Fluctuation Theorem on a Riemannian Manifold

Yifan Cai,* Tao Wang,† and Liu Zhao‡
School of Physics, Nankai University, Tianjin 300071, China

(Dated: January 1, 2024)

Based on the covariant underdamped and overdamped Langevin equations with Stratonovich coupling to multiplicative noises and the associated Fokker-Planck equations on Riemannian manifold, we present the first law of stochastic thermodynamics on the trajectory level. The corresponding fluctuation theorems are also established, with the total entropy production of the Brownian particle and the heat reservoir playing the role of dissipation function.

Keywords: Langevin equation, Fokker-Planck equation, fluctuation theorem, Riemannian manifold

I. INTRODUCTION

Since the early 1990s, fluctuation theorems (FTs) [1–3] have played an indispensable role in understanding the origin of macroscopic irreversibility. Such theorems, often realized in the form of unequal probabilities for the forward and reversed processes, greatly helped in resolving the long lasting puzzles and debates regarding Boltzmann's H-theorem, known as Loschmidt paradox. Based on Sekimoto's work [4] on stochastic energetics, Seifert [5] was able to establish a version of FT associated to the stochastic trajectories described by overdamped Langevin equation (OLE), and subsequent works [6–8] extended the construction to the cases of various generalized forms of Langevin equation (LE).

Can we establish FTs on the trajectory level on curved Riemannian manifold? This is the question we wish to address in this work. In recent years, stochastic thermodynamics has gained increasing importance in understanding phenomena at the mesoscopic scale [9–13]. There are certain realistic scenarios, such as the diffusion of individual protein molecule on a biological membrane, which calls for a construction of LE and FT on Riemannian manifolds. Another motivation for the quest of stochastic thermodynamics and FT on Riemannian manifolds is to take it as a midway step towards general relativistic description of these fields, which is important because, in essence, every physics system must abide by the principles of relativity, whereas the spacetime symmetries in general relativity impose much stronger restrictions, making it harder to break the time reversal symmetry.

Historically, there have been some repeated attempts in the construction of LE on Riemannian manifolds in the mathematical [14–19] and physical [20, 21] literature. However, the first law and other thermodynamic relations were not considered in these works. Xing *et al.* [22, 23] explored an Ito-type nonlinear LE and established a FT on Riemannian manifold. Most of these works have interpreted the word "covariance" in the sense of second order (or *jet bundle*) geometry, which is different from the usual coordinate covariance in standard Riemannian geometry. Ref. [20] seems to be an exception. However, the difference between our work and [20] is acute: the LE in [20] is presented only in the configuration space, something like the OLE to be described in Section II B. However, as will be explained near the end of Section II B, the covariance of LE in [20] actually holds only in flat manifolds and only with respect to linear coordinate transformations, while our work covers both the configuration space (OLE) and phase space (ULE) descriptions, and our formalism is genuinely covariant under general coordinate transformation in the sense of first order geometry.

II. LANGEVIN AND FOKKER-PLANCK EQUATIONS ON RIEMANNIAN MANIFOLD

A. Underdamped case

In Ref. [24], the relativistic covariant underdamped Langevin equation (ULE) on *pseudo-Riemannian spacetime* is established. The same procedure can be used for constructing the LE for a point particle of mass $m > 0$ moving on d-dimensional *Riemannian space* M with positive definite metric $g_{\mu\nu}(x)$, so we directly present the result,

$$\mathrm{d}\tilde{x}^\mu_t = \frac{\tilde{p}^\mu_t}{m}\mathrm{d}t, \tag{1}$$

$$\begin{aligned}\mathrm{d}\tilde{p}^\mu_t = &\left[R^\mu{}_a \circ_S \mathrm{d}\tilde{w}^a_t + \frac{1}{2}R^\mu{}_a \frac{\partial}{\partial p^\nu} R^\nu{}_a \mathrm{d}t\right] \\ &- \frac{1}{m}K^\mu{}_\nu \tilde{p}^\nu_t \mathrm{d}t + f^\mu_{\mathrm{ex}}\mathrm{d}t - \frac{1}{m}\Gamma^\mu{}_{\alpha\beta}\tilde{p}^\alpha_t \tilde{p}^\beta_t \mathrm{d}t,\end{aligned} \tag{2}$$

where $R^\mu{}_a$ represent the stochastic amplitudes which may depend on $\tilde{x}^\mu$ and $\tilde{p}^\mu$, $\Gamma^\mu{}_{\alpha\beta}$ is the Christoffel connection associated with $g_{\mu\nu}(x)$, $K^\mu{}_\nu$ is the tensorial damping coefficient (referred to as damping tensor henceforth),

* caiyifan@mail.nankai.edu.cn
† taowang@mail.nankai.edu.cn
‡ lzhao@nankai.edu.cn

and $\mathrm{d}\tilde{w}_t^a$ are Gaussian noises with probability distribution functions (PDFs)

$$\Pr[\mathrm{d}\tilde{w}_t^a = \mathrm{d}w^a] = \frac{1}{(2\pi \mathrm{d}t)^{d/2}} \exp\left[-\frac{\delta_{ab}\mathrm{d}w^a \mathrm{d}w^b}{2\mathrm{d}t}\right]. \quad (3)$$

The symbol $\circ_S$ represents the Stratonovich coupling which ensures the chain rule in stochastic calculus. Greek indices $\mu, \nu, \cdots$ label spatial directions and latin indices $a, b, \cdots$ label independent noises, which are all running from 1 to d. Tilded variables like $\tilde{x}^\mu, \tilde{p}^\mu$ represent random variables, and the un-tilded symbols like x^μ, p^μ their realizations. In particular, (x^μ, p^μ) denotes the coordinate of the Brownian particle on TM in a concrete realization.

Some remarks are in due here.

(i) The LE described above is a system of stochastic differential equations on the tangent bundle TM regarded as the space of micro states of the Brownian particle. Normally, the space of micro states for a particle is taken to be the cotangent bundle T^*M. However, due to the non-degeneracy of the metric $g_{\mu\nu}$, the tangent and cotangent spaces are dual to each other, and the tangent and cotangent bundles can be used interchangeably (see [25–27] for explicit use of both approaches). Recall that TM is naturally equipped with the Sasaki metric [25, 26, 28]

$$\hat{g} := g_{\mu\nu}\mathrm{d}x^\mu \otimes \mathrm{d}x^\nu + g_{\mu\nu}\theta^\mu \otimes \theta^\nu,$$

where $\theta^\mu = \mathrm{d}p^\mu + \Gamma^\mu{}_{\alpha\beta}p^\alpha \mathrm{d}x^\beta$, together with the invariant volume element [here $g(x) := \det g_{\mu\nu}(x)$]

$$\mathrm{d}^{2d}X = g(x)\ \mathrm{d}x^1 \wedge \mathrm{d}x^2 \wedge ... \wedge \mathrm{d}x^d \wedge \mathrm{d}p^1 \wedge ... \wedge \mathrm{d}p^d.$$

As explained in [25, 26], the Sasaki metric is closely connected to the symplectic structure on TM.

Notice that the metric $g_{\mu\nu}(x)$ on M plays an indispensable role while obtaining the above Sasaki metric and also while describing the last term of eq. (2).

(ii) The choice for the stochastic amplitude $R^\mu{}_a$ is non-unique. Different choices correspond to different Langevin systems, and the result of this work should hold for all choices such that $R^\mu{}_a$ is an invertible matrix function which is differentiable in (x^μ, p^μ) and transforms as a vector for each fixed a.

(iii) Our approach to LE is the *traditional* one, as opposed to the more abstract nonlinear approach used in [22, 23]. Eq. (2) can be viewed as the geodesic equation supplemented by additional force terms, including a stochastic force

$$\xi^\mu := R^\mu{}_a \circ_S \mathrm{d}\tilde{w}_t^a/\mathrm{d}t + \frac{1}{2}R^\mu{}_a \frac{\partial}{\partial p^\nu} R^\nu{}_a,$$

a damping force $f^\mu_{\mathrm{dp}} := -K^\mu{}_\nu \tilde{p}^\nu_t/m$ and an external force f^μ_{ex}. The second term in the stochastic force is known as *additional stochastic force* [29, 30], which is required in order for the Brownian particle to be able to reach thermal equilibrium with the heat reservoir. In one-dimensional case, the stochastic force can also be expressed in the form of post-point rule $\xi = R \circ_p \mathrm{d}\tilde{w}_t/\mathrm{d}t$, hence some authors argued that the post-point rule is better suited for LE with multiplicative noises. However, in higher-dimensional cases, the post-point rule leads to a different result,

$$R^\mu{}_a \circ_p \mathrm{d}\tilde{w}_t^a/\mathrm{d}t = R^\mu{}_a \circ_S \mathrm{d}\tilde{w}_t^a/\mathrm{d}t + \frac{1}{2}R^\nu{}_a \frac{\partial}{\partial p^\nu} R^\mu{}_a.$$

(iv) The Stratonovich coupling maintains the chain rule, which ensures the covariance in the usual sense in Riemannian geometry. This makes an important difference from the previous works [16–19, 22, 23].

(v) Although eqs. (1)-(2) look the same as their relativistic counterparts [24], there are some essential differences. First, the time t used here is absolute, meaning that eqs. (1)-(2) are *non-relativistic*; Second, the heat reservoir hiding behind the stochastic and damping force terms is also non-relativistic, there is no need to worry about the relativistic effects such as the Tolman-Ehrenfest red shift; Last, the momentum space is flat, as opposed to the relativistic case.

The external force term depends on the position x of the Brownian particle and an external control parameter λ, and can be separated into conservative and non-conservative parts,

$$f^\mu_{\mathrm{ex}} = f^\mu_{\mathrm{con}} + f^\mu_{\mathrm{noc}} = -\nabla^\mu U(x, \lambda) + f^\mu_{\mathrm{noc}}.$$

There is some ambiguity in this decomposition, *e.g.*

$$f^\mu_{\mathrm{ex}} = (f^\mu_{\mathrm{con}} + \hat{f}^\mu) + (f^\mu_{\mathrm{noc}} - \hat{f}^\mu) = \hat{f}^\mu_{\mathrm{con}} + \hat{f}^\mu_{\mathrm{noc}},$$

where $\hat{f}^\mu$ is an arbitrary conservative force. This also leads to an ambiguity in the energy

$$E(x, p, \lambda) = \frac{1}{2m} g_{\mu\nu}(x) p^\mu p^\nu + U(x, \lambda)$$

of the Brownian particle. In the extremal case with $\hat{f}^\mu = -f^\mu_{\mathrm{con}}$, The energy will be consisted purely of the kinematic energy.

Since the Stratonovich coupling preserves the chain rule, we have

$$\begin{aligned} \mathrm{d}\tilde{E}_t &= \frac{\partial E}{\partial p^\mu}\mathrm{d}\tilde{p}^\mu_t + \frac{\partial E}{\partial x^\mu}\mathrm{d}\tilde{x}^\mu_t + \mathrm{d}_\lambda U \\ &= \frac{1}{m}(\tilde{p}_t)_\mu(\xi^\mu + f^\mu_{\mathrm{dp}})\mathrm{d}t + \frac{1}{m}(\tilde{p}_t)_\mu f^\mu_{\mathrm{noc}}\mathrm{d}t + \mathrm{d}_\lambda U, \quad (4) \end{aligned}$$

where $\tilde{E}_t$ is the energy considered as a random variable and E its realization. The part of the increase of energy caused by the heat reservoir is purely a thermal effect and thus comprehended as *trajectory heat* $\mathrm{d}\tilde{Q}_t$, the rest part is purely mechanical and should be comprehended as *trajectory work* $\mathrm{d}\tilde{W}_t$,

$$\mathrm{d}\tilde{Q}_t = \frac{(\tilde{p}_t)_\mu}{m}\left(\xi^\mu + f^\mu_{\mathrm{dp}}\right)\mathrm{d}t,$$
$$\mathrm{d}\tilde{W}_t = \frac{(\tilde{p}_t)_\mu}{m} f^\mu_{\mathrm{noc}}\mathrm{d}t + \mathrm{d}_\lambda U. \tag{5}$$

Thus eq. (4) becomes the first law of stochastic thermodynamics on the trajectory level, *i.e.*

$$\mathrm{d}\tilde{E}_t = \mathrm{d}\tilde{Q}_t + \mathrm{d}\tilde{W}_t.$$

The ambiguity in the decomposition of the external force also leads to an ambiguity in the trajectory work. However, in any case, the trajectory heat $\mathrm{d}\tilde{Q}_t = \mathrm{d}\tilde{E}_t - \mathrm{d}\tilde{W}_t$ is always unambiguously defined. Please be reminded that, unlike the usual heat and work in standard thermodynamics which are inexact differentials defined on the space of macro states, the *trajectory heat and work* are only defined on a stochastic trajectory of the Brownian particle.

The Fokker-Planck equation (FPE) associated to eqs. (1)-(2) can also be established on TM. The PDF for the Brownian particle under the measure $\mathrm{d}^{2d}X$ is denoted as $\Phi_t(x,p) := \Pr[\tilde{x}_t = x, \tilde{p}_t = p]$, and is clearly coordinate-independent. By use of the diffusion operator method [31], one can get

$$\partial_t \Phi_t = \frac{\partial}{\partial p^\mu}\left[\frac{1}{2} D^{\mu\nu}\frac{\partial}{\partial p^\nu}\Phi_t + \frac{1}{m}K^\mu{}_\nu p^\nu \Phi_t - f^\mu_{\mathrm{ex}}\Phi_t\right] - \frac{1}{m}\mathcal{L}(\Phi_t), \tag{6}$$

where $D^{\mu\nu} := R^\mu{}_a R^\nu{}_a$ is the diffusion tensor and

$$\mathcal{L} = p^\mu \frac{\partial}{\partial x^\mu} - \Gamma^\mu{}_{\alpha\beta}p^\alpha p^\beta \frac{\partial}{\partial p^\mu}$$

is the Liouville vector field on TM [25, 26].

Let the non-conservative force be temporarily turned off and the external control parameter be fixed. Then, after sufficiently long period of time, the Brownian particle will reach a thermal equilibrium with the heat reservoir, yielding the equilibrium PDF

$$\Phi_t(x,p) = \frac{1}{Z}\exp\left[-\frac{1}{T}\left(\frac{g_{\mu\nu}p^\mu p^\nu}{2m} + U(x,\lambda)\right)\right]. \tag{7}$$

Putting this PDF into the FPE (6), one gets the Einstein relation

$$D^{\mu\nu} = 2TK^{\mu\nu}, \tag{8}$$

which implies that the damping tensor $K^{\mu\nu}$ are not independent of the stochastic amplitudes $R^\mu{}_a$ and that $K^{\mu\nu}$ is invertible as a matrix. As long as only the FPE is concerned, there is an additional freedom in the sign choice of $R^\mu{}_a$, because $D^{\mu\nu}$ appear as a quadratic form in $R^\mu{}_a$.

To facilitate the discussion about FT, we introduce the time-reversal transform (TRT) for the process ranging from t_I to t_F:

$$I : \begin{cases} x^\mu(t) \mapsto x^\mu(t_F + t_I - t) \\ p^\mu(t) \mapsto -p^\mu(t_F + t_I - t) \end{cases},$$

which is often briefly described as $I : (x,p) \mapsto (x,-p)$ for short. Notice that the infinitesimal time increment $\mathrm{d}t$ is not affected by such transformation and remains to be positive.

It is obvious that TRT preserves the metric, *i.e.* $I^*\hat{g} = \hat{g}$. The damping force $f^\mu_{\mathrm{dp}}|_X = -K^\mu{}_\nu p^\nu/m$ reverses sign under TRT, thus the damping tensor must be invariant under TRT,

$$K^{\mu\nu}|_{I(X)} = K^{\mu\nu}|_X, \quad X = (x,p), \quad I(X) = (x,-p).$$

Eq. (8) implies that the diffusion tensor is also invariant under TRT. There is some freedom in choosing the stochastic amplitudes, and hence also in determining their behaviors under TRT. Here we assume the simplest transformation rule,

$$R^\mu{}_a|_{I(X)} = R^\mu{}_a|_X.$$

An immediate consequence is that the additional stochastic force should reverse sign under TRT. The coefficients of $\mathrm{d}t$ in eq. (2) can be classified into even and odd parts under TRT, *i.e.*

$$F^\mu = f^\mu_{\mathrm{ex}} - \frac{1}{m}\Gamma^\mu{}_{\alpha\beta}p^\alpha p^\beta \text{ with } F^\mu|_{I(X)} = F^\mu|_X,$$

and

$$\bar{F}^\mu = \frac{1}{2}R^\mu{}_a \frac{\partial}{\partial p^\nu}R^\nu{}_a - \frac{1}{m}K^\mu{}_\nu p^\nu \text{ with } \bar{F}^\mu|_{I(X)} = -\bar{F}^\mu|_X.$$

Then eq. (2) can be recast in a simpler form,

$$\mathrm{d}\tilde{p}^\mu_t = R^\mu{}_a \circ_S \mathrm{d}\tilde{w}^a_t + F^\mu \mathrm{d}t + \bar{F}^\mu \mathrm{d}t. \tag{9}$$

The last equation is the starting point for introducing the discretized version of the ULE in Appendix B 2.

B. Overdamped case

The stochastic mechanics characterized by LE is a branch of physics with multiple time scales, of which the smallest one is the time scale $\mathrm{d}t$ which allows for a sufficient number of collisions between the Brownian particle and the heat reservoir particles which cause little changes in the state of the Brownian particle [32]. If the temporal resolution Δt greatly exceeds $\mathrm{d}t$ but is still smaller than the relaxation time, the ULE (1)-(2) emerges. If Δt is further increased so that it is much larger than the relaxation time, the overdamped limit emerges.

The relaxation process of the Brownian particle can be viewed either as the process by which the damping force attains a state of mechanical equilibrium with other forces (mechanical relaxation process), or as the process by which the Brownian particle achieves local thermal equilibrium with the heat reservoir (thermodynamic relaxation process). The characteristic timescales associated with both processes are m/κ, where κ is the eigenvalue of $K^{\mu\nu}$. Consequently, the OLE should arise when Δt greatly exceeds m/κ.

The two kinds of relaxation process correspond to two approaches for taking the overdamped limit. From the perspective of mechanical relaxation process, the OLE can be described as the condition for mechanical equilibrium

$$0 = f^\mu_{\rm dp} + f^\mu_{\rm ex} + \xi^\mu. \tag{10}$$

This result is achieved in flat space under the condition that the damping tensor is position- and momentum-independent [33, 34], and is often viewed as a stochastic differential equation in configuration space. However, if the stochastic amplitudes are momentum-dependent, so is the additional stochastic force. Thus the mechanical equilibrium condition cannot be understood as a stochastic differential equation in configuration space. To avoid the above difficulty, let us consider the simpler situation in which the stochastic amplitudes are momentum-independent. Using the thermodynamic relaxation approach [35], it will be shown that, even in this simpler situation, a nontrivial additional stochastic force term still arises in the corresponding OLE.

The overdamped condition implies that the momentum space PDF already reaches the equilibrium form, while the configuration space PDF does not, so that the full PDF $\Phi_t(x,p)$ can be factorized,

$$\Phi_t(x,p) = \rho_t(x) P^s(x,p),$$

where

$$P^s(x,p) := \frac{1}{(2\pi m T)^{-d/2}} \exp\left[-\frac{g_{\mu\nu}(x)p^\mu p^\nu}{2mT}\right] \tag{11}$$

is the (Maxwell) equilibrium PDF in momentum space, and $\rho_t(x) := \int g^{1/2} {\rm d}^d p\, \Phi_t(x,p) = \Pr[\tilde{x}_t = x]$ is the PDF in configuration space. By adding the first order corrections from the near equilibrium states, the overdamped FPE is found to be (see Appendix. A),

$$\partial_t \rho_t = \nabla_\mu \left[\frac{1}{2}\hat{D}^{\mu\nu}\nabla_\nu \rho_t - \hat{K}^\mu{}_\nu f^\nu_{\rm ex}\rho_t\right], \tag{12}$$

where $\hat{K}^{\mu\nu} = (K^{-1})^{\mu\nu}$ and $\hat{D}^{\mu\nu} = 4T^2 (D^{-1})^{\mu\nu}$. Two important properties of eq. (12) are worth of notice: 1) The Einstein relation still holds for the rescaled damping and diffusion tensors

$$\hat{D}^{\mu\nu} = 2T\hat{K}^{\mu\nu};$$

and 2) The Boltzmann distribution

$$\rho_t(x) = \frac{1}{Z_x} {\rm e}^{-U(x,\lambda)/T}$$

is a solution of eq. (12) provided the non-conservative force is turned off and the external parameter is fixed, wherein Z_x represents the configuration space normalization factor, which should not be confused with the normalization factor Z appeared in eq. (7).

Using the diffusion operator method, it can be checked that the LE associated with eq. (12) takes the same form as eq. (10), but with a momentum-independent stochastic force term

$$\xi^\mu = R^\mu{}_a \circ_S {\rm d}\tilde{w}^a_t/{\rm d}t + \frac{1}{2} R^\mu{}_a \nabla_\nu (\hat{K}^\nu{}_\alpha R^\alpha{}_a). \tag{13}$$

Denoting

$$\hat{R}^\mu{}_a := \hat{K}^\mu{}_\nu R^\nu{}_a,\ \hat{F}^\mu = \hat{K}^\mu{}_\nu f^\nu_{\rm ex} + \frac{1}{2}\hat{R}^\mu{}_a \nabla_\nu \hat{R}^\nu{}_a, \tag{14}$$

the OLE can be written in a simpler form

$${\rm d}\tilde{x}^\mu_t = \hat{R}^\mu{}_a \circ_S {\rm d}\tilde{w}^a_t + \hat{F}^\mu {\rm d}t. \tag{15}$$

This equation is similar in form to the LE presented in [21]. However, unlike eq. (14), the additional stochastic force presented in [21] contains only an ordinary coordinate derivative rather than covariant derivative. Consequently, the claimed general covariance of the LE of [20] is questionable: it actually holds only for flat manifolds and only with respects to linear coordinate transformations.

Since the inertial effect can be ignored in the overdamped case, the energy of the Brownian particle contains only the potential energy, *i.e.* $E(x,\lambda) = U(x,\lambda)$. Using eq. (10) and the chain rule, we have

$$\begin{aligned} {\rm d}\tilde{E}_t &= \frac{\partial U}{\partial x^\mu}{\rm d}\tilde{x}^\mu_t + {\rm d}_\lambda U = -(f_{\rm con})_\mu {\rm d}\tilde{x}^\mu_t + {\rm d}_\lambda U \\ &= [\xi_\mu + (f_{\rm dp})_\mu]{\rm d}\tilde{x}^\mu_t + (f_{\rm noc})_\mu {\rm d}\tilde{x}^\mu_t + {\rm d}_\lambda U. \end{aligned}$$

Similar to the underdamped case, the energy absorbed from the heat reservoir is understood as the trajectory heat, and the rest part of the energy increase as trajectory work. Thus we have, thanks to eq. (10),

$$\begin{aligned} {\rm d}\tilde{Q}_t &= [\xi_\mu + (f_{\rm dp})_\mu]{\rm d}\tilde{x}^\mu_t = -(f_{\rm ex})_\mu {\rm d}\tilde{x}^\mu_t, \\ {\rm d}\tilde{W}_t &= (f_{\rm noc})_\mu {\rm d}\tilde{x}^\mu_t + {\rm d}_\lambda U. \end{aligned} \tag{16}$$

III. FLUCTUATION THEOREM

Now we come to the stage for describing FT on Riemannian manifold based on the description of stochastic trajectories. Since the trajectory probability in continuous time is hard to deal with, we adopt the following strategy: first we take a discrete equidistant set of time nodes $t_I = t_0 < t_1 < ... < t_n = t_F$ to rewrite the LE and factorize the corresponding trajectory probabilities, and then take the continuum limit at the end of the calculation.

The stochastic process in discrete time can be viewed as a sequence of random variables, *i.e.* $\tilde{X}_{[t]} = (\tilde{X}_0, \tilde{X}_1, ..., \tilde{X}_n)$ with $\tilde{X}_i = \tilde{X}_{t_i}$ for the underdamped case and $\tilde{x}_{[t]} = (\tilde{x}_0, \tilde{x}_1, ..., \tilde{x}_n)$, with $\tilde{x}_i = \tilde{x}_{t_i}$ for the overdamped case.

Before proceeding, it is necessary to clarify the concepts of *ensemble* and *trajectory* entropy productions. At the time t, the ensemble entropy of the Brownian particle reads

$$S_t = -\int \mathrm{d}^{2d}X\, \Phi_t \ln \Phi_t,$$

and the ensemble entropy production in the process is $\Delta S = S_{t_F} - S_{t_I}$. Notice that (throughout this paper, an overline denotes ensemble average, while $\langle\ \ \rangle$ denotes trajectory average)

$$-\overline{\ln \rho_t} := -\int g^{1/2}\mathrm{d}^d x\ \rho_t \ln \rho_t$$

is *not* the entropy of the overdamped Brownian particle, however, the ensemble entropy production of the Brownian particle can be represented as the difference of $-\overline{\ln \rho_t}$, because the ensemble entropy for the overdamped Brownian particle can be evaluated to be

$$S_t = -\overline{\ln \rho_t} + \frac{d}{2} + \frac{d}{2}\ln(2\pi m T),$$

where the last two terms arise from the momentum space integration of the term involving the distribution $P^s(x,p)$ given in eq. (11). Subtracting the initial value from the final value leaves only the difference of $-\overline{\ln \rho_t}$,

$$\Delta S = \overline{\ln \rho_{t_I}} - \overline{\ln \rho_{t_F}}.$$

In contrast, the trajectory entropy production is defined simply to be the difference between the logarithms of the PDF at the initial and final times, *i.e.*

$$\Delta S_{X_{[t]}} = \ln \Phi_{t_I}(X_0) - \ln \Phi_{t_F}(X_n)$$

for the underdamped and

$$\Delta S_{x_{[t]}} = \ln \rho_{t_I}(X_0) - \ln \rho_{t_F}(X_n)$$

for the overdamped cases.

Now let us consider the underdamped case. It is important to distinguish the terms *process* and *trajectory*: the latter is a realization of the former. The forward process $\tilde{X}_{[t]}$ refers to a stochastic process governed by the ULE (1)-(2), wherein the external control parameter $\lambda_{[t]} = (\lambda_0, \lambda_1, ..., \lambda_n)$ varies over time. Correspondingly, the reversed process $\tilde{X}^-_{[t]}$ also refers to a stochastic process governed by the same LE, but its initial state should be identified with the time-reversal of the final state of the forward process, *i.e.* $\tilde{X}^-_0 = I(\tilde{X}_n)$, and the corresponding external control parameter should satisfy $\lambda^-_i = \lambda_{n-i}$. The reversed trajectory $X_{[t]} = (X_0, X_1, ..., X_n)$ is defined such that $X^-_i := I(X_{n-i})$.

We will prove that the total entropy production, *i.e.* the sum of the trajectory entropy production with the change of the entropy of the heat reservoir, should be

$$\Sigma_{X_{[t]}} = \ln \frac{\Pr[\tilde{X}_{[t]} = X_{[t]}]}{\Pr[\tilde{X}^-_{[t]} = X^-_{[t]}]}. \tag{17}$$

Since the Brownian motion is a Markov process, the trajectory probability can be decomposed into product of transition probabilities,

$$\begin{aligned}&\Pr[\tilde{X}_{[t]} = X_{[t]}]\\ &\quad= \Big(\prod_{i=0}^{n-1} \Pr[\tilde{X}_{i+1} = X_{i+1}|\tilde{X}_i = X_i]\Big)\Pr[\tilde{X}_0 = X_0].\end{aligned}$$

A similar decomposition can be made for $\Pr[\tilde{X}^-_{[t]} = X^-_{[t]}]$. Therefore, we have

$$\begin{aligned}\Sigma_{X_{[t]}} &= \sum_{i=0}^{n-1} \ln \frac{\Pr[\tilde{X}_{i+1} = X_{i+1}|\tilde{X}_i = X_i]}{\Pr[\tilde{X}^-_{i+1} = X^-_{i+1}|\tilde{X}^-_i = X^-_i]} + \ln \frac{\Pr[\tilde{X}_0 = X_0]}{\Pr[\tilde{X}^-_0 = X^-_0]}\\ &= \sum_{i=0}^{n-1} \ln \frac{\Pr[\tilde{X}_{i+1} = X_{i+1}|\tilde{X}_i = X_i]}{\Pr[\tilde{X}^-_{n-i} = I(X_i)|\tilde{X}^-_{n-i-1} = I(X_{i+1})]} + \ln \frac{\Pr[\tilde{X}_0 = X_0]}{\Pr[\tilde{X}_n = X_n]},\end{aligned} \tag{18}$$

where, in the last step, the definitions of the reversed process and reversed trajectory have been used and re-arrange of terms in the summation has been adopted. The last term in eq. (18) is simply the trajectory entropy production, because

$$\Phi_{t_I}(X_0) = \Pr[\tilde{X}_0 = X_0], \quad \Phi_{t_F}(X_n) = \Pr[\tilde{X}_n = X_n].$$

On the other hand, the continuum limit of the first term reads (see Appendix.B 2)

$$\begin{aligned}
&\lim_{n\to+\infty} \sum_{i=0}^{n-1} \ln \frac{\Pr[\tilde{X}_{i+1} = X_{i+1}|\tilde{X}_i = X_i]}{\Pr[\tilde{X}^-_{n-i} = I(X_i)|\tilde{X}^-_{n-i-1} = I(X_{i+1})]} \\
&= -\frac{1}{T}\int_{t_I}^{t_F} \mathrm{d}t \frac{p^\mu}{m} \nabla_\mu (\mathcal{T} + U) + \frac{1}{T}\int_{t_I}^{t_F} \mathrm{d}t \frac{p_\mu}{m} f^\mu_{\rm noc}, \quad (19)
\end{aligned}$$

where $\mathcal{T} := p^\mu p_\mu/2m$ is the kinematic energy. According to eq. (5), the trajectory work should be

$$\Delta W_{X_{[t]}} = \int_{t_I}^{t_F} \mathrm{d}t \left(\frac{p_\mu}{m} f^\mu_{\rm noc} + \frac{\partial U}{\partial \lambda}\frac{\mathrm{d}\lambda}{\mathrm{d}t}\right),$$

and the change of energy is

$$\begin{aligned}
\Delta E_{X_{[t]}} &= \Delta \mathcal{T}_{X_{[t]}} + \Delta U_{X_{[t]}} \\
&= \int_{t_I}^{t_F} \mathrm{d}t \frac{p^\mu}{m}\nabla_\mu \mathcal{T} + \int_{t_I}^{t_F} \mathrm{d}t \left(\frac{p^\mu}{m}\nabla_\mu U + \frac{\partial U}{\partial \lambda}\frac{\mathrm{d}\lambda}{\mathrm{d}t}\right).
\end{aligned}$$

Since the trajectory heat is $\Delta Q_{X_{[t]}} = \Delta E_{X_{[t]}} - \Delta W_{X_{[t]}}$, eq. (18) can also be rewritten as

$$\Sigma_{X_{[t]}} = \Delta S_{X_{[t]}} - \frac{1}{T}\Delta Q_{X_{[t]}} = \Delta S_{X_{[t]}} + \Delta S_{\rm Res}, \quad (20)$$

where, since the heat reservoir maintains in equilibrium, the Clausius equality holds, the change of the entropy of the heat reservoir reads

$$\Delta S_{\rm Res} = \Delta Q_{\rm Res}/T = -\Delta Q_{X_{[t]}}/T.$$

Thus $\Sigma_{X_{[t]}}$ is indeed the total entropy production. Inserting eq. (20) back into eq. (17), we get the desired FT

$$\frac{\Pr[\tilde{X}_{[t]} = X_{[t]}]}{\Pr[\tilde{X}^-_{[t]} = X^-_{[t]}]} = \mathrm{e}^{\Sigma_{X_{[t]}}} = \mathrm{e}^{\Delta S_{X_{[t]}} - \Delta Q_{X_{[t]}}/T}, \quad (21)$$

which tells that the process with positive total entropy production is probabilistically more preferred.

Taking the trajectory average of eq. (21), we get, by use of the Jensen inequality, the following result,

$$\begin{aligned}
\mathrm{e}^{-\left\langle \Sigma_{\tilde{X}_{[t]}}\right\rangle} &\leq \left\langle \mathrm{e}^{-\Sigma_{\tilde{X}_{[t]}}}\right\rangle = \int \mathcal{D}[X_{[t]}] \Pr[\tilde{X}^-_{[t]} = X^-_{[t]}] \\
&= \int \mathcal{D}[X^-_{[t]}] \Pr[\tilde{X}^-_{[t]} = X^-_{[t]}] = 1, \quad (22)
\end{aligned}$$

where $\mathcal{D}[X_{[t]}] = \mathrm{d}X_0 \wedge \mathrm{d}X_1 \wedge ... \wedge \mathrm{d}X_n$ is the measure on the trajectory space. It is easy to prove that the map $X_{[t]} \mapsto X^-_{[t]}$ preserves the measure, *i.e.* $\mathcal{D}[X_{[t]}] = \mathcal{D}[X^-_{[t]}]$. Eq.(22) is the so-called integral FT, which tells that the entropy production is non-negative in any macroscopic process, *i.e.* $\left\langle \Sigma_{\tilde{X}_{[t]}}\right\rangle \geq 0$.

The FT in the overdamped case can be constructed following a similar fashion, however the processes must be described solely in configuration space. The definitions of the reversed process $\tilde{x}^-_{[t]}$ and the reversed trajectory $x^-_{[t]}$ are similar to the underdamped case, with the replacement $X_i \to x_i$. Therefore,

$$\begin{aligned}
&\Sigma_{x_{[t]}} := \ln \frac{\Pr[\tilde{x}_{[t]} = x_{[t]}]}{\Pr[\tilde{x}^-_{[t]} = x^-_{[t]}]} \\
&= \sum_{i=0}^{n-1} \ln \frac{\Pr[\tilde{x}_{i+1} = x_{i+1}|\tilde{x}_i = x_i]}{\Pr[\tilde{x}^-_{n-i} = x_i|\tilde{x}^-_{n-i-1} = x_{i+1}]} + \ln \frac{\Pr[\tilde{x}_0 = x_0]}{\Pr[\tilde{x}_n = x_n]}.
\end{aligned}$$

In the continuum limit, we have (see Appendix.B 1)

$$\begin{aligned}
&\lim_{n\to+\infty} \sum_{i=0}^{n-1} \ln \frac{\Pr[\tilde{x}_{i+1} = x_{i+1}|\tilde{x}_i = x_i]}{\Pr[\tilde{x}^-_{n-i} = x_i|\tilde{x}^-_{n-i-1} = x_{i+1}]} \\
&= \frac{1}{T}\int_{t_I}^{t_F} v_\mu f^\mu_{\rm ex} \mathrm{d}t, \quad (23)
\end{aligned}$$

where v^μ is the velocity of the Brownian particle. According to eq. (16), the trajectory heat is

$$\Delta Q_{x_{[t]}} = -\int_{t_I}^{t_F} v_\mu f^\mu_{\rm ex} \mathrm{d}t,$$

and the trajectory entropy production is

$$\Delta S_{x_{[t]}} = \ln \rho_{t_I}(x_0) - \ln \rho_{t_F}(x_n) = \ln \frac{\Pr[\tilde{x}_0 = x_0]}{\Pr[\tilde{x}_n = x_n]}.$$

Finally, we arrive at the desired FT

$$\frac{\Pr[\tilde{x}_{[t]} = x_{[t]}]}{\Pr[\tilde{x}^-_{[t]} = x^-_{[t]}]} = \mathrm{e}^{\Sigma_{x_{[t]}}} = \mathrm{e}^{\Delta S_{x_{[t]}} - \Delta Q_{x_{[t]}}/T}.$$

The integral FT in the overdamped case can be obtained in complete analogy to the underdamped case, therefore, there is no need to repeat the construction.

IV. CONCLUDING REMARKS

The covariant LE and FPE on a Riemannian manifold are constructed in both underdamped and overdamped cases. The concepts of trajectory heat and trajectory work are clarified, and the first law on the trajectory level is established. The Stratonovich coupling plays an important role in establishing the appropriate form of the first law. In either cases, the corresponding FTs are proved in both differential and integral forms. These results allow for a complete extension of the existing stochastic thermodynamics to arbitrary Riemannian manifolds, which in turn may be helpful in understanding the origin of irreversibility in certain biological scenarios.

During the construction, we also clarified the link between the PDF with the ensemble and trajectory entropy productions, which may also shed some light in the parallel constructions in flat spaces. Moreover, the different forms of the additional stochastic forces in the underdamped and overdamped cases are also worth of notice. Last, we hope the results presented here could be inspiring for an ultimate resolution for the fully general relativistic construction for the FTs.

ACKNOWLEDGEMENT

This work is supported by the National Natural Science Foundation of China under the grant No. 12275138.

[1] D. J. Evans and D. J. Searles, Equilibrium microstates which generate second law violating steady states. *Phys. Rev. E*, 50(2):1645, 1994.
[2] G. Gallavotti and E. G. D. Cohen, Dynamical ensembles in nonequilibrium statistical mechanics. *Phys. Rev. Lett.*, 74(14):2694, 1995. [arXiv:chao-dyn/9410007].
[3] G. E. Crooks, Entropy production fluctuation theorem and the nonequilibrium work relation for free energy differences. *Phys. Rev. E*, 60(3):2721, 1999. [arXiv:cond-mat/9901352].
[4] K. Sekimoto, Langevin equation and thermodynamics. *Prog Theor Phys Supp*, 130:17–27, 1998.
[5] U. Seifert, Entropy production along a stochastic trajectory and an integral fluctuation theorem. *Phys. Rev. Lett.*, 95(4):040602, 2005. [arXiv:cond-mat/0503686].
[6] A Imparato and L Peliti, Fluctuation relations for a driven Brownian particle. *Phys. Rev. E*, 74(2):026106, 2006. [arXiv:cond-mat/0603506].
[7] V. Y. Chernyak, M. Chertkov, and C. Jarzynski, Path-integral analysis of fluctuation theorems for general Langevin processes. *J. Stat. Mech. Theory Exp.*, 2006(08):P08001, 2006. [arXiv:cond-mat/0605471].
[8] T. Ohkuma and T. Ohta, Fluctuation theorems for non-linear generalized Langevin systems. *J. Stat. Mech. Theory Exp.*, 2007(10):P10010, 2007.
[9] R. Randel, H. C. Loebl, and C. C. Matthai, Molecular dynamics simulations of polymer translocations. *Macromol Theory Simul*, 13(5):387–391, 2004.
[10] M. Muthukumar and C.Y. Kong, Simulation of polymer translocation through protein channels. *Proc. Natl. Acad. Sci. U.S.A.*, 103(14):5273–5278, 2006.
[11] U. Seifert, Stochastic thermodynamics, fluctuation theorems and molecular machines. *Rep. Prog. Phys.*, 75(12):126001, 2012. [arXiv:1205.4176].
[12] H. Qian, S. Kjelstrup, A. B. Kolomeisky, and D. Bedeaux, Entropy production in mesoscopic stochastic thermodynamics: nonequilibrium kinetic cycles driven by chemical potentials, temperatures, and mechanical forces. *J. Phys. Condens. Matter*, 28(15):153004, 2016. [arXiv:1601.04018].
[13] I. Golokolenov, A. Ranadive, L. Planat, M. Esposito, N. Roch, X. Zhou, A. Fefferman, and E. Collin, Thermodynamics of a single mesoscopic phononic mode. *Phys. rev. res.*, 5(1):013046, 2023. [arXiv:2208.08395].
[14] K. D. Elworthy, *Stochastic Differential Equations on Manifolds*, pages 165–178. Springer New York, New York, NY, 1998. ISBN: 9781461222248.
[15] E. P. Hsu, *Stochastic analysis on manifolds*, American Mathematical Soc., Number 38, 2002. ISBN: 9780821808023.
[16] P. A. Meyer, A differential geometric formalism for the Ito calculus. In *Stochastic Integrals*, pages 256–270, Berlin, Heidelberg, 1981, Ed. by David Williams.
[17] M. Émery, *Stochastic calculus in manifolds*, Springer Science & Business Media, 2012. ISBN: 9783642750519.
[18] J. Armstrong and D. Brigo, Coordinate-free stochastic differential equations as jets. [arXiv:1602.03931].
[19] J. Armstrong and D. Brigo, Intrinsic stochastic differential equations as jets. *Proceedings of the Royal SocietyA: Mathematical, Physical and Engi*
[20] R. Graham, "Covariant formulation of non-equilibrium statistical thermodynamics," *Z Physik B*, 26: 397-405, 1977.
[21] R. Graham, "Covariant stochastic calculus in the sense of Itô," *Phys. lett. A*, 109 (5): 209, 1985.
[22] M. Ding, Z. Tu, and X. Xing, Covariant formulation of nonlinear Langevin theory with multiplicative gaussian white noises. *Phys. Rev. Res.*, 2(3):033381, 2020. [arXiv:2007.16131].
[23] M. Ding and X. Xing, Covariant nonequilibrium thermodynamics from Ito-Langevin dynamics. *Phys. Rev. Res.*, 4(3):033247, 2022. [arXiv:2105.14534].
[24] Y.F. Cai, T. Wang, and L. Zhao, Relativistic stochastic dynamics I: Langevin equation from observer's perspective, 2023. *J. Stat. Phys.* (2023) 190:193. [arXiv:2306.01982].
[25] O. Sarbach and T. Zannias, "Relativistic kinetic theory: an introduction," in *AIP Conference Proceedings*, volume 1548, pages 134–155, American [arXiv:1303.2899].
[26] O. Sarbach and T. Zannias, "The geometry of the tangent bundle and the relativistic kinetic theory of gases," *Class. Quant. Grav.*, 31(8):085013, 2014. [arXiv:1309.2036].
[27] R. O. Acuna-Cardenas, C. Gabarrete, O. Sarbach, "An introduction to the relativistic kinetic theory on curved spacetimes," *Gen. Rel. Grav.* (2022) 54:23. [arXiv:2106.09235].
[28] S. Sasaki, On the differential geometry of tangent bundles of Riemannian manifolds.

Tohoku Math. J., 10(3):338–354, 1958.
[29] Y. L. Klimontovich, Nonlinear Brownian motion. *Phys-Usp*, 37(8):737, 1994.
[30] Y.F. Cai, T. Wang, and L. Zhao, Relativistic stochastic mechanics II: Reduced Fokker-Planck equation in curved spacetime. *J Stat Phys*, 190:181, 2023. [arXiv:2307.07805].
[31] D. Bakry, I. Gentil, and M. Ledoux, *Analysis and geometry of Markov diffusion operators*, volume 103. Springer, 2013. ISBN: 9783319002279.
[32] K. Sekimoto. *Stochastic energetics*, volume 799. Springer, 2010. ISBN: 9783642054112.
[33] A. S. Bodrova, A. V. Chechkin, A. G. Cherstvy, H. Safdari, I. M. Sokolov, and R. Metzler, Underdamped scaled brownian motion:(non-) existence of the overdamped limit in anomalous diffusion. *Sci. Rep.*, 6(1):30520, 2016. [arXiv:1609.07250].
[34] H. Huang, Quantitative estimate of the overdamped limit for the Vlasov-Fokker-Planck systems. *Partial Differen-tial Equations in Applied Mathematics*, 4:100186, 20 [arXiv:2110.00883].
[35] X. Durang, C. Kwon, and H. Park, Overdamped limit and inverse-friction expansion for Brownian motion in an inhomogeneous medium. *Phys. Rev. E.*, 91(6):062118, 2015. [arXiv:1309.5750].

Appendix A: Overdamped Fokker-Planck equation

Here we outline the procedure for taking the overdamped limit following the line of Ref. [35]. Using the diffusion operator

$$A = \frac{\delta^{ab}}{2} L_a L_b + L_0 \quad \text{with } L_0 = \frac{1}{m}\mathcal{L} - \frac{1}{m} K^{\mu}{}_{\nu} p^{\nu} \frac{\partial}{\partial p^{\mu}}, \quad L_a = R^{\mu}{}_a \frac{\partial}{\partial p^{\mu}},$$

the underdamped FPE can be written as

$$\partial_t \Phi_t = A^{\dagger} \Phi_t, \tag{A1}$$

where $A^{\dagger}$ is the adjoint of A. The FPE can be rewritten as

$$\partial_t \bar{\Phi}_t = \bar{A}^{\dagger} \bar{\Phi}_t,$$

where $\bar{\Phi}_t := [P^s]^{-1/2} \Phi_t$ and $\bar{A}^{\dagger} := [P^s]^{-1/2} A^{\dagger} [P^s]^{1/2}$, and P^s is given in eq. (11). $\bar{A}^{\dagger}$ can be decomposed in terms of the creation and annihilation operators

$$a_{\mu} = \zeta \frac{\partial}{\partial p^{\mu}} + \frac{1}{2\zeta} p_{\mu}, \qquad a^{\dagger}_{\mu} = -\zeta \frac{\partial}{\partial p^{\mu}} + \frac{1}{2\zeta} p_{\mu}$$

which obey the commutation relation $[a_{\mu}, a^{\dagger}_{\nu}] = g_{\mu\nu}$, wherein $\zeta = \sqrt{mT}$.

Since $K^{\mu\nu}$ is a symmetric tensor, its eigenvectors $e^{\mu}{}_{\hat{\nu}}$ constitute an orthonormal basis. The components of a tensor under the orthonormal basis are denoted by adding a hat on its index, *e.g.* $W_{\hat{\mu}} = e^{\nu}{}_{\hat{\mu}} W_{\nu}$, $V_{\hat{\mu}} = e_{\nu\hat{\mu}} V^{\nu}$, where $e_{\nu\hat{\mu}}$ is the dual basis. For convenience, only lower indices are used under the orthonormal basis. The commutator between the creation and the annihilation operators can be rewritten as $[a_{\hat{\mu}}, a^{\dagger}_{\hat{\nu}}] = \delta_{\hat{\mu}\hat{\nu}}$. Moreover, p_{μ} and $\frac{\partial}{\partial p^{\mu}}$ can be decomposed as

$$p_{\mu} = \zeta e_{\mu\hat{\nu}} (a_{\hat{\nu}} + a^{\dagger}_{\hat{\nu}}), \qquad \frac{\partial}{\partial p^{\mu}} = \frac{1}{2\zeta} e_{\mu\hat{\nu}} (a_{\hat{\nu}} - a^{\dagger}_{\hat{\nu}}).$$

Using the above notations, one has

$$\bar{A}^{\dagger} = -\frac{1}{m} \kappa^{\hat{\mu}} N_{\hat{\mu}} + \frac{1}{\zeta} f^{\hat{\mu}}_{\text{ex}} a^{\dagger}_{\hat{\mu}} - \frac{1}{m} \mathcal{L},$$

where $\kappa^{\hat{\mu}}$ represent the eigenvalues of $K^{\mu\nu}$, and $N_{\hat{\mu}} := a^{\dagger}_{\hat{\mu}} a_{\hat{\mu}}$. Let $\psi_0 := [P^s]^{1/2}$, $\psi_{\hat{\mu}} := a^{\dagger}_{\hat{\mu}} \psi_0$. Clearly, we have $a_{\hat{\mu}} \psi_0 = 0$, *i.e.* the ground state ψ_0 of the Fock space generated by $a^{\dagger}_{\hat{\mu}}$ corresponds to the equilibrium distribution P^s, and the excited states to non-equilibrium modifications.

In principle, $\bar{\Phi}_t$ can be expanded as a linear superposition of the eigenstates of $N_{\hat{\mu}}$. However, in the overdamped limit, the time resolution Δt is considered to be much larger than the relaxation time, so, in the first-order approximation near equilibrium state, we have

$$\bar{\Phi}_t(x, p) \approx C^0(x, t) \psi_0(x, p) + C^{\hat{\mu}}(x, t) \psi_{\hat{\mu}}(x, p). \tag{A2}$$

Some important commutation relations are listed below:

$$\begin{aligned}
&[p^{\alpha}, a^{\dagger}_{\hat{\mu}}] = \zeta e^{\alpha}{}_{\hat{\mu}}, \quad [\frac{\partial}{\partial p^{\alpha}}, a^{\dagger}_{\hat{\mu}}] = \frac{1}{2\zeta} e_{\alpha\hat{\mu}}, \\
&[\frac{\partial}{\partial x^{\alpha}}, a^{\dagger}_{\hat{\mu}}] = \frac{1}{2\zeta} \partial_{\alpha} g_{\nu\beta} p^{\beta} e^{\nu}{}_{\hat{\mu}} + a^{\dagger}_{\nu} \partial_{\alpha} e^{\nu}{}_{\hat{\mu}}, \\
&[\mathcal{L}, a^{\dagger}_{\hat{\mu}}] = \zeta e^{\sigma}{}_{\hat{\mu}} \left[\frac{\partial}{\partial x^{\sigma}} - \Gamma^{\nu}{}_{\alpha\sigma} p^{\alpha} \frac{\partial}{\partial p^{\nu}} \right] + \zeta e_{\nu\hat{\alpha}} \nabla_{\hat{\beta}} e^{\nu}{}_{\hat{\mu}} (a_{\hat{\beta}} + a^{\dagger}_{\hat{\beta}}) a^{\dagger}_{\hat{\alpha}}.
\end{aligned} \tag{A3}$$

Using eq. (A3) one gets

$$\mathcal{L}(\bar{\Phi}_t) = \mathcal{L}(C^0)\psi_0 + \mathcal{L}(C^{\hat{\mu}})\psi_{\hat{\mu}} + C^{\hat{\mu}} \mathcal{L}(\psi_{\hat{\mu}}) = \mathcal{L}(C^0)\psi_0 + \mathcal{L}(C^{\hat{\mu}})\psi_{\hat{\mu}} + C^{\hat{\mu}} [\mathcal{L}, a^{\dagger}_{\hat{\mu}}] \psi_0$$

$$= \mathcal{L}(C^0)\psi_0 + \mathcal{L}(C^{\hat{\mu}})\psi_{\hat{\mu}} + \zeta\nabla_\nu e^\nu{}_{\hat{\mu}} C^{\hat{\mu}}\psi_0 + \zeta e_{\nu\hat{\alpha}}(\nabla_{\hat{\beta}} e^\nu{}_{\hat{\mu}})C^{\hat{\mu}}(a^\dagger_{\hat{\beta}} a^\dagger_{\hat{\alpha}}\psi_0),$$

where the property $\left(\frac{\partial}{\partial x^\sigma} - \Gamma^\nu{}_{\alpha\sigma}p^\alpha\frac{\partial}{\partial p^\nu}\right)\psi^0 = 0$ has been used. Since $\mathcal{L}(C^0)$ and $\mathcal{L}(C^{\hat{\mu}})$ still contain momentum, these expressions can be further expanded,

$$\mathcal{L}(C^0)\psi_0 = \partial^{\hat{\mu}} C^0 p_{\hat{\mu}}\psi_0 = \zeta\nabla^{\hat{\mu}} C^0\psi_{\hat{\mu}},$$
$$\mathcal{L}(C^{\hat{\nu}})\psi_{\hat{\nu}} = \zeta\nabla_{\hat{\mu}} C^{\hat{\mu}}\psi_0 + \zeta\nabla_{\hat{\alpha}} C^{\hat{\beta}}(a^\dagger_{\hat{\alpha}} a^\dagger_{\hat{\beta}}\psi_0).$$

Defining an operator $D_{\hat{\mu}} := (\zeta/T)(f^{\hat{\mu}}_{\rm ex} - T\partial_{\hat{\mu}})$ and substituting eq. (A2) into eq. (A1), the evolution equations of C^0 and $C^{\hat{\mu}}$ follow,

$$\partial_t C^0 = -\frac{T}{\zeta}\left[\nabla_{\hat{\mu}} C^{\hat{\mu}} + (\nabla_\nu e^\nu{}_{\hat{\mu}})C^{\hat{\mu}}\right], \tag{A4}$$

$$\partial_t C^{\hat{\mu}} = -\frac{\kappa^{\hat{\mu}}}{m} C^{\hat{\mu}} + \frac{1}{m} D_{\hat{\mu}} C^0. \tag{A5}$$

The overdamped limit means that $\kappa^{\hat{\mu}}$ is large. In such a limit, C^0 and $C^{\hat{\mu}}$ are respectively of orders $O((\kappa^{\hat{\mu}})^0)$ and $O((\kappa^{\hat{\mu}})^{-1})$. Therefore, up to order $O((\kappa^{\hat{\mu}})^0)$, we can safely ignore the left hand side of eq. (A5), yielding

$$C^{\hat{\mu}} \approx \frac{1}{\kappa^{\hat{\mu}}} D_{\hat{\mu}} C^0.$$

Inserting this result into eq. (A4), we get

$$\begin{aligned}\partial_t C^0 &= -\frac{T}{\zeta}\left[\nabla_{\hat{\mu}}\left(\frac{1}{\kappa^{\hat{\mu}}} D_{\hat{\mu}} C^0\right) + (\nabla_\nu e^\nu{}_{\hat{\mu}})\frac{1}{\kappa^{\hat{\mu}}} D_{\hat{\mu}} C^0\right] = -\frac{T}{\zeta}\nabla_\nu\left(e^\nu{}_{\hat{\mu}}\frac{1}{\kappa^{\hat{\mu}}} e^\alpha{}_{\hat{\mu}} D_\alpha C^0\right)\\ &= -\nabla_\nu \hat{K}^{\nu\alpha}\left[g_{\alpha\mu} f^\mu_{\rm ex} C^0 - T\nabla_\alpha C^0\right]\\ &= \nabla_\mu\left[\frac{1}{2}\hat{D}^{\mu\nu}\nabla_\nu C^0 - \hat{K}^\mu{}_\nu f^\nu_{\rm ex} C^0\right],\end{aligned} \tag{A6}$$

where $\hat{K}^{\mu\nu} = e^\mu{}_{\hat{\alpha}} e^\nu{}_{\hat{\alpha}}(\kappa^{\hat{\alpha}})^{-1}$ is inverse of the damping tensor $K^{\mu\nu}$, and $\hat{D}^{\mu\nu} := 2T\hat{K}^{\mu\nu} = 4T^2(D^{-1})^{\mu\nu}$ is the rescaled diffusion tensor. Since

$$\rho_t(x) = \int g^{1/2}{\rm d}^d p\ \Phi_t(x,p) = \int g^{1/2}{\rm d}^d p\ \psi_0(x,p)\bar{\Phi}_t(x,p) = C^0(x,t),$$

the overdamped FPE (12) follows from eq. (A6).

Appendix B: Continuum limit

Let us first introduce some mathematical tricks. Let A be a full-rank square matrix and B be a matrix of the same size, the determinant of $A + B{\rm d}t$ can be expanded as

$$\det[A + B{\rm d}t + O({\rm d}t^2)] = \det[A]\det[I + A^{-1}B{\rm d}t + O({\rm d}t^2)] = \det[A] + \det[A]\,{\rm Tr}[A^{-1}B]{\rm d}t + O({\rm d}t^2). \tag{B1}$$

Let $f(t)$ be a continuous function on $[t_I, t_F]$. Then there is a continuum limit

$$\lim_{n\to+\infty}\prod_{i=0}^{n-1}[1 + f(t_i){\rm d}t + O({\rm d}t^2)] = \exp\left[\int_{t_I}^{t_F}{\rm d}t f(t)\right], \tag{B2}$$

where ${\rm d}t = (t_F - t_I)/n$ and $t_i \in [i{\rm d}t, (i+1){\rm d}t]$. Combining eqs. (B1)-(B2), we have

$$\lim_{n\to+\infty}\prod_{i=0}^{n-1}\frac{\det[A(t_i) + B(t_i){\rm d}t + O({\rm d}t^2)]}{\det[A(t_i) + C(t_i){\rm d}t + O({\rm d}t^2)]} = \exp\left[\int_{t_I}^{t_F}{\rm d}t\,{\rm Tr}[A^{-1}(B - C)]\right], \tag{B3}$$

where $A(t), B(t), C(t)$ are continuous matrix functions on $[t_I, t_F]$.

1. Overdamped case

The discrete time version of the OLE (15) reads

$$\tilde{x}^\mu_{i+1} - \tilde{x}^\mu_i = \hat{F}^\mu(\bar{\lambda}_i)|_{(\tilde{x}_{i+1}+\tilde{x}_i)/2}{\rm d}t + \hat{R}^\mu{}_a|_{(\tilde{x}_{i+1}+\tilde{x}_i)/2}{\rm d}\tilde{w}^a_i,$$

where $\bar{\lambda}_i := (\lambda_{i+1} + \lambda_i)/2$. The choice of discretization rule for $\hat{F}^\mu$ does not affect the continuum limit, but we still take the middle-point rule for consistency of notations. We introduce a function

$${\rm d}w^a(x, y, \lambda) := (\hat{R}^{-1})^a{}_\nu|_{\bar{x}}\left[x^\nu - y^\nu + \hat{F}^\nu(\lambda)|_{\bar{x}}{\rm d}t\right],$$

where $\bar{x} := (x+y)/2$. This function is connected with the Gaussian noise $\mathrm{d}\tilde{w}_i^a$ via the relation

$$\mathrm{d}\tilde{w}_i^a = \mathrm{d}w^a(\tilde{x}_{i+1}, \tilde{x}_i, \bar{\lambda}_i). \tag{B4}$$

The reversed process is governed by same LE with the external control parameter time-reversed, so that

$$\mathrm{d}\tilde{w}_i^a = \mathrm{d}w^a(\tilde{x}_{i+1}^-, \tilde{x}_i^-, \bar{\lambda}_i^-) = \mathrm{d}w^a(\tilde{x}_{i+1}^-, \tilde{x}_i^-, \bar{\lambda}_{n-i-1}).$$

The last equation can also be rewritten as

$$\mathrm{d}\tilde{w}_{n-i-1}^a = \mathrm{d}w^a(\tilde{x}_{n-i}^-, \tilde{x}_{n-i-1}^-, \bar{\lambda}_i). \tag{B5}$$

Using eqs. (B4) and (B5), the transition probabilities of the forward and reversed processes can be expressed as

$$\Pr[\tilde{x}_{i+1} = x_{i+1} | \tilde{x}_i = x_i] = g^{-1/2}(x_{i+1}) \det\left[\frac{\partial \mathrm{d}w^a}{\partial x^\mu}\right](x_{i+1}, x_i, \bar{\lambda}_i) \Pr[\mathrm{d}\tilde{w}_i^a = \mathrm{d}w^a(x_{i+1}, x_i, \bar{\lambda}_i)] \tag{B6}$$

$$\Pr[\tilde{x}_{n-i}^- = x_i | \tilde{x}_{n-i-1}^- = x_{i+1}] = g^{-1/2}(x_i) \det\left[\frac{\partial \mathrm{d}w^a}{\partial x^\mu}\right](x_i, x_{i+1}, \bar{\lambda}_i) \Pr[\mathrm{d}\tilde{w}_{n-i-1}^a = \mathrm{d}w^a(x_i, x_{i+1}, \bar{\lambda}_i)]. \tag{B7}$$

It is easy to calculate the Jacobi matrix of $\mathrm{d}w^a$,

$$\begin{aligned}
\frac{\partial \mathrm{d}w^a}{\partial x^\mu}(x_{i+1}, x_i, \bar{\lambda}_i) &= (\hat{R}^{-1})^a{}_\mu|_{\bar{x}_i} + \frac{1}{2}\left\{\partial_\mu(\hat{R}^{-1})^a{}_\nu v_i^\nu - \partial_\mu[(\hat{R}^{-1})^a{}_\nu \hat{F}^\nu(\bar{\lambda}_i)]\right\}_{\bar{x}_i} \mathrm{d}t, \\
\frac{\partial \mathrm{d}w^a}{\partial x^\mu}(x_i, x_{i+1}, \bar{\lambda}_i) &= (\hat{R}^{-1})^a{}_\mu|_{\bar{x}_i} + \frac{1}{2}\left\{-\partial_\mu(\hat{R}^{-1})^a{}_\nu v_i^\nu - \partial_\mu[(\hat{R}^{-1})^a{}_\nu \hat{F}^\nu(\bar{\lambda}_i)]\right\}_{\bar{x}_i} \mathrm{d}t,
\end{aligned}$$

where $v_i := (x_{i+1} - x_i)/\mathrm{d}t$ is the velocity of the Brownian particle. Using eq. (B3), we can get

$$\lim_{n\to+\infty} \prod_{i=0}^{n-1} \frac{g^{-1/2}(x_{i+1}) \det\left[\frac{\partial \mathrm{d}w^a}{\partial x^\mu}\right](x_{i+1}, x_i, \bar{\lambda}_i)}{g^{-1/2}(x_i) \det\left[\frac{\partial \mathrm{d}w^a}{\partial x^\mu}\right](x_i, x_{i+1}, \bar{\lambda}_i)} = \exp\left[-\int_{t_I}^{t_F} v^\nu (\hat{R}^{-1})^a{}_\nu \nabla_\mu \hat{R}^\mu{}_a \mathrm{d}t\right]. \tag{B8}$$

Therefore, using eq. (3), we arrive at the following continuum limit,

$$\lim_{n\to+\infty} \prod_{i=0}^{n-1} \frac{\Pr[\mathrm{d}\tilde{w}_i^a = \mathrm{d}w^a(x_{i+1}, x_i, \bar{\lambda}_i)]}{\Pr[\mathrm{d}\tilde{w}_{n-i-1}^a = \mathrm{d}w^a(x_i, x_{i+1}, \bar{\lambda}_i)]} = \exp\left[\frac{1}{T}\int_{t_I}^{t_F} v_\mu f_{\mathrm{ex}}^\mu \mathrm{d}t\right] \exp\left[\int_{t_I}^{t_F} v^\nu (\hat{R}^{-1})^a{}_\nu \nabla_\mu \hat{R}^\mu{}_a \mathrm{d}t\right]. \tag{B9}$$

Finally, using eqs. (B6)-(B7) and (B9), we get

$$\lim_{n\to+\infty} \prod_{i=0}^{n-1} \frac{\Pr[\tilde{x}_{i+1} = x_{i+1} | \tilde{x}_i = x_i]}{\Pr[\tilde{x}_{n-i}^- = x_i | \tilde{x}_{n-i-1}^- = x_{i+1}]} = \exp\left[\frac{1}{T}\int_{t_I}^{t_F} v_\mu f_{\mathrm{ex}}^\mu \mathrm{d}t\right],$$

which is essentially identical to eq. (23).

2. Underdamped case

Similarly, we write down the discrete time version of eqs. (1) and (9),

$$\begin{aligned}
\tilde{x}_{i+1}^\mu - \tilde{x}_i^\mu &= \frac{\tilde{p}_{i+1}^\mu + \tilde{p}_i^\mu}{2m}\mathrm{d}t, \\
\tilde{p}_{i+1}^\mu - \tilde{p}_i^\mu &= F^\mu(\bar{\lambda}_i)|_{(\tilde{X}_{i+1}+\tilde{X}_i)/2}\mathrm{d}t + \bar{F}^\mu|_{(\tilde{X}_{i+1}+\tilde{X}_i)/2}\mathrm{d}t + R^\mu{}_a|_{(\tilde{X}_{i+1}+\tilde{X}_i)/2}\mathrm{d}\tilde{w}_i^a.
\end{aligned}$$

Now $\tilde{x}_{i+1}$ should be viewed as a function in $\tilde{p}_{i+1}$ and $\tilde{X}_i$,

$$\tilde{x}_{i+1} = x(\tilde{p}_{i+1}, \tilde{X}_i) = \frac{\tilde{p}_{i+1} + \tilde{p}_i}{2m}\mathrm{d}t + \tilde{x}_i.$$

Denoting $\Delta(X_{i+1}, X_i) = \delta^d(x_{i+1} - x_i - (p_{i+1} + p_i)\mathrm{d}t/2m)$, we have the conditional probability

$$\Pr[\tilde{x}_{i+1} = x_{i+1} | \tilde{p}_{i+1} = p_{i+1}, \tilde{X}_i = X_i] = g^{-1/2}(x_{i+1})\Delta(X_{i+1}, X_i).$$

Notice that $\Delta(X_{i+1}, X_i) = \Delta(I(X_i), I(X_{i+1}))$. Defining

$$\mathrm{d}w^a(X, Y, \lambda) := (R^{-1})^a{}_\nu|_{\bar{X}}\left[p^\nu - k^\nu - F^\nu(\lambda)|_{\bar{X}}\mathrm{d}t + \bar{F}^\nu|_{\bar{X}}\mathrm{d}t\right],$$

where $X = (x, p)$, $Y = (y, k)$ and $\bar{X} = (X+Y)/2$, we have

$$\mathrm{d}\tilde{w}_i^a = \mathrm{d}w^a(\tilde{X}_{i+1}, \tilde{X}_i, \bar{\lambda}_i) = \mathrm{d}w^a(\tilde{p}_{i+1}, x(\tilde{p}_{i+1}, \tilde{X}_i), \tilde{X}_i, \bar{\lambda}_i).$$

Similarly, for the reversed process, we have

$$\mathrm{d}\tilde{w}_{n-i-1}^a = \mathrm{d}w^a(\tilde{p}_{n-i}^-, x(\tilde{p}_{n-i}^-, \tilde{X}_{n-i-1}^-), \tilde{X}_{n-i-1}^-, \bar{\lambda}_i).$$

Now introduce the matrix

$$T^a{}_\mu(X, Y, \lambda) := \frac{\partial \mathrm{d}w^a}{\partial p^\mu}(X, Y) + \frac{\partial x^\nu}{\partial p^\mu}\frac{\partial \mathrm{d}w^a}{\partial x^\nu}(X, Y)$$

$$= \frac{\partial \mathrm{d}w^a}{\partial p^\mu}(X,Y) + \frac{1}{2m}\frac{\partial \mathrm{d}w^a}{\partial x^\mu}(X,Y)\mathrm{d}t.$$

The transition probabilities of the forward and reversed processes can be written as

$$\begin{aligned}
&\Pr[\tilde{X}_{i+1} = X_{i+1}|\tilde{X}_i = X_i] = \Pr[\tilde{x}_{i+1} = x_{i+1}|\tilde{p}_{i+1} = p_{i+1}, \tilde{X}_i = X_i]\Pr[\tilde{p}_{i+1} = p_{i+1}|\tilde{X}_i = X_i],\\
&\quad = g^{-1}(x_{i+1})\Delta(X_{i+1}, X_i)\det[T^a{}_\mu](X_{i+1}, X_i, \bar{\lambda}_i)\Pr[\mathrm{d}\tilde{w}^a_i = \mathrm{d}w^a(X_{i+1}, X_i, \bar{\lambda}_i)],
\end{aligned} \tag{B10}$$

and

$$\begin{aligned}
&\Pr[\tilde{X}^-_{n-i} = I(X_i)|\tilde{X}^-_{n-i-1} = I(X_{i+1})] = g^{-1}(x_i)\Delta(I(X_i), I(X_{i+1}))\det[T^a{}_\mu](I(X_i), I(X_{i+1}), \bar{\lambda}_i)\\
&\quad \times \Pr[\mathrm{d}\tilde{w}^a_i = \mathrm{d}w^a(I(X_i), I(X_{i+1}), \bar{\lambda}_i)].
\end{aligned} \tag{B11}$$

The Jacobi matrices in $T^a{}_\mu(X,Y,\lambda)$ can be explicitly calculated, yielding

$$\begin{aligned}
T^a{}_\mu(X_{i+1}, X_i, \bar{\lambda}_i) &= (R^{-1})^a{}_\mu|_{\bar{X}_i} + \frac{1}{2}\left\{m\frac{\partial}{\partial p^\mu}(R^{-1})^a{}_\nu A^\nu_i - \frac{\partial}{\partial p^\mu}[(R^{-1})^a{}_\nu F^\nu(\bar{\lambda}_i)] - \frac{\partial}{\partial p^\mu}[(R^{-1})^a{}_\nu \bar{F}^\nu]\right\}_{\bar{X}_i}\mathrm{d}t,\\
T^a{}_\mu(I(X_i), I(X_{i+1}), \bar{\lambda}_i) &= (R^{-1})^a{}_\mu|_{\bar{X}_i} + \frac{1}{2}\left\{-m\frac{\partial}{\partial p^\mu}(R^{-1})^a{}_\nu A^\nu_i + \frac{\partial}{\partial p^\mu}[(R^{-1})^a{}_\nu F^\nu(\bar{\lambda}_i)] - \frac{\partial}{\partial p^\mu}[(R^{-1})^a{}_\nu \bar{F}^\nu]\right\}_{\bar{X}_i}\mathrm{d}t,
\end{aligned}$$

where $A^\mu_i := (p^\mu_{i+1} - p^\mu_i)/(m\mathrm{d}t)$ is the coordinates acceleration, and terms of order $O(\mathrm{d}t^2)$ are omitted. Using these results together with eq. (B3), we can get

$$\begin{aligned}
&\lim_{n\to+\infty}\prod_{i=0}^{n-1}\frac{\det[T^a{}_\mu](X_{i+1}, X_i, \bar{\lambda}_i)}{\det[T^a{}_\mu](I(X_i), I(X_{i+1}), \bar{\lambda}_i)}\frac{g^{-1}(x_{i+1})}{g^{-1}(x_i)}\\
&= \exp\left[\int_{t_I}^{t_F}\left(R^\mu{}_a\frac{\partial}{\partial p^\mu}[(R^{-1})^a{}_\nu(ma^\nu - f^\nu_{\mathrm{ex}})] + \frac{2}{m}\Gamma^\mu{}_{\mu\alpha}p^\alpha\right)\mathrm{d}t\right]\exp\left[-\int_{t_I}^{t_F} g^{-1}\frac{p^\mu}{m}\frac{\partial}{\partial x^\mu}g\mathrm{d}t\right]\\
&= \exp\left[\int_{t_I}^{t_F} R^\mu{}_a\frac{\partial}{\partial p^\mu}(R^{-1})^a{}_\nu(ma^\nu - f^\nu_{\mathrm{ex}})\mathrm{d}t\right],
\end{aligned}$$

where $a^\mu = A^\mu + \Gamma^\mu{}_{\alpha\beta}p^\alpha p^\beta/m$ is covariant acceleration. Using eq.(3), the following continuum limit can be derived,

$$\begin{aligned}
&\lim_{n\to+\infty}\prod_{i=0}^{n-1}\frac{\Pr[\mathrm{d}\tilde{w}^a_i = \mathrm{d}w^a(X_{i+1}, X_i, \bar{\lambda}_i)]}{\Pr[\mathrm{d}\tilde{w}^a_i = \mathrm{d}w^a(I(X_i), I(X_{i+1}), \bar{\lambda}_i)]}\\
&\quad = \exp\left[-\frac{1}{T}\int_{t_I}^{t_F}\frac{p_\mu}{m}(ma^\mu - f^\mu_{\mathrm{ex}})\mathrm{d}t\right]\exp\left[-\int_{t_I}^{t_F} R^\mu{}_a\frac{\partial}{\partial p^\mu}(R^{-1})^a{}_\nu(ma^\nu - f^\nu_{\mathrm{ex}})\right].
\end{aligned} \tag{B12}$$

Finally, using eqs. (B10) and (B11), we get

$$\begin{aligned}
&\lim_{n\to+\infty}\prod_{i=0}^{n-1}\frac{\Pr[\tilde{X}_{i+1} = X_{i+1}|\tilde{X}_i = X_i]}{\Pr[\tilde{X}^-_{n-i} = I(X_i)|\tilde{X}^-_{n-i-1} = I(X_{i+1})]} = \exp\left[-\frac{1}{T}\int_{t_I}^{t_F}\frac{p_\mu}{m}(ma^\mu - f^\mu_{\mathrm{ex}})\mathrm{d}t\right]\\
&= \exp\left[-\frac{1}{T}\int_{t_I}^{t_F}\mathrm{d}t\frac{p^\mu}{m}\nabla_\mu(\mathcal{T} + U) + \frac{1}{T}\int_{t_I}^{t_F}\mathrm{d}t\frac{p_\mu}{m}f^\mu_{\mathrm{noc}}\right].
\end{aligned}$$

This result is identical to eq. (19).