Nonequilibrium Microcanonical and Canonical Ensembles and Their Equivalence

Raphaël Chetrite[1,*] and Hugo Touchette[2,3,†]

[1]*Laboratoire J. A. Dieudonné, UMR CNRS 6621, Université de Nice Sophia-Antipolis, Nice 06108, France*
[2]*National Institute for Theoretical Physics (NITheP), Stellenbosch 7600, South Africa*
[3]*Institute of Theoretical Physics, University of Stellenbosch, Stellenbosch 7600, South Africa*

Generalizations of the microcanonical and canonical ensembles for paths of Markov processes have been proposed recently to describe the statistical properties of nonequilibrium systems driven in steady states. Here, we propose a theory of these ensembles that unifies and generalizes earlier results and show how it is fundamentally related to the large deviation properties of nonequilibrium systems. Using this theory, we provide conditions for the equivalence of nonequilibrium ensembles, generalizing those found for equilibrium systems, construct driven physical processes that generate these ensembles, and rederive in a simple way known and new product rules for their transition rates. A nonequilibrium diffusion model is used to illustrate these results.

Equilibrium properties of many-particle systems are determined by having recourse to statistical ensembles, such as the microcanonical ensemble for systems with constant energy or the canonical ensemble for systems at constant temperature [1]. For nonequilibrium systems modeled with Markov processes, similar ensembles can be constructed by defining probability distributions on the trajectories or paths of these systems, which are either conditioned on some constraints, as in the microcanonical ensemble, or have the form of a Gibbs distribution, as in the canonical ensemble. Such nonequilibrium path ensembles have been considered, in particular, by Evans [2–4] in the context of sheared fluids, and have come to play an important role in recent studies of the glass transition, e.g., in Lennard-Jones liquids [5–7], and dynamical phase transitions in kinetically constrained models [8–11] and quantum systems [12–15].

It is known from [2–4] that microcanonical path ensembles can be generated in the long-time limit by a Markov process whose transition rates are those of the original process, modified by a factor taking into account the microcanonical constraint. This result, extended to canonical path ensembles by Jack and Sollich [16], is important physically as it shows that path ensembles can be generated by an effective or driven process. However, as of now, it is not clear for which systems this result holds and how a microcanonical driven process is related to a canonical one. These problems relate to the issue of ensemble equivalence, which is fundamental in equilibrium statistical mechanics [17].

Here, we present a systematic approach to path ensembles that addresses these problems and show how they are fundamentally related, via the theory of large deviations [18], to dynamical fluctuations of nonequilibrium systems. Our theory contains many results obtained before, but also simplifies and generalizes them in many ways. In particular, we are able to treat an important class of nonequilibrium systems not covered in previous studies,

namely, diffusions, and to provide a direct proof of known constraint rules satisfied by driven processes [19–21] in addition to deriving new ones. From the connection with large deviations, we also state explicit conditions for the equivalence of microcanonical and canonical path ensembles, similar to those of equilibrium systems, indicating that these ensembles are not always equivalent. To highlight the case of diffusions, we illustrate our results for a driven periodic Langevin equation, used, for example, to model manipulated colloidal particles [22].

We consider, as a general model of physical systems driven in nonequilibrium steady states, an ergodic Markov process X_t evolving over a time $t \in [0, T]$. The specific form of this process depends on the system considered. It can be a jump process if one considers discrete-state processes, such as noisy chemical and biological reactions [22] or interacting particles evolving on a lattice [23]. Alternatively, it can be a diffusion, or a set of coupled stochastic differential equations in general, if one considers systems such as colloidal particles immersed in liquids [22], stochastic thermostated systems [24], or coupled noisy oscillators. In both cases, the effect of external reservoirs and fields can be modeled in terms of boundary conditions or directly at the level of transition rates, drift, and diffusion coefficients.

The statistical properties of X_t are determined by its Markov generator L, which governs the evolution of functions of X_t or by its adjoint $L^\dagger$, which governs the evolution of the probability density function (PDF) $p(x, t)$. Formally, we can also characterize X_t, following the path integral approach to Markov processes, by the functional PDF $P[x]$ of the paths $\{x(t)\}_{t=0}^{T}$, which enters in the calculation of expectations of the form

$$\langle A_T \rangle = \int \mathcal{D}[x] P[x] A_T[x], \qquad (1)$$

where $\mathcal{D}[x]$ denotes the path integral element and $A_T[x]$ is a general observable of X_t. Physical examples of observables include particle and energy currents, shear induced by external forces, or the work performed to drive a process in a nonequilibrium steady state.

In general, one is interested in studying not just the average value of an observable A_T, but also its typical (i.e., most probable) value and fluctuations determined by its PDF $P(A_T = a)$. Following the case of equilibrium systems, it is also of interest to study the properties of a process X_t when some observable acting as a constraint is varied. For example, one can study how a fluid initially at equilibrium goes to a nonequilibrium state under shearing [2–4] or how the properties of a glassy system vary as a function of its dynamical activity [5–7]. As another example recently considered in [25–27], one can study how the stationary distribution of a nonequilibrium interacting particle system changes as its current J_T integrated over a time T is observed to be far from its typical value.

In terms of probabilities, these situations are described by conditioning the path PDF $P[x]$ of X_t on the constraint or fluctuation $A_T = a$ observed

$$P^a[x] = P[x|A_T = a] = \frac{P[x, A_T = a]}{P(A_T = a)}. \quad (2)$$

This defines, similarly to equilibrium, a microcanonical path ensemble, also referred to as a conditional or constrained ensemble [2–4]. Alternatively, we can follow Gibbs and define a canonical path ensemble by

$$P_k[x] = \frac{e^{TkA_T[x]}P[x]}{W_T(k)}, \qquad W_T(k) = \langle e^{TkA_T} \rangle, \quad (3)$$

which replaces the constraint $A_T = a$ by an exponential involving the parameter k conjugated to A_T [28]. In other works, P_k is also referred to as the *biased*, *tilted*, or *s ensemble* [16,29,30]. The analogy with equilibrium ensembles is obvious if one views paths as microstates and A_T as the energy. From this, we see that the normalization constant $W_T(k)$ is the analog of the partition function.

Unlike its equilibrium counterpart, the canonical path ensemble does not always have a physical interpretation in terms of heat baths or driving fields. However, from a theoretical point of view, it is interesting to ask, as Gibbs did, whether P_k is equivalent to P^a, i.e., whether the properties of a system (equilibrium or stationary) in the canonical ensemble reproduce those of the microcanonical ensemble. This equivalence is known to hold for many equilibrium systems [17] and justifies using the canonical ensemble for actual calculations because of its simpler, unconstrained form. The same problem applies here: if we can demonstrate that P_k and P^a are equivalent, then any calculations of typical quantities involving P^a can be done with P_k. Proving equivalence also provides a way to physically obtain the canonical ensemble from the microcanonical ensemble by Gibbs conditioning.

Below, we show that equivalence holds in the stationary limit where $T \to \infty$, under some conditions related to the fluctuations of A_T. Moreover, we show that there exists a Markov process with generator L_k, different from L, whose path PDF is equivalent to P^a and P_k in the same limit. Physically, this means that the microcanonical and canonical path ensembles are generated by a specific process, called as before the driven process, involving additional forces compared to those driving X_t.

This is essentially the result of Evans mentioned earlier, with the difference that we now treat the problem of equivalence explicitly. Moreover, we obtain this result for general Markov processes, as well as a general class of observables that are (i) integrated in time, and (ii) depend on the state of the process and its transitions over time. For jump processes, these observables take the form

$$A_T = \frac{1}{T} \int_0^T f(X_t)dt + \frac{1}{T} \sum_{0 \leq t \leq T : \Delta X_t \neq 0} g(X_{t^-}, X_{t^+}), \quad (4)$$

where f and g are arbitrary functions and the sum is over all jumps $\Delta X_t = X_{t^+} - X_{t^-} \neq 0$ of the process. For diffusion processes, observables satisfying (i) and (ii) above are of the form

$$A_T = \frac{1}{T} \int_0^T f(X_t)dt + \frac{1}{T} \int_0^T g(X_t) \circ dX_t, \quad (5)$$

where $\circ$ denotes the Stratonovich integral [31]. These observables are very general: they include as special cases all the quantities mentioned so far (currents, shear, activity) in addition to quantities such as the integrated work, heat, and entropy production considered in the context of stochastic thermodynamics [32].

The fundamental assumption underlying our results is that the observable A_T satisfies a large deviation principle [18]. This means that the stationary PDF of A_T scales in the limit $T \to \infty$ as

$$P(A_T = a) = e^{-TI(a)+o(T)}, \quad (6)$$

where $I(a)$ is a function, called the rate function, that does not depend on T. This function plays a central role in equilibrium and nonequilibrium statistical physics [18], as it characterizes the typical values and fluctuations of A_T. In large deviation theory, it is obtained from the dominant eigenvalue Λ_k of the so-called tilted operator $\mathcal{L}_k$, whose form depends on the observable considered. For observables having the jump form of (4), we have [33]

$$\mathcal{L}_k = Le^{kg} + kf, \quad (7)$$

whereas for diffusion observables defined in (5), we have

$$\mathcal{L}_k = \hat{F}(\nabla + kg) + (\nabla + kg)\frac{D}{2}(\nabla + kg) + kf, \quad (8)$$

where $\hat{F}$ is the corrected drift entering in the symmetrized form of the generator of the stochastic differential equation considered and D is its diffusion matrix [34].

Assuming a large deviation principle for A_T, we can already solve the equivalence problem by appealing to general results of [17] stated for equilibrium ensembles, but which equally apply to nonequilibrium ensembles. The definition of equivalence used here is the following: we say that two path PDFs P and Q are asymptotically equivalent if

$$\lim_{T \to \infty} \frac{1}{T} \ln \frac{P[x]}{Q[x]} = 0 \qquad (9)$$

almost everywhere with respect to P. Essentially, this means that P and Q are equivalent if they are equal up to subexponential terms in T and for almost all paths.

This notion of equivalence is slightly stronger than the one used in [17], but also implies that the typical values of any observable obtained under P as $T \to \infty$ are the same as those obtained under Q in this limit. Moreover, the equivalence of P^a and P_k in the sense of (9) turns out to be determined, as in [17], by the convexity of the rate function $I(a)$ as follows: (i) If $I(a)$ is convex at the point a, then there exists a k such that P_k is asymptotically equivalent to P^a, which means, physically, that the two ensembles describe the same long-time stochastic dynamics. Moreover, if $I(a)$ is differentiable, then equivalence holds for $k = I'(a)$, providing a generalization of the temperature-entropy relation. (ii) If $I(a)$ is nonconvex at a, then there is no k for which P_k is asymptotically equivalent with P^a. In this case, the microcanonical process obtained by conditioning on $A_T = a$ has no canonical counterpart for any k.

The proof of these results, as well as the discussion of a third, more technical case of equivalence, known as partial equivalence, are deferred to a longer paper [35]. What is important to note here is that the equivalence of path ensembles is not guaranteed—it depends on the convexity of the rate function $I(a)$, in a way similar to equilibrium ensembles, whose equivalence depends on the concavity of the thermodynamic entropy [17]. In most works on path ensembles, this equivalence is assumed.

With this result, we now turn to the problem of generating or representing the microcanonical and canonical path ensembles by a Markov process. *A priori*, it is not clear that such a Markov representation exists, since neither P^a nor P_k is Markovian: the former has a conditioning that depends on the future time T, while the latter involves the constant $W_T(k)$ which breaks the multiplicative structure of the path PDF. However, as mentioned before, it is possible to construct a Markov process with generator L_k, whose path PDF is asymptotically equivalent to P_k, in the same sense as above, and which, in the case of ensemble equivalence, must therefore also be equivalent to P^a.

To construct this Markov process, we consider the following similarity transform of L

$$L^h = h^{-1} L h - h^{-1}(Lh), \qquad (10)$$

which is a generalization of Doob's h transform [36–38] defined for positive functions h [39]. For the purpose of

defining L_k, the main property of L^h that we need to know is that it is a Markov generator, and so defines a Markov process which has an explicit probability ratio with respect to the original path PDF $P[x]$. Combining the expression of this ratio, found in [38], with $\mathcal{L}_k$, it is then possible to construct the equivalent Markov process with generator L_k and path PDF $Q[x]$ in such a way that the limit of (9) with P/Q replaced by Q/P_k vanishes. The full derivation is deferred again to [34].

The final result is that L_k is unique and is obtained by applying the h transform to $\mathcal{L}_k$ with $h = r_k$, where r_k is the (right) eigenvector of $\mathcal{L}_k$ associated with the dominant eigenvalue Λ_k, so that

$$L_k = \mathcal{L}_k^{r_k}. \qquad (11)$$

Noticing that the similarity does not act on the diagonal part of $\mathcal{L}_k$ involving f, we can also write

$$L_k = r_k^{-1} \mathcal{L}_k|_{f=0} r_k + kf - \Lambda_k. \qquad (12)$$

This explicit expression for the generator of the Markov process asymptotically equivalent to the canonical path ensemble is the main result of this Letter. It applies to any Markov processes, including jump processes and diffusions, as well as any observables of these processes, provided that they satisfy a large deviation principle.

Before discussing a specific application, it is instructive to show that (12) recovers previous results obtained for physical jump processes characterized by jump rates $W(x, y)$ from state x to y. In this case, we obtain from (12) and (7) that the rates $W_k(x, y)$ of the driven process associated with the canonical path ensemble have the form

$$W_k(x, y) = r_k^{-1}(x) W(x, y) e^{kg(x,y)} r_k(y), \qquad (13)$$

which is the result obtained by Jack and Sollich [16] and essentially also Evans [2–4]. From this, we immediately obtain

$$W_k(x, y) W_k(y, x) = W(x, y) W(y, x) e^{k[g(x,y)+g(y,x)]}, \qquad (14)$$

which reduces to

$$W_k(x, y) W_k(y, x) = W(x, y) W(y, x) \qquad (15)$$

if g is an antisymmetric function. This identity relating the driven rates to the original rates of X_t is one of the product constraint rules recently derived in [19–21] by more involved methods. Another constraint found by these authors relates the escape rate $\nu(x) = (W1)(x)$ of X_t to the escape rate $\nu_k(x) = (W_k1)(x)$ of the driven process

$$\nu(x) - \nu(y) = \nu_k(x) - \nu_k(y). \qquad (16)$$

In our formalism, this exit rate constraint is simply obtained from (12) for $f = 0$ and any function g. In the case where $f \neq 0$, we obtain the more general identity,

$$\nu_k(x) - \nu(x) + kf(x) = \Lambda_k. \qquad (17)$$

To see how these results can be applied beyond jump processes, consider now a general d-dimensional diffusion having the (Itô) form

$$dX_t = F(X_t)dt + \sigma(X_t)dW_t, \qquad (18)$$

where F is a d-dimensional force field, σ is a $d \times d$ noise matrix, and dW_t is a d-dimensional Wiener noise [31]. Assuming that this process is ergodic, we can derive from its generator and (12) that the driven process generating the canonical path ensemble of X_t is a diffusion with the same noise as above but with the modified force

$$F_k = F + D(kg + \nabla \ln r_k), \qquad (19)$$

where $D = \sigma\sigma^T$ is the diffusion matrix. If the canonical path ensemble is equivalent to the microcanonical path ensemble, then the driven diffusion with the force F_k also generates the process X_t conditioned on the fluctuation $A_T = a$. As a consequence of this result, we see that the driven process associated with an equilibrium system [40] is itself an equilibrium system if g is gradient and is otherwise modified to a nonequilibrium system not satisfying detailed balance. If g is gradient, we also obtain from (19) the identity

$$\nabla \times (D^{-1}F_k) = \nabla \times (D^{-1}F) \qquad (20)$$

for two- and three-dimensional systems [41], which can be seen as a diffusive version of the jump process constraint rules.

To illustrate these results, let us consider the forced periodic diffusion evolving according to

$$\dot{\theta}(t) = \gamma - V'[\theta(t)] + \xi(t), \qquad \theta \in [0, 2\pi), \quad (21)$$

where $V = V_0 \cos(\theta)$, γ is a constant frequency driving the system in a nonequilibrium steady state breaking detailed balance, and $\xi(t)$ is a Gaussian white noise. This model arises in many physical applications, e.g., ion conduction, noisy Josephson junctions [42], in addition to Brownian motors and colloidal particles [22], and serves as an important model for testing new ideas and results about nonequilibrium systems; see, e.g., [43–45].

An important observable of this system is the velocity averaged over a time T

$$A_T = \frac{1}{T}\int_0^T d\theta(t) = \frac{\theta(T) - \theta(0)}{T}, \qquad (22)$$

which is proportional to the mean current. This is an observable having the form (5) with $f = 0$ and $g = 1$, for which the tilted generator is, according to (8),

$$\mathcal{L}_k = (\gamma - V' + k)\frac{d}{d\theta} + \frac{1}{2}\frac{d^2}{d\theta^2} + \frac{k^2}{2} + k(\gamma - V'). \quad (23)$$

The dominant eigenfunction $r_k(\theta)$ of this non-Hermitian operator is easily found by expanding the eigenvalue equation in Fourier modes. The effective force $F_k(\theta)$ of the driven process obtained from $r_k(\theta)$ is shown in Fig. 1
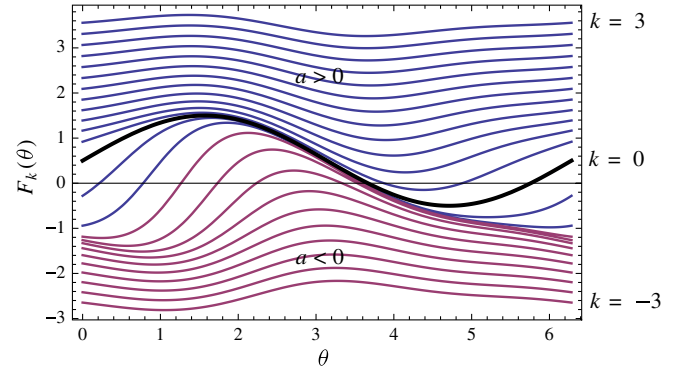

FIG. 1 (color online). Force $F_k(\theta)$ of the driven Markov process associated with the canonical version of the periodic diffusion with $V_0 = 1$ and $\gamma = 0.5$. The different curves are obtained for $k \in [-3, 3]$ in steps of 0.25. The force of the base process corresponding to $k = 0$ is shown in black. Blue curves yield positive stationary currents, $A_T = a > 0$, while purple curves yield negative stationary currents.

for $V_0 = 1$ and $\gamma = 0.5$, in which case the stationary (long-time) value of the current A_T is $a^* \approx 0.176$.

We see from the plot that the effective force of the driven process generating the canonical path ensemble is always periodic but changes in a nontrivial way with k. For $k > 0$, F_k drives the system to a positive stationary value of the current greater than a^*, while for $k < 0$, F_k drives the system to positive currents for k down to some value $k = 0.5$, below which the current becomes negative in the stationary limit. Thus, by choosing k properly, we can force the system to assume any stationary, nonequilibrium current. In the limit $k \to \pm\infty$, the constant k^2 term in the tilted generator dominates and leads to $F_k \approx k$, which means, physically, that the effective dynamics behind large currents in the canonical ensemble is a simple potential-free diffusion with a constant drive in the direction of the current. Since the rate function of A_T is known to be convex for this model, this also means that the periodic diffusion (21) is equivalent to a driven diffusion with constant frequency when it is conditioned microcanonically on large currents.

To conclude, we have presented a general theory of microcanonical and canonical path ensembles describing the typical states and fluctuations of Markovian systems driven in nonequilibrium steady states. This theory enables one to determine, similarly to equilibrium systems, when these two ensembles are equivalent (that is, when they describe the same stationary states) and shows how the nonequilibrium dynamics associated with each ensemble is generated physically by an effective Markov process representing a driven nonequilibrium system. It also shows in a clear and rigorous way that the properties of this driven process are intimately related to the large deviation properties of nonequilibrium systems. In the future, we will provide proofs of the results presented here, in addition to investigating new problems arising from these results—in

---

*Raphael.Chetrite@unice.fr

†htouchet@alum.mit.edu

[1] F. Reif, *Fundamentals of Statistical and Thermal Physics* (McGraw-Hill, New York, 1965).

[2] R. M. L. Evans, Phys. Rev. Lett. **92**, 150601 (2004).

[3] R. M. L. Evans, J. Phys. A **38**, 293 (2005).

[4] R. M. L. Evans, Contemp. Phys. **51**, 413 (2010).

[5] M. Merolle, J. P. Garrahan, and D. Chandler, Proc. Natl. Acad. Sci. U.S.A. **102**, 10837 (2005).

[6] L. O. Hedges, R. L. Jack, J. P. Garrahan, and D. Chandler, Science **323**, 1309 (2009).

[7] D. Chandler and J. P. Garrahan, Ann. Rev. Chem. Phys. **61**, 191 (2010).

[8] J. P. Garrahan, R. L. Jack, V. Lecomte, E. Pitard, K. van Duijvendijk, and F. van Wijland, Phys. Rev. Lett. **98**, 195702 (2007).

[9] J. P. Garrahan, R. L. Jack, V. Lecomte, E. Pitard, K. van Duijvendijk, and F. van Wijland, J. Phys. A **42**, 075007 (2009).

[10] R. L. Jack and J. P. Garrahan, Phys. Rev. E **81**, 011111 (2010).

[11] T. Speck and J. P. Garrahan, Eur. Phys. J. B **79**, 1 (2011).

[12] C. Ates, B. Olmos, J. P. Garrahan, and I. Lesanovsky, Phys. Rev. A **85**, 043620 (2012).

[13] J. P. Garrahan, A. D. Armour, and I. Lesanovsky, Phys. Rev. E **84**, 021115 (2011).

[14] S. Genway, J. P. Garrahan, I. Lesanovsky, and A. D. Armour, Phys. Rev. E **85**, 051122 (2012).

[15] J. M. Hickey, S. Genway, I. Lesanovsky, and J. P. Garrahan, Phys. Rev. A **86**, 063824 (2012).

[16] R. L. Jack and P. Sollich, Prog. Theor. Phys. Suppl. **184**, 304 (2010).

[17] H. Touchette, Europhys. Lett. **96**, 50010 (2011).

[18] H. Touchette, Phys. Rep. **478**, 1 (2009).

[19] A. Baule and R. M. L. Evans, Phys. Rev. Lett. **101**, 240601 (2008).

[20] A. Baule and R. M. L. Evans, J. Stat. Mech. (2010) P03030.

[21] A. Simha, R. M. L. Evans, and A. Baule, Phys. Rev. E **77**, 031117 (2008).

[22] U. Seifert, Rep. Prog. Phys. **75**, 126001 (2012).

[23] B. Derrida, J. Stat. Mech. (2007) P07023.

[24] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNola, and J. R. Haak, J. Chem. Phys. **81**, 3684 (1984).

[25] V. Popkov, G. M. Schütz, and D. Simon, J. Stat. Mech. (2010) P10007.

[26] V. Popkov and G. Schütz, J. Stat. Phys. **142**, 627 (2011).

[27] V. Belitsky and G. Schütz, J. Stat. Phys. **152**, 93 (2013).

[28] We use k rather than β to make clear that k is not the inverse temperature in general.

[29] V. Lecomte, C. Appert-Rolland, and F. van Wijland, Phys. Rev. Lett. **95**, 010601 (2005).

[30] V. Lecomte, C. Appert-Rolland, and F. van Wijland, J. Stat. Phys. **127**, 51 (2007).

[31] D. W. Stroock and S. R. S. Varadhan, *Multidimensional Diffusion Processes* (Springer, New York, 1979).

[32] K. Sekimoto, Lect. Notes Phys. **799**, 1 (2010).

[33] This follows by combining the Feynman-Kac and Girsanov formulae [31].

[34] All terms in these expressions are operators: Le^{kg} is the operator with components $(Le^{kg})(x, y) = L(x, y)e^{kg(x,y)}$, while kf is a diagonal operator with components $kf(x)\delta(x - y)$.

[35] R. Chetrite and H. Touchette (to be published).

[36] J. L. Doob, *Classical Potential Theory and Its Probabilistic Counterpart* (Springer, New York, 1984).

[37] K. L. Chung and J. B. Walsh, *Markov Processes, Brownian Motion, and Time Symmetry* (Springer, Berlin, 2005), 2nd ed.

[38] R. Chetrite and S. Gupta, J. Stat. Phys. **143**, 543 (2011).

[39] Multiplications in this expression have the same sense as in $\mathcal{L}_k$, except for (Lh) which stands for the application or action of L on h.

[40] An equilibrium system has a stationary distribution ρ, drift, and diffusion related by $F = (D/2)\nabla\ln\rho + (1/2)\nabla D$.

[41] A similar formula involving exterior derivatives holds for dimensions greater than 3.

[42] H. Risken, *The Fokker-Planck Equation: Methods of Solution and Applications* (Springer, Berlin, 1996), 3rd ed.

[43] V. Blickle, T. Speck, C. Lutz, U. Seifert, and C. Bechinger, Phys. Rev. Lett. **98**, 210601 (2007).

[44] J. R. Gomez-Solano, A. Petrosyan, S. Ciliberto, R. Chetrite, and K. Gawedzki, Phys. Rev. Lett. **103**, 040601 (2009).

[45] T. Nemoto and S.-I. Sasa, Phys. Rev. E **83**, 030105 (2011); **84**, 061113 (2011).