lambda)$ exhibits the large deviation behavior as

$$g_Q(\lambda) \sim e^{-w\tau e(\lambda)} \tag{32}$$

with

$$e(\lambda) = \begin{cases} \lambda(1 - \lambda) & \text{for } -\beta' < \lambda < 1, \\ -\infty & \text{otherwise}, \end{cases} \tag{33}$$

where the divergence is evident as $\lambda \to 1$ from below and $-\beta'$ from above. Each of them is due to the boundary term at the final and initial points, respectively.

As the PDF $P_\tau(Q)$ is given by the Fourier transformation of $g_Q(\lambda)$ as in Eq. (11), we expect for large τ

$$P_\tau(Q) \sim e^{-w\tau h(p)} \quad \text{with} \quad p \equiv Q/(w\tau), \tag{34}$$

where p is a properly scaled variable for the heat. Then, the LDF $h(p)$ is simply given by the Legendre transform of $e(\lambda)$, given by

$$h(p) = \max_\lambda [e(\lambda) - \lambda p]. \tag{35}$$

We find

$$h(p) = \begin{cases} -p & \text{for } p < -1, \\ (p-1)^2/4 & \text{for } -1 \leqslant p \leqslant 2\beta' + 1, \\ \beta' p - \beta'(1 + \beta') & \text{for } p > 2\beta' + 1. \end{cases} \tag{36}$$

Note that the nonanalytic behavior of the LDF $h(p)$ originates from the divergence of $e(\lambda)$ due to the prefactor singularity in $g_Q(\lambda)$.

The detailed FT for the heat is examined by

$$f_\tau(p) = \frac{1}{\omega\tau} \ln\left[\frac{P_\tau(w\tau p)}{P_\tau(-w\tau p)}\right], \tag{37}$$

where $f_\tau(-p) = -f_\tau(p)$. If the transient detailed FT is satisfied, then $f_\tau(p) = p$ for any τ. In the $\tau \to \infty$ limit, we can easily find $f_\infty(p) = -h(p) + h(-p)$, yielding

$$f_\infty(p) = \begin{cases} p & \text{for} \quad 0 \leqslant p < 1, \\ p - (p-1)^2/4 & \text{for} \quad 1 \leqslant p < 2\beta' + 1, \\ (1 - \beta')p + \beta'(1 + \beta') & \text{for} \quad p \geqslant 2\beta' + 1. \end{cases} \tag{38}$$

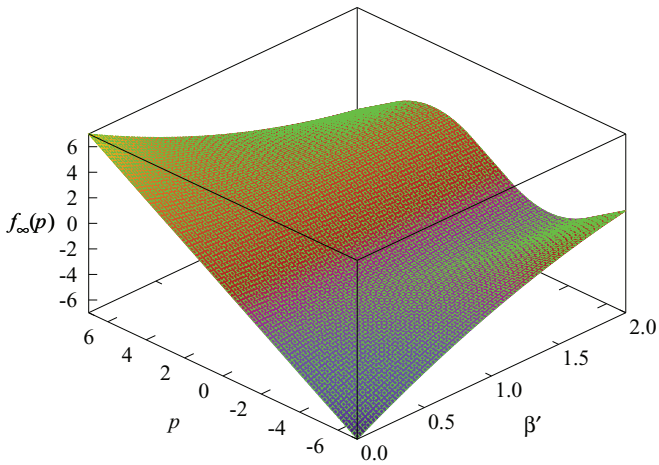

FIG. 1. (Color online) $f_\infty(p)$ as in Eq. (38) is plotted as a function of p and β'.

Indeed, the detailed FT for the heat holds only for $|p| \leqslant 1$ (region I), outside of which $f_\infty(p)$ deviates significantly from the FT relation. Its deviation depends on the initial temperature β' and differs in $1 \leqslant |p| < 2\beta' + 1$ (region II) and in $|p| \geqslant 2\beta' + 1$ (region III), as seen in Figs. 1 and 2. It is interesting to note that the FT is restored for all p in the $\beta' = 0$ (infinite initial-temperature) limit, where region II disappears and $f_\infty(p)$ approaches p in region III. We will be back to this limit later in this section. The extended FT discussed by van Zon and Cohen [28,29] is a special case at $\beta' = \beta = 1$.

B. Finite-time corrections

It is difficult to compare the results in the $\tau \to \infty$ limit with those in the simulations or experiments, due to huge sampling errors in the PDF tail representing rare events. In particular, the FT violation appears in this tail region. It is thus desirable to estimate finite-time corrections analytically. We want to evaluate the LDF up to $\mathcal{O}(\tau^{-1})$.

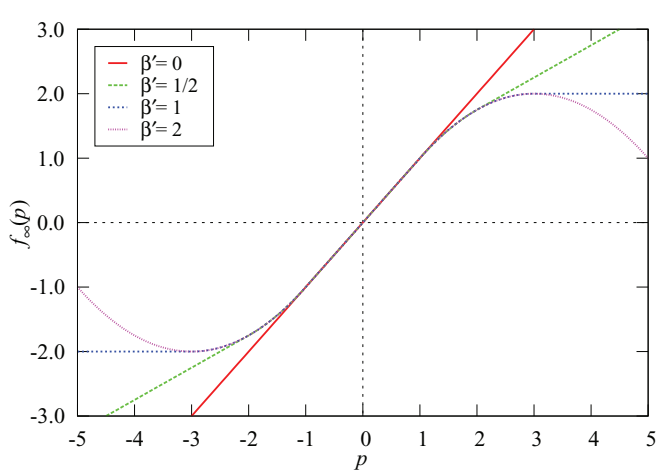

FIG. 2. (Color online) $f_\infty(p)$ for $\beta' = 0$, 1/2, 1, and 2.

From Eq. (11), the PDF for the heat Q $(=w\tau p)$ is written as

$$P_\tau(w\tau p) = \int_{-i\infty}^{+i\infty} d\lambda \phi(\lambda) e^{-w\tau H(\lambda,p)}, \tag{39}$$

with the prefactor

$$\phi(\lambda) = \frac{\beta'^{d/2}}{2\pi i[(\beta'+\lambda)(1-\lambda)]^{d/2}} e^{w[3\lambda/2 - 2\lambda^2 + \lambda^2/2(\beta'+\lambda)]}, \tag{40}$$

and

$$H(\lambda,p) = -\lambda p + \lambda(1-\lambda). \tag{41}$$

The prefactor shows singularities at $\lambda = \lambda_1 = 1$ and $\lambda_2 = -\beta'$, which are simple poles for $d = 2$, but branch points for $d = 1,3$. Later, we choose a branch cut on the real axis of λ for $\lambda > \lambda_1$ and $\lambda < \lambda_2$ when $d \neq 2$. In addition, there is an essential singularity at $\lambda = \lambda_2$, which will cause a little more complication in the following integration.

The integral for large τ can be approximated by the integral along the steepest descent path passing through a saddle point in the complex plane of λ. In the conventional saddle-point approximation, a saddle point $\lambda_0^*(p)$ is chosen by extremizing $H(\lambda,p)$ such as $dH/d\lambda|_{\lambda_0^*} = 0$, yielding $\lambda_0^* = (1-p)/2$. However, the integral may diverge due to the prefactor $\phi(\lambda)$ when the saddle point approaches one of its singularities.

In this study, we adopt the modified saddle point integral method [23] and search for the modified saddle points $\lambda^*(p)$ by extremizing

$$S(\lambda,p) \equiv H(\lambda,p) - (w\tau)^{-1} \ln \phi(\lambda), \tag{42}$$

with

$$\left.\frac{dS}{d\lambda}\right|_{\lambda=\lambda^*} = 0. \tag{43}$$

There are multiple saddle points for a given p. However, it can be shown that there always exists a saddle point $\lambda^*(p)$ on the real-λ axis between the two singularities, i.e., $\lambda_2 < \lambda^*(p) < \lambda_1$. This saddle point is τ dependent and sometimes approaches the singularities asymptotically for large τ. For $-1 < p < 2\beta' + 1$, λ^* approaches the conventional saddle point λ_0^*, otherwise one of the singularities such as λ_1 for $p < -1$ and λ_2 for $p > 2\beta' + 1$, respectively.

When the modified saddle point λ^* is nearby the singularities, the integral along the steepest descent path should be performed with special care, because it becomes a non-Gaussian integral, described in detail in the Appendix of Ref. [23].

Now we present the results for different regions of p as follows.

1. Central region of the PDF

Sufficiently deep inside of the interval of $-1 < p < 2\beta' + 1$, the saddle point λ^* is given by

$$\lambda^* = \frac{1-p}{2} + \mathcal{O}(\tau^{-1}), \tag{44}$$

which approaches λ_0^* for large τ and is far enough from the singularities at λ_1 and λ_2. Thus, one can apply the conventional

saddle point approximation [see Eq. (A.20) in Ref. [23]], which yields

$$P_\tau(w\tau p) = i\left[\frac{2\pi}{w\tau|H''(\lambda_0^*)|}\right]^{1/2}\phi(\lambda_0^*)e^{-w\tau H(\lambda_0^*)}$$
$$= \sqrt{\frac{\beta'^d}{\pi w\tau}}\frac{2^{d-1}}{[(1+p)(2\beta'+1-p)]^{d/2}}$$
$$\times\ e^{-w\tau[(1-p)^2/4]+(3/4)w(1-p)}$$
$$\times e^{-w(1-p)^2[2(1-p)+4\beta'-1]/4(2\beta'+1-p)}. \quad (45)$$

This result is exact up to $\mathcal{O}(\tau^{-1})$ for the τ-dependent LDF defined as

$$h_\tau(p) \equiv -\frac{1}{w\tau}\ln P_\tau(w\tau p)$$
$$= h_c(p) + \frac{\ln\tau}{2w\tau} - \frac{r_c(p)}{w\tau}, \quad (46)$$

where $h_c(p) = (1-p)^2/4$ and $r_c(p)$ is the logarithm of the $\mathcal{O}(1)$ terms in the multiplicative factor and also in the exponent in Eq. (45). The usual asymptotic LDF in Eq. (36) is obtained as $h(p) = \lim_{\tau\to\infty} h_\tau(p)$.

2. *Left wing of the PDF*

The saddle point λ^* approaches the singularity at $\lambda = \lambda_1$ $(= 1)$ from below, in the left side of the central region ($p \lesssim -1$). Let us write $\delta\lambda_1 = \lambda^* - \lambda_1$ (<0). For small $\delta\lambda_1$ and large τ, the saddle-point equation (43) is expanded in $\delta\lambda_1$ as

$$-2\delta\lambda_1 - (p+1) + \frac{d}{2w\tau\delta\lambda_1} \approx 0. \quad (47)$$

Its proper solution is

$$\delta\lambda_1 = \frac{1}{4}\left[-(p+1) - \sqrt{(p+1)^2 + \frac{4d}{w\tau}}\right]. \quad (48)$$

For $(p+1)\tau^{1/2} \ll -1$, Eq. (48) becomes

$$\delta\lambda_1 \approx \frac{d}{2w(p+1)}\tau^{-1}, \quad (49)$$

which determines the PDF in the region of $p < -1$. Note that the saddle point is already very close to λ_1 with a distance of $\mathcal{O}(\tau^{-1})$.

For $|p+1|\tau^{1/2} \ll 1$, Eq. (48) becomes

$$\delta\lambda_1 \approx -\frac{1}{2}\sqrt{\frac{d}{w}}\tau^{-1/2}, \quad (50)$$

which determines the PDF in a narrow region around $p = -1$. This region vanishes in the $\tau \to \infty$ limit. In this case, the distance between the saddle point and λ_1 shrinks more slowly with a distance of $\mathcal{O}(\tau^{-1/2})$.

The steepest descent integration passing through the saddle point near the singularity becomes problematic, mainly because the singular prefactor cannot be expanded around the singularity. However, the integration can be still performed only with the expansion of the exponent $H(\lambda, p)$ around the saddle point. The price to pay is that one should perform a non-Gaussian integration along the steepest descent path. The integration results are explicitly given in the Appendix of Ref. [23] for general power-law singularities. Here, we just briefly sketch the integration method.

We expand $H(\lambda, p)$ in powers of $\delta\lambda_1$ and use a new variable v defined as $v = 1 + (\lambda - \lambda^*)/\delta\lambda_1$. Then, Eq. (39) can be written as

$$P_\tau(w\tau p) = C_1\int_{1-i\infty}^{1+i\infty} dv\,\frac{e^{w\tau[(p+1)\delta\lambda_1 v + \delta\lambda_1^2 v^2]}}{v^{d/2}}, \quad (51)$$

where

$$C_1 = \frac{1}{2\pi i}\frac{\beta'^{d/2}(-\delta\lambda_1)^{1-d/2}}{(1+\beta')^{d/2}}e^{w[\tau p - \beta'/2(\beta'+1)]}. \quad (52)$$

This integral can be simplified by modifying the integral contour $[1-i\infty, 1+i\infty]$ into a composite of two straight lines of $[-i\infty, -i\epsilon]$ and $[+i\epsilon, +i\infty]$ and a semicircle with an infinitesimally small radius ϵ to avoid the singular point at the origin ($v = 0$). By changing the variable to y as $v = iy$, the integration along the two straight lines becomes a real-valued integral and the contribution from the semicircle contour can be also done, using the polar coordinate representation. Summing up these contributions, one can finally come up with a single real-valued integral expression as in Eq. (A16) of Ref. [23]. Then, it is possible to integrate even the tail part of the PDF numerically with very high precision.

In this paper, we just present the results only in the two scaling regimes of $(p+1)\tau^{1/2} \ll -1$ and $|p+1|\tau^{1/2} \ll 1$. In addition, we restrict ourselves to the case of $1 \leqslant d < 4$ for simplicity. For $(p+1)\tau^{1/2} \ll -1$, we find

$$P_\tau(w\tau p) = \frac{2\beta'^{d/2}(w\tau|p+1|)^{(d-2)/2}}{\pi(1+\beta')^{d/2}}\Gamma(2-d/2)$$
$$\times\ \frac{\sin\frac{d\pi}{2}}{2-d}\,e^{w[\tau p - \beta'/2(\beta'+1)]}, \quad (53)$$

where the term $\sin(d\pi/2)/(2-d)$ goes to $\pi/2$ as $d \to 2$. The τ-dependent LDF is given as

$$h_\tau(p) = h_l(p) - \frac{d-2}{2w\tau}\ln\tau - \frac{r_l(p)}{w\tau}, \quad (54)$$

where $h_l(p) = -p$ and $r_l(p)$ comes from the $\mathcal{O}(1)$ terms.

For $|p+1|\tau^{1/2} \ll 1$ (a narrow scaling region between the center and the left wing), we find

$$P_\tau(w\tau p) = \frac{2\beta'^{d/2}(w\tau)^{(d-2)/4}}{\pi(1+\beta')^{d/2}}\Gamma(3/2-d/4)$$
$$\times\ \frac{\cos\frac{d\pi}{4}}{2-d}\,e^{w[\tau p - \beta'/2(\beta'+1)]}, \quad (55)$$

where the term $\cos(d\pi/4)/(2-d)$ goes to $\pi/4$ as $d \to 2$. The τ-dependent LDF is

$$h_\tau(p) = h_l(p) - \frac{d-2}{4w\tau}\ln\tau - \frac{r_{l,c}(p)}{w\tau}, \quad (56)$$

where $h_l(p)$ is the same as that in Eq. (54) and $r_{l,c}(p)$ also comes from the $\mathcal{O}(1)$ terms.

3. *Right wing of the PDF*

In the right side of the central region ($p \gtrsim 2\beta' + 1$), the saddle point approaches the singularity at $\lambda = \lambda_2$ $(= -\beta')$. In this case, we have an additional complication due to the

essential singularity in the prefactor. Let us write $\delta\lambda_2 = \lambda^* - \lambda_2$ (> 0). The saddle-point equation (43) is expanded in terms of $\delta\lambda_2$ as

$$-2\delta\lambda_2 - (p - 1 - 2\beta') + \frac{d}{2w\tau\delta\lambda_2} + \frac{\beta'^2}{2\tau\delta\lambda_2^2} \approx 0. \quad (57)$$

Compared to Eq. (47), it contains a more divergent (fourth) term for finite β' and leads to different scaling behavior of $\delta\lambda_2$. [The case for $\beta' \ll (\delta\lambda_2)^{1/2}$ will be discussed in the next subsection].

For $[p - (2\beta' + 1)]\tau^{1/3} \gg 1$, we get

$$\delta\lambda_2 \approx \frac{\beta'}{\sqrt{2[p - (2\beta' + 1)]}}\tau^{-1/2}, \quad (58)$$

which determines the PDF in the region of $p > 2\beta' + 1$.

For $|p - (2\beta' + 1)|\tau^{1/3} \ll 1$, we get

$$\delta\lambda_2 \approx \left(\frac{\beta'^2}{4}\right)^{1/3} \tau^{-1/3}, \quad (59)$$

which determines the PDF in a narrow region around $p = 2\beta' + 1$ between the center and the right wing of the PDF.

Similar to the left wing, by expanding $H(\lambda, p)$ in powers of $\delta\lambda_2$ and using a new variable $v = 1 + (\lambda - \lambda^*)/\delta\lambda_2$, we find

$$P_\tau(w\tau p) = C_2 \int_{1-i\infty}^{1+i\infty} dv \frac{e^{w\tau[(p-2\beta'-1)\delta\lambda_2 v + \delta\lambda_2^2 v^2]}}{v^{d/2}} \times e^{w\beta'^2/(2\delta\lambda_2 v)}, \quad (60)$$

where

$$C_2 = \frac{1}{2\pi i}\frac{\beta'^{d/2}\delta\lambda_2^{1-d/2}}{(1+\beta')^{d/2}} e^{w\tau[\beta'(1+\beta') - \beta' p]} e^{-w[(5/2)\beta' + 2\beta'^2]}. \quad (61)$$

Note that the integrand in Eq. (60) has an exponentially diverging term near $v = 0$, which makes useless the previous contour deformation in the left wing in this case. This makes it difficult to evaluate the integral systematically. Thus, we try to employ again the saddle point method to evaluate this integral up to $\mathcal{O}(\tau^{-1})$.

First, for $[p - (2\beta' + 1)]\tau^{1/3} \gg 1$, we plug $\delta\lambda_2$ given in Eq. (58) into the integrand of Eq. (60). Then, the integral without the multiplicative constant C_2 can be written as

$$I_a = \int_{1-i\infty}^{1+i\infty} dv \frac{e^{w\beta'\sqrt{\tau(p-2\beta'-1)/2}(v+1/v) + [w\beta'^2/2(p-2\beta'-1)]v^2}}{v^{d/2}}. \quad (62)$$

Since $\sqrt{\tau(p - 2\beta' - 1)} \gg \tau^{1/3}$, one can use the saddle-point approximation for the integral. The saddle point is approximately determined from $\frac{d}{dv}(v + 1/v) = 0$ (the second term in the exponent is much smaller than the first one), yielding $v^* \approx 1$. This saddle point is far from the singularity at $v = 0$, so the conventional saddle point integral is sufficient. The curvature proportional to $\frac{d^2}{d^2 v}(v + 1/v)|_{v=1} = 2$ is positive, so the steepest descent path is coincident with the original contour. As a result, we find

$$I_a = \frac{i2^{1/4}\sqrt{\pi}e^{w\beta'\sqrt{2\tau(p-2\beta'-1)} + w\beta'^2/2(p-2\beta'-1)}}{(w\beta')^{1/2}[\tau(p - 2\beta' - 1)]^{1/4}}. \quad (63)$$

Multiplying it by C_2, we get

$$P_\tau(w\tau p) = \sqrt{\frac{\beta'}{\pi w}} \frac{2^{(d-5)/4} e^{-w\beta'[\tau(p-\beta'-1) - \sqrt{2\tau(p-2\beta'-1)}]}}{(1+\beta')^{d/2}[\tau(p - 2\beta' - 1)]^{(3-d)/4}} \times e^{-(w\beta'/2)[5+4\beta' - \beta'/(p-2\beta'-1)]}. \quad (64)$$

The τ-dependent LDF is given as

$$h_\tau(p) = h_r(p) - \frac{\beta'\sqrt{2(p - 2\beta' - 1)}}{\sqrt{\tau}} - \frac{d-3}{4w\tau}\ln\tau - \frac{r_r(p)}{w\tau}, \quad (65)$$

where $h_r(p) = \beta' p - \beta'(1 + \beta')$ and $r_r(p)$ comes from the $\mathcal{O}(1)$ terms.

Next, for $|p - 2\beta' - 1|\tau^{1/3} \ll 1$, $\delta\lambda_2 \sim \tau^{-1/3}$ as in Eq. (59). Again, by the power counting, one can easily simplify the integral in Eq. (60) without C_2 as

$$I_b = \int_{1-i\infty}^{1+i\infty} dv \frac{e^{w\tau\delta\lambda_2^2 v^2 + w\beta'^2/(2\delta\lambda_2 v)}}{v^{d/2}}. \quad (66)$$

A nuisance comes in when we calculate the LDF exactly up to $\mathcal{O}(\tau^{-1})$ [or $\ln P_\tau$ up to $\mathcal{O}(1)$] because higher-order expansions are needed for $\delta\lambda_2$ in a very narrow region like $|p - 2\beta' - 1| \sim \tau^{-\alpha}$ with $1/3 \leqslant \alpha \leqslant 2/3$. In fact, we need to divide this region into infinitely many intervals in order to calculate the finite-time correction to the LDF exactly up to $\mathcal{O}(\tau^{-1})$. This can be done with a straightforward calculation in principle, but requires a lengthy one, involving high-order calculations of $\delta\lambda_2$ from Eq. (57).

In this paper, we consider only the simplest case of $|p - 2\beta' - 1|\tau^{2/3} \ll 1$. Then, both terms in the exponent of Eq. (66) scale as $\sim\tau^{1/3}$ and the saddle point is determined by $\frac{d}{dv}(v^2 + 2/v) = 0$, which gives $v^* \approx 1$. The curvature is also positive, so the steepest path is again coincident with the original. As a result, we get

$$I_b = i\sqrt{\frac{2^{4/3}\pi}{3w\tau^{1/3}\beta'^{4/3}}} e^{3w\tau^{1/3}(\beta'/2)^{4/3}}. \quad (67)$$

Multiplying it by C_2, we obtain

$$P_\tau(w\tau p) = \frac{2^{(d-3)/3}\tau^{(d-3)/6}\beta'^{d/6}}{\sqrt{3\pi w}(1+\beta')^{d/2}} e^{-w\beta'\tau(p-\beta'-1)} \times e^{3w\tau^{1/3}(\beta'/2)^{4/3} - (w\beta'/2)(5+4\beta')}. \quad (68)$$

The τ-dependent LDF is

$$h_\tau(p) = h_r(p) - \frac{3}{\tau^{2/3}}\left(\frac{\beta'}{2}\right)^{4/3} - \frac{d-3}{6w\tau}\ln\tau - \frac{r_{r,c}(p)}{w\tau}, \quad (69)$$

where $h_r(p)$ is the same as that in Eq. (65) and $r_{r,c}(p)$ also comes from the $\mathcal{O}(1)$ terms. An extension to higher dimensions ($d \geqslant 4$) is straightforward.

4. FT violations

We examine the detailed FT for the heat by varying the initial temperature β'. We present $f_\tau(p)$ defined in Eq. (37) such that $f_\tau(p) = -h_\tau(p) + h_\tau(-p)$. All the results in this

subsection are summarized into

$$f_\tau(p) = \begin{cases} p + \frac{1}{\tau}\left[\frac{2\beta'^2 p}{(2\beta'+1-p)(2\beta'+1+p)} + \frac{d}{2w}\ln\frac{(1-p)(2\beta'+1+p)}{(1+p)(2\beta'+1-p)}\right] & \text{for } 0 \leqslant p < 1, \\ p - \frac{(1-p)^2}{4} - \frac{d}{4w\tau}\ln\tau + \frac{r_c(p)-r_{l,c}(-p)}{w\tau} & \text{for } |p-1| \ll \tau^{-1/2}, \\ p - \frac{(1-p)^2}{4} - \frac{d-1}{2w\tau}\ln\tau + \frac{r_c(p)-r_l(-p)}{w\tau} & \text{for } 1 < p < 2\beta'+1, \\ (1-\beta')p + \beta'(1+\beta') + \frac{3}{\tau^{2/3}}\left(\frac{\beta'}{2}\right)^{4/3} - \frac{2d-3}{6w\tau}\ln\tau + \frac{r_{r,c}(p)-r_l(-p)}{w\tau} & \text{for } |p-2\beta'-1| \ll \tau^{-2/3}, \\ (1-\beta')p + \beta'(1+\beta') + \beta'\sqrt{\frac{2(p-2\beta'-1)}{\tau}} - \frac{d-1}{4w\tau}\ln\tau + \frac{r_r(p)-r_l(-p)}{w\tau} & \text{for } p-2\beta'-1 \gg \tau^{-1/3}, \end{cases} \tag{70}$$

which converge to Eq. (38) for large τ with various finite-time corrections.

Inside of the region I ($|p| < 1$), the detailed FT is violated for finite τ by the amount of $\mathcal{O}(1/\tau)$, and the FT is restored ($f_\infty(p) = p$) in the infinite-τ limit. In all other regions, the FT is violated even in the infinite-τ limit. We present the figures for $f_\tau(p)$, Fig. 3 for $\beta' = 0.1$ and Fig. 4 for $\beta' = 2$. They show similar trends to $f_\infty(p)$ as in Fig. 2. The FT holds approximately well only in the central region (I).

C. FT in the $\beta' \to 0$ limit

In this subsection, we establish the transient detailed FT for the heat in general, from the standard stochastic thermodynamics where the time-integrated quantities are defined at the level of dynamic trajectories [16,17,36].

A trajectory starting from $t = 0$ to τ, is denoted by $\mathbf{q}(t) \equiv \{q_t; t \in [0,\tau]\}$ with a set of state variables q_t. The probability to find a trajectory $\mathbf{q}$ in a given dynamic process can be written as

$$P(\mathbf{q}) = P_0(q_0)\Pi(\mathbf{q}|q_0), \tag{71}$$

where $P_0(q_0)$ is the probability distribution of the initial state q_0 and $\Pi(\mathbf{q}|q_0)$ is the conditional probability for the trajectory $\mathbf{q}$ starting from q_0.

We also define the time-reverse trajectory $\mathbf{q}^\dagger$ with $\mathbf{q}^\dagger(t) = \epsilon\mathbf{q}(\tau - t)$ with $\epsilon\mathbf{q}$ representing the mirrored trajectory with a parity ϵ for each state variable [37,38]. This trajectory starts at the mirrored state of the final state of the original trajectory: $q_0^\dagger = \epsilon q_\tau$. The trajectory probability for $\mathbf{q}^\dagger$ is similarly written as

$$P(\mathbf{q}^\dagger) = P_0(q_0^\dagger)\,\Pi(\mathbf{q}^\dagger|q_0^\dagger). \tag{72}$$

It is well known [16,17] that the heat production for a given trajectory $\mathbf{q}$ is identical to the logarithm of the ratio of two conditional probabilities as

$$\beta Q_\tau[\mathbf{q}] = \ln\frac{\Pi(\mathbf{q}|q_0)}{\Pi(\mathbf{q}^\dagger|q_0^\dagger)}, \tag{73}$$

where β is the inverse temperature of the heat bath.

By choosing various initial ensembles for the original and the time-reverse processes [$P_0(q_0)$ and $P_0(q_0^\dagger)$], one can derive FT's for different thermodynamic quantities. For example, when one chooses the initial distribution of the time-reverse process as the final distribution of the original process [$P_0(q_0^\dagger) = P_\tau(q_\tau)$], then the total entropy production summing the change of the Shannon entropy in the system and the heat production in the heat bath becomes simply a logarithm of the ratio of two trajectory probabilities such that $\Delta S_{\text{total}} = \ln[P(\mathbf{q})/P(\mathbf{q}^\dagger)]$. Since the ΔS_{total} is written as the logarithm of the two normalized probability distributions (a typical property of the relative entropy), the integral FT should hold for ΔS_{total} for any finite τ and any initial ensemble with $P_0(q_0)$ [16,17]. For the transient detailed FT, we need the so-called involution condition, which requires the steady-state initial ensemble.

It is also well known that the choice of the equilibrium Boltzmann ensembles as the initial ensembles for both the original and time-reverse processes yields the transient integral

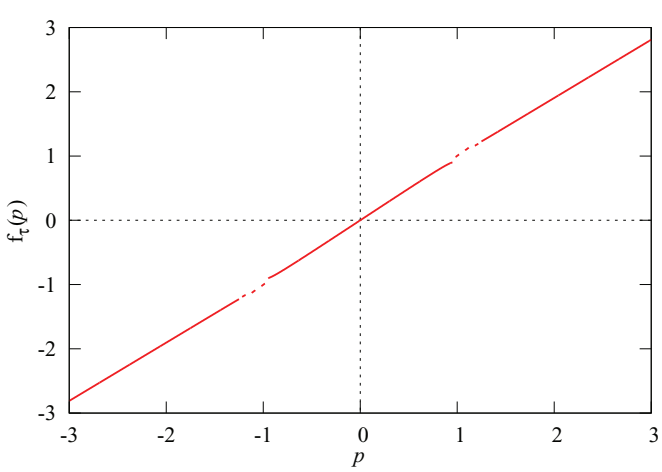

FIG. 3. (Color online) $f_\tau(p) = \frac{1}{w\tau}\ln\frac{P_\tau(w\tau p)}{P_\tau(-w\tau p)}$ is drawn for $\beta' = 0.1$, $\tau = 20$, and $d = 3$.

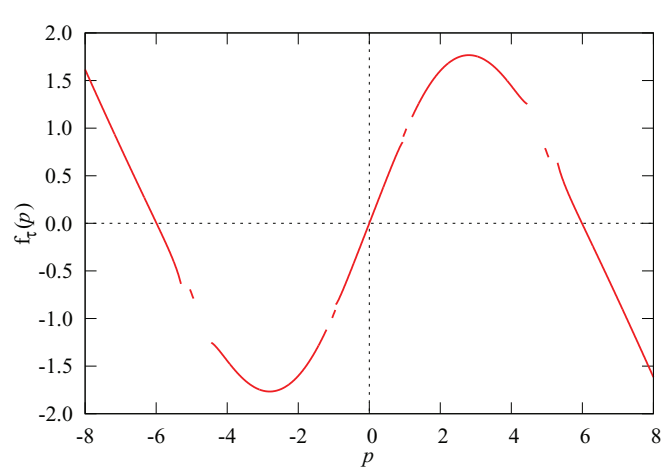

FIG. 4. (Color online) $f_\tau(p)$ is drawn for $\beta' = 2$, $\tau = 20$, and $d = 3$.

and detailed FT's for the work, where the involution condition is automatically satisfied with this choice.

We can also derive the FT for the heat in a similar manner by choosing the uniform (state-independent) distributions as the initial distributions for both processes. Then, the logarithm of the ratio of trajectory probabilities is simply the heat production as in Eq. (73), due to the cancellation of $P_0(q_0)$ and $P_0(q_0^\dagger)$. Since these initial distributions are obviously involutory to each other, not only the integral but also the detailed FT should hold for any finite τ. Even though the uniform distribution cannot be realized in the infinite-state space, one may consider it as the infinite-temperature ($\beta' \to 0$) limit of the Boltzmann distribution.

In the $\tau \to \infty$ limit, we have already shown that the detailed FT is satisfied by taking the $\beta' \to 0$ limit as in Eq. (38). However, the finite-time corrections in Eq. (70) seem to indicate that the $\beta' \to 0$ limit does not restore the FT for finite τ. This suggests the noncommutativity between the $\tau \to \infty$ limit and the $\beta' \to 0$ limit, which indeed turns out to be true.

All the complications come from the calculation of the PDF in the right wing. The saddle point equation in Eq. (57) has the β'-dependent divergent (fourth) term. In the case that β' is small and approaches zero for large τ such that $\beta' \ll (\delta\lambda_2)^{1/2}$, this fourth term can be ignored with respect to the third term. Then, all the subsequent calculations are very similar to those for the left wing of the PDF. The results are summarized below.

For $(p-1)\tau^{1/2} \gg 1$, $\delta\lambda_2 \approx \frac{d}{2w(p-1)}\tau^{-1}$ and

$$h_\tau(p) = h_r(p) - \frac{d-2}{2w\tau}\ln\tau - \frac{\tilde{r}_r(p)}{w\tau}, \tag{74}$$

with $h_r(p) = \beta' p - \beta'(1+\beta')$. Note that $\tilde{r}_r(p) = r_l(-p)$ in Eq. (54) for $\beta' \ll (\delta\lambda_2)^{1/2}$.

For $|p-1|\tau^{1/2} \ll 1$, $\delta\lambda_2 \approx \frac{1}{2}\sqrt{d/w}\,\tau^{-1/2}$ and

$$h_\tau(p) = h_r(p) - \frac{d-2}{4w\tau}\ln\tau - \frac{\tilde{r}_{r,c}(p)}{w\tau}, \tag{75}$$

where $\tilde{r}_{r,c}$ is a p-independent constant and equal to $r_{l,c}$ in Eq. (56) for $\beta' \ll (\delta\lambda_2)^{1/2}$. In both cases, $\delta\lambda_2$ decays with τ; so does β'.

In the calculation of $f_\tau = -h_\tau(p) + h_\tau(-p)$, nice cancellations occur between the finite-time correction terms up to $\mathcal{O}(\tau^{-1})$, and the FT is fully restored in the region I ($|p| < 1$). However, in the other regions, we still have an extra term such as in $f_\tau(p) = -h_r(p) + h_l(-p) \approx p + \beta'(1-p)$. This extra term may still be bigger than $\mathcal{O}(\tau^{-1})$ with the τ-dependent β', satisfying the condition $\beta' \ll (\delta\lambda_2)^{1/2}$. Therefore, the full FT for finite τ should be restored only in the $\beta' \to 0$ limit before any large-τ limit is taken.

IV. LDF AND FT FOR WORK

The generating function $g_W(\lambda)$ for the work is Gaussian in λ without any singularity. Thus, its Fourier integration in Eq. (11) can be evaluated exactly. Using a scaled variable for the work $\tilde{p} \equiv W/(w\tau)$, we find an exact PDF including all the exponentially decaying terms as

$$P_\tau(w\tau\tilde{p}) = \frac{e^{-w[\tau(\tilde{p}-1)+1-\alpha]^2/2[2\tau-3+1/\beta'+\alpha b(\alpha)]}}{\{2\pi w[2\tau-3+1/\beta'+\alpha b(\alpha)]\}^{1/2}}, \tag{76}$$

where the exact generating function in Eq. (A7) was integrated with $\alpha = e^{-\tau}$ and $b(\alpha) = 2(2-1/\beta') - \alpha(1-1/\beta')$.

Then, the τ-dependent LDF is simply given by $h_\tau(\tilde{p}) = -(w\tau)^{-1}\ln P_\tau(w\tau\tilde{p})$ as

$$h_\tau(\tilde{p}) = h_W(\tilde{p}) + \frac{\ln\tau}{2w\tau} - \frac{r_W(\tilde{p})}{w\tau}, \tag{77}$$

where $h_W(\tilde{p}) = (\tilde{p}-1)^2/4$ and $r_W(\tilde{p})$ comes from the $\mathcal{O}(1)$ terms. The FT-examining function $f_\tau(\tilde{p})$ becomes

$$f_\tau(\tilde{p}) = \frac{2\tilde{p}(\tau-1+\alpha)}{2\tau-3+1/\beta'+\alpha b(\alpha)}. \tag{78}$$

At $\beta' = 1$, $b(\alpha) = 2$ and $f_\tau(\tilde{p}) = \tilde{p}$ exactly for any τ. Thus, the transient detailed FT holds exactly at $\beta' = 1$. For $\beta' \neq 1$, $f_\tau(\tilde{p}) \approx \tilde{p} + \mathcal{O}(\tau^{-1})$ for large τ, so the detailed FT is satisfied only in the $\tau \to \infty$ limit.

V. DISCUSSION

The memory of the initial state is shown not to vanish, but to remain in the rare events for the time-accumulated quantities such as heat and work. This interesting phenomenon is manifested in the form of PDF particularly in the tail region corresponding to the rare events. The FT for finite τ depends on the initial ensemble. For example, the work satisfies the transient detailed FT only with an initial equilibrium Boltzmann distribution, while the heat does only with an initial uniform distribution. Common sense suggests that the heat and work satisfy the detailed FT's simultaneously in the long-time limit, since both quantities are equivalent to each other on average. However, it turns out that the FT's for the heat and work in the long-time limit deviate in different ways. In this limit, the FT for work is satisfied with any initial ensemble, while the FT for heat is not valid except for the uniform initial ensemble. This discrepancy originates from the unboundedness of the (potential) energy fluctuations ΔU in the heat $Q = W - \Delta U$.

In this paper, we investigate the PDF's for the work and heat generated in the Brownian motion in a sliding harmonic potential with a general initial ensemble characterized by the Boltzmann distribution with the inverse temperature β', generally different from the inverse temperature β of the reservoir. The heat PDF is calculated analytically up to $\mathcal{O}(\tau^{-1})$ with the measuring time τ and the work PDF is obtained exactly for any finite τ.

We explicitly show that the transient detailed FT holds for the work only at $\beta' = \beta$, set to unity, and for the heat only at $\beta' = 0$, as expected. On the other hand, the detailed FT in the long-time limit holds at any β' for the work, but only at $\beta' = 0$ for the heat (one should be careful about the order of the two limiting procedures of $\beta' \to 0$ and $\tau \to \infty$). This is due to the presence of singularities in the boundary terms for the heat, which represents the persistence of the initial memory. Physically, it can be argued that the highly energetic particles (high U) in the initial ensemble dominantly contribute to the events of positive large heat production ($Q \gg 1$) by losing its energy through dissipation [14]. This is why the right wing of the heat PDF depends strongly on the initial temperature β', but its left wing depends on it only very weakly. It is interesting to note that there is no threshold value of β' for the dominance

of the initial ensemble, in contrast to other cases where a finite critical threshold is found [14].

It may be an interesting task to find a systematic deviation of the FT for time-integrated quantities with an arbitrary initial ensemble, for example, by generalizing the recent study on the relation of heat fluctuations in Ref. [15]. It is also interesting to apply our method to other solvable nonequilibrium systems, such as a linear diffusion system with a nonconservative force [35,39] and a motion under a breathing harmonic potential [40].

ACKNOWLEDGMENTS

This work was supported by the EDISON program through NRF Grant No. 2012M3C1A6035307 (K.K.) and also by the Basic Science Research Program through NRF Grants No. 2013R1A1A2011079 (C.K.) and No. 2013R1A1A2A10009722 (H.P.). We thank Korea Institute for Advanced Study for providing computing resources (KIAS Center for Advanced Computation Abacus) for this work. H.P. also thanks the Galileo Galilei Institute for Theoretical Physics for the hospitality and the INFN for partial support during the completion of this work.

APPENDIX: GENERATING FUNCTIONS

The exact formula for the generating function for the heat is given by

$$g_Q(\lambda) = \left(\frac{\beta'}{\beta'+\lambda}\right)^{d/2} \frac{1}{C^{d/2}}\, e^{-wG_Q(\lambda,\tau)} \tag{A1}$$

with

$$\begin{aligned} G_Q(\lambda,\tau) &= \tau\lambda(1-\lambda) - \frac{3}{2}\lambda + \frac{\lambda^2}{2C}(1-\lambda)(4-B) \\ &\quad + \alpha\left[2\lambda - \frac{2\lambda^2}{C}(1-\lambda)(2-B)\right] \\ &\quad + \alpha^2\left\{-\frac{\lambda}{2} + \frac{\lambda^2}{2C}[7-6B-(4-3B)\lambda]\right\} \\ &\quad - \alpha^3\frac{2\lambda^2}{C}(1-B) + \alpha^4\frac{\lambda^2}{2C}(1-B), \end{aligned} \tag{A2}$$

where

$$B = (\beta'+\lambda)^{-1}, \quad C = 1-\lambda+\alpha^2\lambda(1-B), \quad \alpha = e^{-\tau}. \tag{A3}$$

By setting $\alpha = 0$ (in the long-time limit), we get Eq. (26) in Sec. II B. At $\beta' = \beta = 1$, we find

$$g_Q(\lambda) = \frac{e^{-w\lambda(1-\lambda)\left[\tau+[(1-\alpha)/\tilde{C}][2\lambda^2(1-\alpha)-(1/2)(3-\alpha)]\right]}}{\tilde{C}^{d/2}}, \tag{A4}$$

with $\tilde{C} = 1-\lambda^2+\lambda^2\alpha^2$. Note that our result is slightly different from that in Ref. [29].

In the limit of $\beta' = 0$, $g_Q(\lambda)$ vanishes as $\sim\beta'^{d/2}$. However, note that its amplitude in this limit

$$g_Q(\lambda)\,\beta'^{-d/2} = \frac{e^{-w\lambda(1-\lambda)[\tau-2(1-\alpha)/(1+\alpha)]}}{[\lambda(1-\lambda)(1-\alpha^2)]^{d/2}}, \tag{A5}$$

perfectly satisfies the GC symmetry for any α (a finite time).

The generating function for the work is given by

$$g_W(\lambda) = e^{-wG_W(\lambda,\tau)} \tag{A6}$$

with

$$\begin{aligned} G_W(\lambda,\tau) &= \tau\lambda(1-\lambda) - \lambda + \frac{\lambda^2}{2}\left(3-\frac{1}{\beta'}\right) \\ &\quad + \alpha\left[\lambda - \lambda^2\left(2-\frac{1}{\beta'}\right)\right] + \alpha^2\frac{\lambda^2}{2}\left(1-\frac{1}{\beta'}\right). \end{aligned} \tag{A7}$$

By setting $\alpha = 0$, we get Eq. (31) in Sec. II C. At $\beta' = \beta = 1$, we find

$$g_W(\lambda) = e^{-w\lambda(1-\lambda)(\tau-1+\alpha)}, \tag{A8}$$

which agrees with the result in Ref. [29] and satisfies perfectly the GC symmetry for any α (a finite time).

[1] D. J. Evans, E. G. D. Cohen, and G. P. Morriss, Phys. Rev. Lett. **71**, 2401 (1993).
[2] D. J. Evans and D. J. Searles, Phys. Rev. E **50**, 1645 (1994).
[3] G. Gallavotti and E. G. D. Cohen, Phys. Rev. Lett. **74**, 2694 (1995); J. Stat. Phys. **80**, 931 (1995).
[4] J. Kurchan, J. Phys. A: Math. Gen. **31**, 3719 (1998).
[5] G. E. Crooks, Phys. Rev. E **60**, 2721 (1999).
[6] J. L. Lebowitz and H. Spohn, J. Stat. Phys. **95**, 333 (1999).
[7] C. Jarzynski, Phys. Rev. Lett. **78**, 2690 (1997); Phys. Rev. E **56**, 5018 (1997).
[8] G. M. Wang, E. M. Sevick, E. Mittag, D. J. Searles, and D. J. Evans, Phys. Rev. Lett. **89**, 050601 (2002).
[9] D. M. Carberry, J. C. Reid, G. M. Wang, E. M. Sevick, D. J. Searles, and D. J. Evans, Phys. Rev. Lett. **92**, 140601 (2004).
[10] F. Douarche, S. Joubaud, N. B. Garnier, A. Petrosyan, and S. Ciliberto, Phys. Rev. Lett. **97**, 140603 (2006).
[11] J. R. Gomez-Solano, L. Bellon, A. Petrosyan, and S. Ciliberto, Europhys. Lett. **89**, 60003 (2010).
[12] S. Ciliberto, S. Joubaud, and A. Petrosyan, J. Stat. Mech. (2010) P12003.
[13] H. Park (unpublished).
[14] J. S. Lee, C. Kwon, and H. Park, Phys. Rev. E **87**, 020104(R) (2013).
[15] J. D. Noh and J.-M. Park, Phys. Rev. Lett. **108**, 240603 (2012).
[16] U. Seifert, Phys. Rev. Lett. **95**, 040602 (2005).
[17] M. Esposito and C. Van den Broeck, Phys. Rev. Lett. **104**, 090601 (2010).
[18] H. Touchette, Phys. Rep. **478**, 1 (2009).

[19] *Large Deviations in Physics*, Lecture Notes in Physics Vol. 885, edited by A. Vulpiani, F. Cecconi, M. Cencini, A. Puglisi, and D. Vergni (Springer, New York, 2014).

[20] T. Bodineaua and B. Derrida, C. R. Phys. **8**, 540 (2007).

[21] D. Lacoste, A. W. C. Lau, and K. Mallick, Phys. Rev. E **78**, 011915 (2008).

[22] K. Saito and A. Dhar, Phys. Rev. Lett. **99**, 180601 (2007).

[23] J. S. Lee, C. Kwon, and H. Park, J. Stat. Mech. (2013) P11002.

[24] $A = \beta(W - \Delta F)$ being taken, the GC symmetry is exact for any finite τ with the initial Boltzmann ensemble characterized by the inverse temperature β of the heat bath.

[25] G. M. Wang, J. C. Reid, D. M. Carberry, D. R. M. Williams, E. M. Sevick, and D. J. Evans, Phys. Rev. E **71**, 046142 (2005).

[26] G. Pesce, G. Volpe, A. Imparato, G. Rusciano, and A. Sasso, J. Optics **13**, 044006 (2011).

[27] The free energy difference is $\Delta F = 0$ in this simply driven system.

[28] R. van Zon and E. G. D. Cohen, Phys. Rev. E **67**, 046102 (2003).

[29] R. van Zon and E. G. D. Cohen, Phys. Rev. Lett. **91**, 110601 (2003); Phys. Rev. E **69**, 056121 (2004).

[30] J. Farago, J. Stat. Phys. **107**, 781 (2002); Physica A **331**, 69 (2004).

[31] P. Visco, J. Stat. Mech. (2006) P06006.

[32] A. Puglisi, L. Rondoni, and A. Vulpiani, J. Stat. Mech. (2006) P08010.

[33] A. Pal and S. Sabhapandit, Phys. Rev. E **87**, 022138 (2013).

[34] In the discrete-time representation, $\int_0^\tau dt\mathcal{L}$ is replaced with $-\frac{1}{4D}\sum_{i=1}^{N}(\frac{\mathbf{X}_i-\mathbf{X}_{i-1}}{\Delta t} + \tau_r^{-1}\mathbf{X}_i^{(\alpha)})^2 - \alpha d\tau_r^{-1}\tau$, where $\mathbf{X}_i^{(\alpha)} = \alpha\mathbf{X}_i + (1-\alpha)\mathbf{X}_{i-1}$ for $\alpha = 0$ (prepoint, Ito), $\alpha = 1/2$ (midpoint, Stratonovich), and $\alpha = 1$ (postpoint). The first term is a contribution from the Lagrangian, while the second term should be kept for a specific representation of α. The simplest one is $\alpha = 0$, but the path integral is independent of α.

[35] C. Kwon, J. D. Noh, and H. Park, Phys. Rev. E **83**, 061145 (2011).

[36] K. Sekimoto, Prog. Theor. Phys. Suppl. **130**, 17 (1998).

[37] R. E. Spinney and I. J. Ford, Phys. Rev. Lett. **108**, 170603 (2012); Phys. Rev. E **85**, 051113 (2012); I. J. Ford and R. E. Spinney, *ibid.* **86**, 021127 (2012).

[38] H. K. Lee, C. Kwon, and H. Park, Phys. Rev. Lett. **110**, 050602 (2013).

[39] J. D. Noh, C. Kwon, and H. Park, Phys. Rev. Lett. **111**, 130601 (2013).

[40] C. Kwon, J. D. Noh, and H. Park, Phys. Rev. E **88**, 062102 (2013).