



# Stochastic Limit-Cycle Oscillations of a Nonlinear System Under Random Perturbations

Yu-Chen Cheng<sup>1</sup> · Hong Qian<sup>1</sup>

Received: 19 October 2020 / Accepted: 9 February 2021 / Published online: 25 February 2021  
© The Author(s), under exclusive licence to Springer Science+Business Media, LLC part of Springer Nature 2021

## Abstract

Dynamical systems with  $\varepsilon$  small random perturbations appear in both continuous mechanical motions and discrete stochastic chemical kinetics. The present work provides a detailed analysis of the central limit theorem (CLT), with a time-inhomogeneous Gaussian process, near a deterministic limit cycle in  $\mathbb{R}^n$ . Based on respectively the theory of random perturbations of dynamical systems and the WKB approximation that codes the large deviations principle (LDP), results are developed in parallel from both standpoints of stochastic trajectories and transition probability density and their relations are elucidated. We show rigorously the correspondence between the local Gaussian fluctuations and the curvature of the large deviation rate function near its infimum, connecting the CLT and the LDP of diffusion processes. We study uniform asymptotic behavior of stochastic limit cycles through the interchange of limits of time  $t \rightarrow \infty$  and  $\varepsilon \rightarrow 0$ . Three further characterizations of stochastic limit cycle oscillators are obtained: (i) An approximation of the probability flux near the cycle; (ii) Two special features of the vector field for the cyclic motion; (iii) A local entropy balance equation along the cycle with clear physical meanings. Lastly and different from the standard treatment, the origin of the  $\varepsilon$  in the theory is justified by a novel scaling hypothesis via constructing a sequence of stochastic differential equations.

**Keywords** Stochastic limit cycles · Central limit theorem · Large deviation principle · Random perturbations of dynamical systems · WKB approximation · Entropy balance · Scaling hypothesis

**Mathematics Subject Classification** 37H05 · 60G07 · 60G15 · 82C31

---

Communicated by Udo Seifert.

---

H.Q. is partially supported by the Olga Jung Wan Endowed Professorship.

---

✉ Yu-Chen Cheng  
yuchench@uw.edu

Hong Qian  
hqian@uw.edu

<sup>1</sup> Department of Applied Mathematics, University of Washington, Seattle, WA 98195-3925, USA

## 1 Introduction

*Newtonian mechanics* represents the world in terms of featureless point masses with their positions and momenta. In contrast, classical *chemical kinetics* represents the world in terms of the number densities of interacting populations of individual molecules, each with a large internal degrees of freedom, as chemical species. What is possibly an appropriate representation for complex biological systems and processes? To answer this question, it is necessary to give a more precise meaning to the too widely used term “complex” [1]. Let us consider one class of complex systems, the living biological cells in terms of a biochemical kinetic description. In this paradigm, a *complex system* consists of many interacting sub-populations of individuals with stochastic state transitions; the system as a whole actively exchanges matters, energy, or information with its environment [2]. One sees a remarkable resemblance between this kinetic description of cells and many other biological systems with complex “individuals”. In fact, the biochemist’s perspective captures a repeated hierarchical structure of the complex world: An ecological system is a community of various biological organism; a human body consists of over 30 trillion cells; and a cell involves a large number of interacting non-living biopolymers. This view echos the philosophy of P. W. Anderson’s hierarchical structure of science [3].

While the “stochasticity” in chemical kinetics mainly originates from internal states of individual macromolecules, uncertainties in mechanical motions in biology, such as protein motor proteins in axonal transport and hemodynamics of cardiovascular systems, are chiefly a consequence of *coarse graining*: A highly complex many-body systems can be represented by simple statistical laws. One of the best examples of this is Kramers’ rate theory for barrier crossing between two basins, which condenses a very complex dynamics into a simple exponentially distributed time with a single parameter. A problem becomes simple if we focus on the emergent behavior of an assembly of a large numbers of atoms at the macroscopic scale with a much longer time scale. Indeed, experimentalists would find that the macromolecular movement obeys simple laws under certain approximations, for example Fick’s law. The bridge between complexity and simplicity is *uncertainty and its statistics*. This is the fundamental idea of the theory of Brownian motion [4].

### 1.1 Stochastic Models of Complex Systems

As can be seen from the above discussion, both representations have their own values for complex systems. Once we choose one of them to describe a system of interest, then the following question is what mathematical model we should adopt. In stochastic chemical kinetics, there is a success of the well-established scaling hypothesis in the continuous-time non-negative integer valued Markov jump process [5]. Consider a continuous stirred chemical reaction vessel of volume  $V$ , in which the numbers of molecules of various species  $\mathbf{n}_V(t)$  is a Markov jump process that can be described by a *master equation*

$$\frac{\partial P(\mathbf{n}_V, t)}{\partial t} = \sum_{\mathbf{r}} [W(\mathbf{n}_V - \mathbf{r}, \mathbf{r})P(\mathbf{n}_V - \mathbf{r}, t) - W(\mathbf{n}_V, \mathbf{r})P(\mathbf{n}_V, t)], \quad (1.1)$$

where  $W(\mathbf{n}_V, \mathbf{r})$  is the transition probability per unit time from  $\mathbf{n}_V$ ,  $\mathbf{n}_V + \mathbf{r}$ , and both  $\mathbf{n}_V$  and  $\mathbf{r}$  are  $q$ -dimensional vectors. As the system’s size  $V \rightarrow \infty$ ,  $\mathbf{n}_V(t)$  follows the law of large number,  $V^{-1}\mathbf{n}_V(t) \rightarrow \mathbf{c}(t)$ , the concentration of  $q$  species.

With a proper scaling by the size  $V$  and the assumption that  $W$  and  $P$  are smooth enough functions, we can take the Kramers-Moyal expansion of the master equation (1.1) [6–8]

$$\varepsilon \frac{\partial p(\mathbf{x}, t)}{\partial t} = \sum_{\mathbf{k}} \binom{1}{\mathbf{k}!} \left( \varepsilon \frac{\partial}{\partial \mathbf{x}} \right)^{\mathbf{k}} [\alpha_{\mathbf{k}}(\mathbf{x}) p(\mathbf{x}, t)], \tag{1.2}$$

where  $\varepsilon = 1/V$ ,  $\mathbf{x} = \mathbf{n}_V/V$ ,  $p(\mathbf{x}, t) = VP(\mathbf{n}_V, t)$ , and  $\mathbf{k} = (k_1, k_2, \dots, k_q)$ ,  $\sum_{\mathbf{k}} = \sum_{k_q} \dots \sum_{k_2} \sum_{k_1}$ ,  $\mathbf{k}! = \prod_i (k_i!)$ , and  $\alpha_{\mathbf{k}}(\mathbf{x}) = \sum_{\mathbf{r}} (\prod_i r_i^{k_i}) w(\mathbf{x}, \mathbf{r})$ ,  $w(\mathbf{x}, \mathbf{r}) = W(\mathbf{X}, \mathbf{r})/V$ . The solutions of the differential equation (1.2) with the infinite terms represents the exact time-dependent probability density of the scaled number of population  $\mathbf{n}_V/V$ . Then a natural question arises: could we obtain a corresponding diffusion process from this infinite order differential equation? The truth is that we can only get a ‘‘local diffusion process approximation’’ for the scaled Markov jump process due to the following reason.

A common method to attack Eq. (1.2) is by truncating the higher order terms to the second order of  $\varepsilon$  to obtain a Fokker-Planck equation (FPE). However, van Kampen [9] pointed out that this method may fail if  $\mathbf{n}_V(t)$  has large sizes of jumps. A concrete example was provided in the work [10]: There exists an inconsistency between the stationary solutions of Kramers-Moyal FPE and the original master equation for the Schlögl’s model of a chemical reaction system which has bistable steady states. The reason for the failure of Kramers-Moyal FPE is that we are only able to observe either the deterministic behavior of the process at the scale of the law of large number or the Gaussian fluctuations at the scale of the central limit theorem; however, there is no one scale to obtain both. To keep the first two order terms of the Kramers-Moyal expansion simultaneously to represent a diffusion process at a single scale is incorrect. Therefore, van Kampen [9] suggested the  $\Omega$  expansion which allows us to get a deterministic trajectory as  $\varepsilon \rightarrow 0$  and a local approximation near the deterministic trajectory at the scale  $O(\sqrt{\varepsilon})$  separately.

In the present work, we focus on the continuous representation of complex systems. We always start with random perturbations of dynamical systems represented by a sequence of stochastic differential equations (SDEs) parameterized by a small parameter  $\varepsilon$

$$d\mathbf{X}_\varepsilon(t) = \mathbf{b}(\mathbf{X}_\varepsilon)dt + [2\varepsilon\mathbf{D}]^{\frac{1}{2}}d\mathbf{B}(t), \tag{1.3}$$

where  $\mathbf{X}_\varepsilon \in \mathbb{R}^n$ ,  $\mathbf{b} : \mathbb{R}^n \rightarrow \mathbb{R}^n$  stands for a drift function, the  $\mathbb{R}^n \times \mathbb{R}^n$  diffusion matrix  $\mathbf{D}$  is constant and positive semidefinite symmetric, and  $\mathbf{B}$  is the standard  $n$ -dimensional Brownian motion. This Langevin type equation is widely applicable for complex systems related to mechanics, and it gives us a clear picture of the entire dynamics including the drift and diffusion at one scale. Furthermore, by the rigorous mathematical theory of semigroup [11], every diffusion process represented by a SDE has a unique FPE to characterize the corresponding transition probability density  $p_\varepsilon(\mathbf{x}, t)$

$$\frac{\partial p_\varepsilon}{\partial t} = -\nabla \cdot \mathbf{J}[p_\varepsilon], \quad \mathbf{J}[p_\varepsilon] \equiv \mathbf{b}(\mathbf{x})p_\varepsilon - \varepsilon\mathbf{D}\nabla p_\varepsilon. \tag{1.4}$$

This line of reasoning to relate diffusion processes and FPEs has no ambiguity unlike the Kramers–Moyal FPEs.

### 1.2 Random Perturbations of Diffusion Processes

Our analysis of random perturbations of diffusion processes is by expansion in powers of  $\varepsilon$  for the sequence of SDEs (1.3), which follows the work of Freidlin and Wentzell [12].

As  $\varepsilon \rightarrow 0$ , the sequence of SDEs converges to an ordinary differential equation (ODE) of the emergent deterministic trajectory by the Law of large number (LLN). To shift the sequence of SDEs to its deterministic trajectory with normalization by the scale  $O(\sqrt{\varepsilon})$ , the rescaled sequence of SDEs converges to a time-inhomogeneous Gaussian process by the central limit theorem (CLT). For rare events in  $O(1)$ , they have the probability asymptotic to zero exponentially fast by the Large deviation principle (LDP). In comparison to Freidlin and Wentzell, there is another celebrated theory for the LDP by Donsker and Varadhan [13–16]. The main difference between them is that the Freidlin-Wentzell theory is about large deviations from a deterministic trajectory by small noise but the Donsker-Varadhan theory is regarding large deviations of certain process expectations for large time with the ergodic theorem.

In the present paper, we provide a trajectory-based proof for an emergent time-inhomogeneous Gaussian process in  $\mathbb{R}^n$  near a deterministic trajectory under the CLT, which follows the proof for the particular case of  $\mathbb{R}^1$  in the Freidlin-Wentzell's textbook [12] (The idea of proof for  $\mathbb{R}^n$  was suggested in the book but without details) and we further obtain a *Lyapunov differential equation* for the covariance of this Gaussian process. In the field of statistical physics, this Lyapunov differential equation was mentioned in the works [9, 17, 18]. However, all of those previous works were based on the small noise expansion of the associated FPEs and each approach has some limitations: In [9, 18], the dynamics was restricted to one dimension; In [17], the dynamics was for elementary processes in chemical reactions. Our approach for the Lyapunov differential equation is trajectory-based without transferring the original SDE problem to the problem of perturbations of partial differential equations (PDEs) and it is applicable for rather general multi-dimensional diffusion processes.

In contradistinction to the Freidlin-Wentzell theory and the Donsker-Varadhan theory, which are both from the standpoint of trajectories of systems, there is another approach of the LDP based on the PDEs: A logarithmic transformation to the differential generator of diffusion processes was proposed by Fleming in 1978 [19] then the PDE-based approach was applied to the LDP through solving the Hamilton–Jacobian equations (HJEs) by Evans and Ishii [20] and others. Feng and Kurtz [21] generalized this approach by refining techniques on the viscosity solutions of HJEs so that the scope of applications of it is compatible with the Freidlin-Wentzell theory and the Donsker-Varadhan theory. This rigorous mathematical PDE-based approach is corresponding to the WKB method of solving FPEs, which was introduced early by theoretical physicists [22, 23]. In the present work, our analysis of stochastic limit cycles is carried out in parallel with both the small random perturbations of SDEs and the WKB approximation of the corresponding transition probability density, which can be regarded as an example of a link between the trajectory-based and the PDE-based methods. The contradistinction provides a more comprehensive portrayal of the stochastic limit cycle.

### 1.3 Time-Inhomogeneous Gaussian Processes from a Transient State to an Invariant Set

By relating those two methods, one of the important results obtained in this paper is the correspondence between the local Gaussian fluctuations along a deterministic path, limit cycle or not, and the curvature of the leading order term in the WKB method near its infimum. In the early works, the connection between the CLT and the LDP of random processes can be found in the analysis of action functional for Gaussian random processes [12] and the LDP for the empirical measures of centered stationary Gaussian processes [24, 25], in which the former follows the Freidlin-Wentzell theory and the later follows the Donsker-Varadhan

theory. In the present paper, our work on the analysis of the CLT and the LDP of nonlinear systems with stochastic limit cycles, from a transient state to infinite time limit on an invariant set, are beyond those theories.

Globally, from the standpoint of probability, the existence of stationary distribution in the whole space for a stochastic stable limit cycle has been proved by Holland [26]. With the WKB method, characterizations of the stationary large deviation rate function near the cycle were studied in the previous work [27–31]. Locally, from the standpoint of trajectories, the dynamics is attracted to an invariant set but still capable of escaping from the set due to the multi-dimensional fluctuations except the part tangential to the cycle. In the long run, the Gaussian fluctuations along the direction tangential to the cycle is eventually smeared out and the rest of fluctuations in the hyperplane perpendicular to the cycle are outward and damped out by the dissipation toward the limit cycle [32].

In this paper, equipped with the Lyapunov equation for the covariance of the time-inhomogeneous Gaussian process, we characterize the fluctuations along the limit cycle by *asymptotic analysis*. Via a careful study of the interchange of limits of time  $t \rightarrow \infty$  and  $\varepsilon \rightarrow 0$ , with a coordinate transformation and dimension reduction on the cycle, we show that the Lyapunov equation becomes a  $n - 1 \times n - 1$  *periodic Riccati differential equation* [33–36] and the solution of equation is a positive definite matrix. We further characterize the curvature of large deviation rate function on the limit cycle by the correspondence between the covariance matrix and the curvature established in our theory.

The importance of stochastic limit cycle oscillations in physics was emphasized by Keizer [17] and van Kampen [9]. In their books, specific examples with careful studies were provided but a general analysis was missing. Our analysis by both the trajectory-based and probability-based methods, from a transient state to an invariant set, helps us to paint a clear picture of dynamics in different scopes of space and time. Additionally, the present work can be regarded as an extension of the linear approximation theory of a stochastic nonlinear system with a fixed point as the steady state, which can be found in [37–39], to an invariant set in  $\mathbb{R}^n$ .

Furthermore, we want to point out that stochastic limit cycles have been widely studied and applied to biological systems in two different scenarios: (1) the finite fluctuation results [40–45] and (2) the zero noise limit with the WKB large deviation [43,46–48]. Our work includes the following new results beyond the scopes of those two: First, we apply CLT to describe the local fluctuations around the deterministic trajectory. It provides us a new scope to investigate the limiting behavior of finite noise dynamics around its most probable path, and this scale is different from the scale of WKB large deviation. Second, for the WKB method, we include not only the large-deviation rate function but also the prefactor, which is the next order to the leading order of the large-deviation rate function. In the present work, we show that the prefactor plays an important role in stochastic limit cycles since the leading order term vanishes on the deterministic trajectory of limit cycles.

## 1.4 Organization of the Paper

In Sect. 2, we start with a rather general small-noise diffusion process represented by a sequence of non-linear multidimensional SDEs. Based on both the trajectory-based approach and the WKB method, key lemmas regarding the time-inhomogeneous Gaussian processes and a link to the large deviation rate function are provided. In Sect. 3, we apply the lemmas to stochastic limit cycle oscillators. This approach is distinct to the previous works [27–29]: (i) The works [27,28] are regarding fluctuations of limit cycles in chemical systems (The former [27] focused on analysis of a stationary FPE and the later [28] applied the WKB method

directly to a master equation.); (ii) The work [29] focused on the case of one-dimensional motion on a circle. In Sect. 4, we introduce the scaling hypothesis of diffusion processes to construct a sequence of dynamics parameterized by  $\varepsilon$ . This scaling hypothesis not only serves as an useful mathematical tool of asymptotic analysis but also a scientific theory to justify the origin of  $\varepsilon$ .

## 2 Preliminaries

As we mentioned in Sect. 1, both discrete chemical kinetics and continuous mechanical motions successfully depict complex systems via introducing uncertainty. Based on the probability theory, the former is conventionally characterized by Markov jump processes (continuous-time and discrete-state) with the corresponding transition probability captured by master equations, and the later is popularly described by diffusion processes (continuous-time and continuous-state) with SDEs. By introducing a parameter of the size of systems, at proper scales, both representations have their corresponding FPEs of the transition probability density and the HJEs of the large deviation rate function. In the present work, based on the continuous representation, we follow the direction SDE–FPE–HJE in a sequence.

### 2.1 Expansion in Powers of a Small Parameter for Diffusion Processes

Let us start from a sequence of SDEs defined in Eq. (1.3)

$$d\mathbf{X}_\varepsilon(t) = \mathbf{b}(\mathbf{X}_\varepsilon)dt + [2\varepsilon\mathbf{D}]^{\frac{1}{2}}d\mathbf{B}(t), \tag{2.1}$$

and by the LLN, it converges to the following ordinary differential equation (ODE) as  $\varepsilon \rightarrow 0$

$$d\mathbf{x}(t) = \mathbf{b}(\mathbf{x})dt. \tag{2.2}$$

Let  $\hat{\mathbf{x}}(t)$  be the solution of this ODE with a given initial condition  $\hat{\mathbf{x}}(0) = \hat{\mathbf{x}}_0$ .

We shall note that a direct application of small noise expansions for the process (2.1) by  $\mathbf{X}_\varepsilon(t) = \sum_{i=0}^n \varepsilon^i \mathbf{X}_i(t)$  may fail for certain types of drift functions  $\mathbf{b}$  [18]. Therefore, we need to expand  $\mathbf{X}_\varepsilon(t)$  with a proper scale: By the scale of the CLT, we can define a random process  $\mathbf{Z}_\varepsilon(t)$  near the deterministic trajectory  $\hat{\mathbf{x}}(t)$

$$\mathbf{Z}_\varepsilon(t) \equiv \frac{\mathbf{X}_\varepsilon(t) - \hat{\mathbf{x}}(t)}{\sqrt{\varepsilon}} \tag{2.3}$$

and substitute Eqs. (2.3) into (2.1), we can derive that

$$\begin{aligned} d\mathbf{X}_\varepsilon(t) &= d\hat{\mathbf{x}}(t) + \sqrt{\varepsilon}d\mathbf{Z}_\varepsilon(t) = \mathbf{b}(\mathbf{X}_\varepsilon)dt + [2\varepsilon\mathbf{D}]^{\frac{1}{2}}d\mathbf{B}(t) + O(\varepsilon) \\ \Rightarrow d\mathbf{X}_\varepsilon(t) &= d\hat{\mathbf{x}}(t) + \sqrt{\varepsilon}d\mathbf{Z}_\varepsilon(t) = (\mathbf{b}(\hat{\mathbf{x}}) + \sqrt{\varepsilon}\mathbf{A}(\hat{\mathbf{x}})\mathbf{Z}_\varepsilon)dt + [2\varepsilon\mathbf{D}]^{\frac{1}{2}}d\mathbf{B}(t) + O(\varepsilon) \\ \Rightarrow d\mathbf{Z}_\varepsilon(t) &= \mathbf{A}(\hat{\mathbf{x}})\mathbf{Z}_\varepsilon dt + [2\mathbf{D}]^{\frac{1}{2}}d\mathbf{B}(t) + O(\sqrt{\varepsilon}), \end{aligned} \tag{2.4}$$

where  $\mathbf{A}(\hat{\mathbf{x}}(t))$  is the Jacobian matrix of  $\mathbf{b}(\mathbf{x})$  evaluated at  $\mathbf{x} = \hat{\mathbf{x}}(t)$ . We then follow the usual approach [12] of perturbation theory to obtain an expansion in powers of the small parameter  $\sqrt{\varepsilon}$

$$\mathbf{Z}_\varepsilon(t) = \mathbf{Z}(t) + \sqrt{\varepsilon}\mathbf{Z}^{(1)}(t) + \dots + \sqrt{\varepsilon}^k\mathbf{Z}^{(k)}(t) + \dots. \tag{2.5}$$

Apply the expansion of  $\mathbf{Z}_\varepsilon(t)$  in Eq. (2.5) to its SDE in Eq. (2.4), we can obtain a SDE for the zeroth approximation of  $\mathbf{Z}_\varepsilon(t)$

$$d\mathbf{Z}(t) = \mathbf{A}(\hat{\mathbf{x}})\mathbf{Z}dt + [2\mathbf{D}]^{\frac{1}{2}}d\mathbf{B}(t). \tag{2.6}$$

The following lemma is about the solution of  $\mathbf{Z}(t)$ :

**Lemma 2.1** *If each element of the Jacobian matrix  $\mathbf{A}(\hat{\mathbf{x}}(t))$  is continuous for all  $t \geq 0$ , then for every  $t > 0$ ,  $\mathbf{Z}(t)$  is a Gaussian random variable  $\mathbf{Z}(t) \sim \mathcal{N}(\boldsymbol{\mu}(t), \boldsymbol{\Sigma}(t))$  with*

$$\frac{d\boldsymbol{\mu}(t)}{dt} = \mathbf{A}(\hat{\mathbf{x}})\boldsymbol{\mu}, \quad \boldsymbol{\mu}(0) = \hat{\boldsymbol{\mu}}_0, \tag{2.7}$$

$$\frac{d\boldsymbol{\Sigma}(t)}{dt} = \mathbf{A}(\hat{\mathbf{x}})\boldsymbol{\Sigma} + \boldsymbol{\Sigma}\mathbf{A}(\hat{\mathbf{x}})^T + 2\mathbf{D}, \quad \boldsymbol{\Sigma}(0) = \hat{\boldsymbol{\Sigma}}_0, \tag{2.8}$$

where  $\boldsymbol{\mu}_0$  and  $\boldsymbol{\Sigma}_0$  are given initial conditions.

**Proof** Under the assumption that each element of  $\mathbf{A}(\hat{\mathbf{x}}(t))$  is continuous for all  $t \geq 0$ , there exists a fundamental matrix  $\mathbf{M}(t) \in \mathbb{R}^n \times \mathbb{R}^n$  satisfied the linear homogenous ordinary differential equation

$$d\mathbf{M}(t) = \mathbf{A}(\hat{\mathbf{x}})\mathbf{M}dt \quad \text{for all } t > 0. \tag{2.9}$$

Let  $\mathbf{Z}_0$  be the given initial condition for the dynamics (2.6), we can verify the equation

$$\mathbf{Z}(t) = \mathbf{M}(t) \left( \mathbf{Z}_0 + \int_0^t \mathbf{M}^{-1}(s)[2\mathbf{D}]^{\frac{1}{2}}d\mathbf{B}(s) \right) \quad \text{for all } t > 0, \tag{2.10}$$

by differentiating the both sides of it with the Itô lemma and Eqs. (2.6) and (2.9) as follows

$$\begin{aligned} & d \left[ \mathbf{M}(t) \left( \mathbf{Z}_0 + \int_0^t \mathbf{M}^{-1}(s)[2\mathbf{D}]^{\frac{1}{2}}d\mathbf{B}(s) \right) \right] \\ &= \mathbf{M}d \left( \mathbf{Z}_0 + \int_0^t \mathbf{M}^{-1}[2\mathbf{D}]^{\frac{1}{2}}d\mathbf{B} \right) + d\mathbf{M} \left( \mathbf{Z}_0 + \int_0^t \mathbf{M}^{-1}[2\mathbf{D}]^{\frac{1}{2}}d\mathbf{B} \right) \\ & \quad + d\mathbf{M}d \left( \mathbf{Z}_0 + \int_0^t \mathbf{M}^{-1}[2\mathbf{D}]^{\frac{1}{2}}d\mathbf{B} \right) \\ &= [2\mathbf{D}]^{\frac{1}{2}}d\mathbf{B} + \mathbf{A}(\hat{\mathbf{x}})\mathbf{M}dt\mathbf{M}^{-1}\mathbf{Z} + \mathbf{A}(\hat{\mathbf{x}})\mathbf{M}dt\mathbf{M}^{-1}[2\mathbf{D}]^{\frac{1}{2}}d\mathbf{B} \\ &= [2\mathbf{D}]^{\frac{1}{2}}d\mathbf{B} + \mathbf{A}(\hat{\mathbf{x}})\mathbf{M}\mathbf{M}^{-1}\mathbf{Z}dt \\ &= d\mathbf{Z}(t). \end{aligned}$$

By Eq. (2.10), for any constant vector  $\mathbf{a} \in \mathbb{R}^n$ , we have that

$$\mathbf{a}^T \mathbf{Z}(t) = \mathbf{a}^T \mathbf{M}(t) \left( \mathbf{Z}_0 + \int_0^t \mathbf{M}^{-1}(s)[2\mathbf{D}]^{\frac{1}{2}}d\mathbf{B}(s) \right) = \sum_{i=1}^n \int_0^t f_i(s)dB_i(s), \tag{2.11}$$

where  $B_i \in \mathbb{R}^1$  is a collection independent and identically distributed random variables from the standard Brownian motion  $\mathbf{B} = (B_1, B_2, \dots, B_n)$ , and  $f_i : \mathbb{R}^1 \rightarrow \mathbb{R}^1$  is a collection of deterministic functions. Therefore,  $\mathbf{a}^T \mathbf{Z}(t)$  has to be a one-dimensional Gaussian random variable since it is a linear combination of a collection of independent one-dimensional Gaussian random variables. Furthermore, by [49], using the moment generating functions, arbitrary linear combinations of the random vector  $\mathbf{Z}(t)$  being an univariate Gaussian random variable implies that  $\mathbf{Z}(t)$  is a multivariate Gaussian random variable.

Next, we want to find expressions of the mean and the covariance of  $\mathbf{Z}(t)$  for every  $t$ . By the property of the standard Brownian motion, given a matrix  $\mathbf{F}(t)$  which is independent of  $\mathbf{B}$ , we have that

$$\mathbb{E} \left[ \int_0^t \mathbf{F}(s) d\mathbf{B}(s) \right] = \mathbf{0}, \quad \text{for all } t \geq 0. \tag{2.12}$$

With (2.10) and (2.12), the first and second moment of  $\mathbf{Z}(t)$  should satisfy

$$\begin{aligned} \mathbb{E}[\mathbf{Z}(t)] &= \mathbf{M}(t)\mathbf{Z}_0 \\ \mathbb{E}[\mathbf{Z}(t)\mathbf{Z}(t)^T] &= \mathbb{E} \left[ \mathbf{M}(t) \left( \mathbf{Z}_0 + \int_0^t \mathbf{M}^{-1}[\mathbf{2D}]^{\frac{1}{2}} d\mathbf{B} \right) \left( \mathbf{Z}_0 + \int_0^t \mathbf{M}^{-1}[\mathbf{2D}]^{\frac{1}{2}} d\mathbf{B} \right)^T \mathbf{M}(t)^T \right] \\ &= \mathbf{M}(t)\mathbf{Z}_0\mathbf{Z}_0^T\mathbf{M}(t)^T + \mathbf{M}(t)\mathbb{E} \left[ \left( \int_0^t \mathbf{M}^{-1}[\mathbf{2D}]^{\frac{1}{2}} d\mathbf{B} \right) \left( \int_0^t \mathbf{M}^{-1}[\mathbf{2D}]^{\frac{1}{2}} d\mathbf{B} \right)^T \right] \mathbf{M}(t)^T \\ &= \mathbf{M}(t)\mathbf{Z}_0\mathbf{Z}_0^T\mathbf{M}(t)^T + 2\mathbf{M}(t) \left( \int_0^t \mathbf{M}^{-1}(s)\mathbf{D}\mathbf{M}^{-T}(s)ds \right) \mathbf{M}(t)^T, \end{aligned} \tag{2.13}$$

where we applied the Itô isometry to the last equation. Since  $\boldsymbol{\mu}(t) = \mathbb{E}[\mathbf{Z}(t)] = \mathbf{M}(t)\mathbf{Z}_0$ , and  $\boldsymbol{\Sigma}(t) = \mathbb{E}[\mathbf{Z}(t)\mathbf{Z}(t)^T] - \mathbb{E}[\mathbf{Z}(t)]\mathbb{E}[\mathbf{Z}(t)^T] = 2\mathbf{M}(t) \left( \int_0^t \mathbf{M}^{-1}(s)\mathbf{D}\mathbf{M}^{-T}(s)ds \right) \mathbf{M}(t)^T$ , by taking derivatives of them with respect to time, we thus obtain dynamics of  $\boldsymbol{\mu}(t)$  and  $\boldsymbol{\Sigma}(t)$  as follows

$$\frac{d\boldsymbol{\mu}(t)}{dt} = \mathbf{A}(\hat{\mathbf{x}})\boldsymbol{\mu}, \quad \boldsymbol{\mu}(0) = \hat{\boldsymbol{\mu}}_0, \tag{2.14}$$

$$\frac{d\boldsymbol{\Sigma}(t)}{dt} = \mathbf{A}(\hat{\mathbf{x}})\boldsymbol{\Sigma} + \boldsymbol{\Sigma}\mathbf{A}(\hat{\mathbf{x}})^T + 2\mathbf{D}, \quad \boldsymbol{\Sigma}(0) = \hat{\boldsymbol{\Sigma}}_0, \tag{2.15}$$

where  $\hat{\boldsymbol{\mu}}_0$  and  $\hat{\boldsymbol{\Sigma}}_0$  are given initial conditions.

By Lemma 2.1, we obtained a time-inhomogeneous Gaussian process from a multi-dimensional nonlinear diffusion process at the scale of the CLT with the covariance captured by the Lyapunov differential equation (2.15). Additionally, this lemma can be applied to solve the FPE

$$\frac{\partial p_\varepsilon}{\partial t} = -\nabla \cdot \mathbf{J}[p_\varepsilon], \quad \mathbf{J}[p_\varepsilon] \equiv \mathbf{b}(\mathbf{x})p_\varepsilon - \varepsilon\mathbf{D}\nabla p_\varepsilon \tag{2.16}$$

with certain boundary conditions. Since the function  $\mathbf{b}$  is nonlinear and multi-dimensional, this type of PDE problems may not be easy to solve directly. Under certain conditions [12,50], the diffusion process (2.1) is associated with this FPE. By expanding the solution of the FPE (2.16) and the diffusion process (2.1) respectively

$$p_\varepsilon(\mathbf{x}, t) = \frac{1}{\sqrt{\varepsilon}} \hat{p}_\varepsilon(\mathbf{z}, t) \quad \text{and} \quad \hat{p}_\varepsilon(\mathbf{z}, t) = \hat{p}_0(\mathbf{z}, t) + \sum_{n=1}^{\infty} (\sqrt{\varepsilon})^n \hat{p}_n(\mathbf{z}, t), \tag{2.17}$$

$$\mathbf{Z}_\varepsilon(t) = \frac{\mathbf{X}_\varepsilon(t) - \hat{\mathbf{x}}(t)}{\sqrt{\varepsilon}} \quad \text{and} \quad \mathbf{Z}_\varepsilon(t) = \mathbf{Z}(t) + \sum_{n=1}^{\infty} (\sqrt{\varepsilon})^n \mathbf{Z}^{(n)}(t), \tag{2.18}$$

we can check that  $\hat{p}_0(\mathbf{z}, t)$  is the probability density of  $\mathbf{Z}(t)$ . Following from Lemma 2.1, we thus obtain an approximation solution of the FPE by having the dynamics of the mean and covariance of  $\mathbf{Z}(t)$ . To transform the FPE problem into the problem of a diffusion process, the above example is an application of Lemma 2.1 to attack a complicated boundary value problem.



## 2.2 Approximations by the Asymptotic Theory Akin to the WKB

In Sect. 1.1, we pointed out that using the Kramers-Moyal Fokker-Planck equation for a mater equation may fail in some cases. Instead, the deterministic behavior and the local fluctuation of a jump Markov process can be obtained by the  $\Omega$  expansion with respect to two different scales. In addition to the  $\Omega$  expansion, another approach by the WKB approximation has been applied to give a full analysis of master equations [8,22,28,51]. In this method, the WKB ansatz is assumed for the solution of the Kramers-Moyal expansion of a master equation without truncating higher-order terms, so this method has no problem unlike the Kramers-Moyal Fokker-Planck equation. As the success of the WKB approximation of master equations for Markov jump processes, this method has also been used to the associated Fokker-Planck equations of diffusion processes [23,52]. Here we want to link our trajectory-based approach in Sect. 2.1 to the probability-based approach by the WKB approximation.

Recall that the path behaviors of the diffusion process is described by the n-dimensional SDE (2.1)

$$d\mathbf{X}_\varepsilon(t) = \mathbf{b}(\mathbf{X}_\varepsilon)dt + \sqrt{2\varepsilon}\mathbf{D}d\mathbf{B}_t. \tag{2.19}$$

In order to link the SDE (2.19) to the WKB ansatz, we need to find its probability-density representation by a FPE. From Eqs. (2.16) to (2.18), we have illustrated a way to convert a PDE problem to a SDE problem; on the other hand, by the semigroup approaches [11,50], we can also convert a SDE problem to a PDE problem. Under certain conditions [50], the original SDE problem can be characterized by the solution of the FPE

$$\frac{\partial p_\varepsilon}{\partial t} = -\nabla \cdot \mathbf{J}[p_\varepsilon], \quad \mathbf{J}[p_\varepsilon] \equiv \mathbf{b}(\mathbf{x})p_\varepsilon - \varepsilon\mathbf{D}\nabla p_\varepsilon. \tag{2.20}$$

We shall note that, as  $\varepsilon \rightarrow 0$ , the FPE reduces to a first-order differential equation so the perturbation of the solution  $p_\varepsilon$  follows the singular perturbation theory. To attack this singular perturbation problem, we adopt an asymptotic series of  $p_\varepsilon$  with a proper scaled variable  $\mathbf{z} = (\mathbf{x} - \hat{\mathbf{x}})/\sqrt{\varepsilon}$ ,

$$p_\varepsilon(\mathbf{x}, t) = \frac{1}{\sqrt{\varepsilon}}\hat{p}_\varepsilon(\mathbf{z}, t) \quad \text{and} \quad \hat{p}_\varepsilon(\mathbf{z}, t) = \sum_{n=0}^{\infty} (\sqrt{\varepsilon})^n \hat{p}_n(\mathbf{z}, t). \tag{2.21}$$

In parallel, there is another complete asymptotic theory for the solution of FPE akin to the WKB ansatz [22,23]

$$p_\varepsilon(\mathbf{x}, t) = a(\varepsilon, t) \exp \left[ -\frac{1}{\varepsilon} \sum_{n=0}^{\infty} \phi_n(\mathbf{x}, t) \varepsilon^n \right], \tag{2.22}$$

where  $a(\varepsilon, t)$  is a normalization factor. The expansion (2.22) with the series  $\varepsilon^{-1}\phi_0 + \phi_1 + \varepsilon\phi_2 + \dots$  was justified by the *extensive property* of  $p_\varepsilon(\mathbf{x}, t)$ , i.e. it keeps the form (2.22) as time evolves [22].

We will connect those two types of expansions in Lemma 2.4. To prove the lemma, we first give two useful lemmas on the asymptotic evaluation of various integrals [53]. All the proofs can be found in Appendix 1.

**Lemma 2.2** For sufficiently smooth scalar functions  $f(\mathbf{x})$  and  $h(\mathbf{x})$ ,  $\mathbf{x} \in \mathbb{R}^n$ ,

$$\int_{\mathbb{R}^n} f(\mathbf{x})e^{-h(\mathbf{x})/\varepsilon} \, d\mathbf{x} = \sqrt{\frac{2\pi\varepsilon}{\det[\nabla\nabla h(\mathbf{x}^*)]}} e^{-\frac{h(\mathbf{x}^*)}{\varepsilon}} \left[ f(\mathbf{x}^*) + \varepsilon\eta(\mathbf{x}^*) + O(\varepsilon^2) \right], \tag{2.23}$$

$$\frac{\int_{\mathbb{R}^n} f(\mathbf{x})e^{-h(\mathbf{x})/\varepsilon} \, d\mathbf{x}}{\int_{\mathbb{R}^n} e^{-h(\mathbf{x})/\varepsilon} \, d\mathbf{x}} = f(\mathbf{x}^*) + \varepsilon \left[ \frac{f''_{ij}(\mathbf{x}^*)\Xi_{ij}}{2} - \frac{f'_i(\mathbf{x}^*)h'''_{jkl}(\mathbf{x}^*)\Xi_{i\mu}^{\frac{1}{2}}\Xi_{j\nu}^{\frac{1}{2}}\Xi_{k\rho}^{\frac{1}{2}}\Xi_{\ell\kappa}^{\frac{1}{2}}\Theta_{\mu\nu\rho\kappa}}{6} \right] + O(\varepsilon^2), \tag{2.24}$$

as  $\varepsilon \rightarrow 0$ , in which Einstein's summation rule is adopted,  $\mathbf{x}^*$  is the global minimum of  $h(\mathbf{x})$ , and

$$\begin{aligned} \eta(\mathbf{x}^*) = & \frac{f''_{ij}(\mathbf{x}^*)\Xi_{ij}}{2} - \left[ \frac{f'_i(\mathbf{x}^*)h'''_{jkl}(\mathbf{x}^*)}{6} + \frac{f(\mathbf{x}^*)h''''_{ijkl}(\mathbf{x}^*)}{24} \right] \Xi_{i\mu}^{\frac{1}{2}}\Xi_{j\nu}^{\frac{1}{2}}\Xi_{k\rho}^{\frac{1}{2}}\Xi_{\ell\kappa}^{\frac{1}{2}}\Theta_{\mu\nu\rho\kappa} \\ & + \frac{f(\mathbf{x}^*)[h''''_{ijk}(\mathbf{x}^*)]^2}{72} \Xi_{i\mu}^{-\frac{1}{2}}\Xi_{i\mu'}^{-\frac{1}{2}}\Xi_{j\nu}^{-\frac{1}{2}}\Xi_{j\nu'}^{-\frac{1}{2}}\Xi_{k\rho}^{-\frac{1}{2}}\Xi_{k\rho'}^{-\frac{1}{2}}\Lambda_{\mu\mu'\nu\nu'\rho\rho'}. \end{aligned} \tag{2.25}$$

The covariance matrix  $\Xi = [\nabla\nabla h(\mathbf{x}^*)]^{-1}$ , and the multi-indexed  $\Theta_{ijkl}$  and  $\Lambda_{\mu\mu'\nu\nu'\rho\rho'}$  are

$$\Theta_{\mu\nu\rho\kappa} = \int_{\mathbb{R}^n} \frac{y_\mu y_\nu y_\rho y_\kappa}{(2\pi)^{n/2}} \exp\left[-\frac{\mathbf{y}^T \mathbf{y}}{2}\right] \, d\mathbf{y}, \tag{2.26}$$

$$\Lambda_{\mu\mu'\nu\nu'\rho\rho'} = \int_{\mathbb{R}^n} \frac{y_\mu y_{\mu'} y_\nu y_{\nu'} y_\rho y_{\rho'}}{(2\pi)^{n/2}} \exp\left[-\frac{\mathbf{y}^T \mathbf{y}}{2}\right] \, d\mathbf{y}. \tag{2.27}$$

By Lemma 2.2, we can obtain the following lemma, which is very useful in the integrals with respect to a probability density approximated by the WKB method.

**Lemma 2.3** For sufficiently smooth functions  $f(\mathbf{x})$ ,  $g(\mathbf{x})$ , and  $h(\mathbf{x})$ ,  $\mathbf{x} \in \mathbb{R}^n$ ,

$$\begin{aligned} & \frac{\int_{\mathbb{R}^n} f(\mathbf{x})g(\mathbf{x})e^{-\frac{h(\mathbf{x})}{\varepsilon}} \, d\mathbf{x}}{\int_{\mathbb{R}^n} g(\mathbf{x})e^{-\frac{h(\mathbf{x})}{\varepsilon}} \, d\mathbf{x}} \\ & = f(\mathbf{x}^*) + \varepsilon \left[ f'_i(\mathbf{x}^*)(\log g)'_i(\mathbf{x}^*)\Xi_{ij} + \frac{f''_{ij}(\mathbf{x}^*)\Xi_{ij}}{2} - \frac{f'_i(\mathbf{x}^*)h'''_{jkl}(\mathbf{x}^*)\Xi_{i\mu}^{\frac{1}{2}}\Xi_{j\nu}^{\frac{1}{2}}\Xi_{k\rho}^{\frac{1}{2}}\Xi_{\ell\kappa}^{\frac{1}{2}}\Theta_{\mu\nu\rho\kappa}}{6} \right] \\ & + O(\varepsilon^2), \end{aligned} \tag{2.28}$$

as  $\varepsilon \rightarrow 0$ , in which Einstein's summation rule is adopted,  $\mathbf{x}^*$  is the global minimum of  $h(\mathbf{x})$ ,

Now, we are ready to relate the two kinds of asymptotic series. Recall that the two expansions are

$$\hat{p}_\varepsilon(\mathbf{z}, t) = \sum_{n=0}^\infty (\sqrt{\varepsilon})^n \hat{p}_n(\mathbf{z}, t), \tag{2.29}$$

$$p_\varepsilon(\mathbf{x}, t) = a(\varepsilon, t) \exp \left[ -\frac{1}{\varepsilon} \sum_{n=0}^\infty \phi_n(\mathbf{x}, t) \varepsilon^n \right], \tag{2.30}$$

where  $\mathbf{z} = (\mathbf{x} - \hat{\mathbf{x}})$  and  $p_\varepsilon(\mathbf{x}, t) = \hat{p}_\varepsilon(\mathbf{z}, t)/\sqrt{\varepsilon}$ . Since we will focus on the first two orders in the WKB approximation, we specifically denote that  $\phi_0 := \varphi$  and  $\phi_1 := \ln \omega$ . These two functions have particular meanings:  $\varphi$  is known as the large deviation rate function [12,54], and  $\omega$  is known as the prefactor [23,29], or the phase space factor [9], or degeneracy in the classical statistical mechanical terminology [52].

Recall that the time-dependent matrix  $\Sigma(t)$  in Lemma 2.1 is the covariance matrix of  $\mathbf{Z}(t)$  and we have checked that  $\mathbf{Z}(t)$  has the density  $\hat{p}_0(\mathbf{z}, t)$ . Therefore, for  $\Sigma(t)$ , it has the formula  $\Sigma_{ij}(t) = \int_{\mathbb{R}^n} z_i z_j \hat{p}_0(\mathbf{z}, t) d\mathbf{z}$ , for  $1 \leq i \leq n, 1 \leq j \leq n$ . Here we further define a time dependent first moment vector  $\mathbf{m}(t)$  with respect to  $\hat{p}_1(\mathbf{x}, t)$  by  $m_i(t) = \int_{\mathbb{R}^n} z_i \hat{p}_1(\mathbf{z}, t) d\mathbf{z}$ , for  $1 \leq i \leq n$ . Note that the functions  $\hat{p}_0$  and  $\hat{p}_1$  are given in the expansion (2.29). Under this framework,  $\Sigma(t)$  and  $\mathbf{m}(t)$  must satisfy the differential equations

$$\frac{d\Sigma(t)}{dt} = \mathbf{A}(\hat{\mathbf{x}}(t))\Sigma + \Sigma\mathbf{A}(\hat{\mathbf{x}}(t))^T + 2\mathbf{D}, \quad \Sigma(0) = \hat{\Sigma}_0, \tag{2.31}$$

$$\frac{d\mathbf{m}(t)}{dt} = \mathbf{A}(\hat{\mathbf{x}}(t))\mathbf{m} + \mathbf{H}(\hat{\mathbf{x}}(t))\Sigma, \quad \mathbf{m}(0) = \hat{\mathbf{m}}_0, \tag{2.32}$$

in which the initial conditions  $\hat{\Sigma}_0$  and  $\hat{\mathbf{m}}_0$  are given by the distribution of  $\mathbf{X}_\varepsilon(0)$ . For example, if the initial probability density of  $\mathbf{X}_\varepsilon(0)$  is purely Gaussian, then  $\hat{p}_1(\mathbf{z}, 0) = 0$  for all  $\mathbf{z}$  hence  $\hat{\mathbf{m}}_0 = 0$ . Furthermore,  $\mathbf{A}(\mathbf{x})$  is the Jacobian matrix of  $\mathbf{b}(\mathbf{x})$ , and  $\mathbf{H}(\mathbf{x})$  is a rank 3 tensor with  $\mathbf{H}_i(\mathbf{x})$  being the Hessian matrix of  $b_i(\mathbf{x})$ ,  $1 \leq i \leq n$ . Eq. (2.31) of  $\Sigma(t)$  is from Lemma 2.1, and Eq. (2.32) can be verified by plugging the expansion (2.29) into the the Fokker-Planck equation (2.20) and using integration by parts.

Based on the above setup,  $\Sigma(t)$ ,  $\mathbf{m}(t)$  are related to the expansion (2.29) and  $\varphi(\mathbf{x}, t)$ ,  $\omega(\mathbf{x}, t)$  are related to the expansion (2.30), then we have the following lemma for their correspondence. Recall that  $\hat{\mathbf{x}}(t)$  is the emergent deterministic trajectory of the diffusion process  $\mathbf{X}_\varepsilon(t)$  as  $\varepsilon \rightarrow 0$ .

**Lemma 2.4**  $\Sigma(t)$ ,  $\mathbf{m}(t)$ ,  $\varphi(\hat{\mathbf{x}}(t), t)$ , and  $\omega(\hat{\mathbf{x}}(t), t)$  must satisfy the equations

$$\Sigma(t) = [\nabla\nabla\varphi(\hat{\mathbf{x}}(t), t)]^{-1}, \tag{2.33}$$

$$\mathbf{m}(t) = [\nabla\nabla\varphi(\hat{\mathbf{x}}(t), t)]^{-1} \nabla \log \omega(\hat{\mathbf{x}}(t), t) - \frac{1}{6} \nabla\nabla\nabla\varphi(\hat{\mathbf{x}}(t), t)\Theta, \tag{2.34}$$

where  $(\nabla\nabla\nabla\varphi(\mathbf{x}, t)\Theta)_i = \sum_{jkl\mu\nu\rho\kappa} \varphi'''_{jkl}(\mathbf{x}, t) \mathcal{E}_{i\mu}^{\frac{1}{2}} \mathcal{E}_{j\nu}^{\frac{1}{2}} \mathcal{E}_{k\rho}^{\frac{1}{2}} \mathcal{E}_{l\kappa}^{\frac{1}{2}} \Theta_{\mu\nu\rho\kappa}$ ,  $\mathcal{E} = [\nabla\nabla\varphi(\mathbf{x}, t)]^{-1}$ , and  $\Theta$  is defined by Eq. (2.26) in Lemma 2.2.

**Proof** By the change of variable  $\mathbf{z} = (\mathbf{x} - \hat{\mathbf{x}})/\sqrt{\varepsilon}$ , we have the following two equations

$$\int (\sqrt{\varepsilon}\mathbf{z})(\sqrt{\varepsilon}\mathbf{z})^T \hat{p}_\varepsilon(\mathbf{z}, t) d\mathbf{z} = \int (\mathbf{x} - \hat{\mathbf{x}})(\mathbf{x} - \hat{\mathbf{x}})^T p_\varepsilon(\mathbf{x}, t) d\mathbf{x}, \tag{2.35}$$

$$\int (\sqrt{\varepsilon}\mathbf{z}) \hat{p}_\varepsilon(\mathbf{z}, t) d\mathbf{z} = \int (\mathbf{x} - \hat{\mathbf{x}}) p_\varepsilon(\mathbf{x}, t) d\mathbf{x}. \tag{2.36}$$

Plug the expansion (2.29) into the left side of (2.35), by Lemma 2.1, the left side of (2.35) becomes

$$\varepsilon \Sigma(t) + o(\varepsilon). \tag{2.37}$$

For a fixed  $t$ , the point  $\mathbf{x} = \hat{\mathbf{x}}(t)$  on the deterministic trajectory is the global minimum of  $\varphi(\mathbf{x}, t)$ . Therefore, to plug the expansion (2.30) into the right side of (2.35), by Lemma 2.3, we have

$$\varepsilon [\nabla \nabla \varphi(\hat{\mathbf{x}}(t), t)]^{-1} + o(\varepsilon). \tag{2.38}$$

With Eqs. (2.37) and (2.38), we thus obtain

$$\Sigma(t) = [\nabla \nabla \varphi(\hat{\mathbf{x}}(t), t)]^{-1}. \tag{2.39}$$

By a similar approach, applying Lemma 2.1 to the left side of (2.36) and Lemma 2.3 to the right side of it, we have that

$$\mathbf{m}(t) = [\nabla \nabla \varphi(\hat{\mathbf{x}}(t), t)]^{-1} \nabla \ln \omega(\hat{\mathbf{x}}(t), t) - \frac{1}{6} \nabla \nabla \nabla \varphi(\hat{\mathbf{x}}(t), t) \Theta. \tag{2.40}$$

The leading order  $\varphi(\mathbf{x}, t)$  of the time-dependent WBK ansatz (2.22) is known as a time-dependent large deviation rate function. Our work (Lemma 2.4) relates the curvature of the time-dependent large deviation rate function near its infimum with the local Gaussian fluctuations of diffusion processes. In the case of independent and identically distributed (i.i.d.) random variables sampling, the inverse of the curvature of large deviation rate function equivalent to the variance of each random variable is one of the important properties of the rate function [55,56]. Equation (2.33) in Lemma 2.4 can be regarded as an extension of this property to the case of random processes. The Freidlin-Wentzell theory [12] gives a clear definition of the large deviation rate function of random processes. From the trajectory standpoint, the action functional is defined as [12]

$$S_{0,t}(\xi) = \frac{1}{4} \int_0^t [\dot{\xi}_s - \mathbf{b}(\xi_s)] \mathbf{D}^{-1} [\dot{\xi}_s - \mathbf{b}(\xi_s)] ds, \tag{2.41}$$

where  $\xi$  is the set of all smooth paths of the process (2.1) on the interval  $[0, t]$ . Then the time-dependent large deviation rate function is the minimum of action among the set of  $\xi$

$$\varphi(\mathbf{x}, t) = \min_{\xi_0=\mathbf{x}_0, \xi_t=\mathbf{x}} S_{0,t}(\xi), \tag{2.42}$$

in which  $\mathbf{x}_0$  and  $\mathbf{x}$  are the initial and end conditions of the process, respectively.

By Eqs. (2.41) and (2.42),  $\varphi(\mathbf{x}, t)$  is no longer just a mathematical concept of the leading order term of the logarithmic asymptotics of probability densities. Borrowing the concept from classical mechanics, the integrand in the integral (2.41) is called the Lagrangian of the action and there is a corresponding Hamiltonian of the system defined by the Legendre dual of the Lagrangian [57]

$$H(\xi, \mathbf{p}) = \mathbf{b}(\xi) \cdot \mathbf{p} + \mathbf{D}\mathbf{p} \cdot \mathbf{p}. \tag{2.43}$$

Furthermore, based on the Hamiltonian given in Eq. (2.43), the large deviation rate function has to satisfy the Hamilton–Jacobi equation

$$\frac{\partial \varphi(\mathbf{x}, t)}{\partial t} = -H(\mathbf{x}, \nabla \varphi). \tag{2.44}$$

Finding solutions of the Hamilton–Jacobi equation (2.44) is still an open problem. By Lemma 2.4, with the dynamics of  $\Sigma(t)$  in Eq. (2.31) and  $\mathbf{m}(t)$  in Eq. (2.32), if the solution of the prefactor  $\omega(\mathbf{x}, t)$  is given, e.g., an uniform prefactor, then we can derive the dynamics of  $\nabla\nabla\varphi(\hat{\mathbf{x}}(t), t)$  and  $\nabla\nabla\nabla\varphi(\hat{\mathbf{x}}(t), t)$ . These results can help us approximate the solution of the HJE near its infimum: by the multivariable Taylor’s expansion, for  $\|\mathbf{x} - \hat{\mathbf{x}}\| < \delta$ , we have a *third order approximation*

$$\varphi(\mathbf{x}, t) = \frac{1}{2} [(\mathbf{x} - \hat{\mathbf{x}}(t)) \cdot \nabla]^2 \varphi(\hat{\mathbf{x}}(t), t) + \frac{1}{6} [(\mathbf{x} - \hat{\mathbf{x}}(t)) \cdot \nabla]^3 \varphi(\hat{\mathbf{x}}(t), t) + o(\delta^3), \tag{2.45}$$

and the first two terms on the right side can be numerically solved with the dynamics of  $\nabla\nabla\varphi(\hat{\mathbf{x}}(t), t)$  and  $\nabla\nabla\nabla\varphi(\hat{\mathbf{x}}(t), t)$  obtained by our theory.

To summarize the novelty and the significance of Lemma 2.4 in the following three points:

1. We show rigorously that the covariance matrix of the local time-inhomogeneous Gaussian process near a deterministic trajectory is equivalent to the inverse of the curvature of the time-dependent large deviation rate function near its infimum.
2. By having the dynamics of  $\nabla\nabla\varphi(\hat{\mathbf{x}}(t), t)$  and  $\nabla\nabla\nabla\varphi(\hat{\mathbf{x}}(t), t)$  from the lemma, we can obtain a third-order approximation of the solution of the HJE (2.44) near its infimum.
3. For analyzing a stochastic stable limit cycle in the next section, with the Lyapunov differential equation of  $\Sigma(t)$  (2.31) and the equation  $\Sigma(t) = [\nabla\nabla\varphi(\hat{\mathbf{x}}(t), t)]^{-1}$  in the lemma, we can further study the asymptotic behaviors of the time-inhomogeneous Gaussian process from a transient state to an invariant set, and relate this result to the previous works [28–31] on the curvature of the stationary large deviation rate function near a limit cycle.

### 3 Main Results of Stochastic Limit-Cycle Oscillations

In this section, we focus on nonlinear stochastic complex systems having stable limit cycles at the macroscopic scale. In chapter XIII. 7 of the textbook by van Kampen [9], the author proposed two examples of stochastic systems with stable limit cycles: one is dynamics of the Brusselator in the chemical reaction and the other one is the generalized Ginzburg–Landau equation in statistical mechanics. The model of the former started from a master equation for the Markov jump process and the later began with a SDE for the diffusion process. More examples of stochastic chemical kinetics with limit cycle oscillators were thoroughly discussed in the previous works [27,28]. In contradistinction to stochastic chemical kinetics, our work follows the idea of the second example in the van Kampen’s book along the line of the continuous representation of complex systems: we start with a randomly perturbed diffusion process satisfied the sequence of SDEs (2.1). Recall it has the form

$$d\mathbf{X}_\varepsilon(t) = \mathbf{b}(\mathbf{X}_\varepsilon)dt + [2\varepsilon\mathbf{D}]^{\frac{1}{2}}d\mathbf{B}(t). \tag{3.1}$$

In this section, we assume  $\mathbf{D}$  is positive definite in particular. Furthermore, there is an emergent deterministic dynamics as  $\varepsilon \rightarrow 0$ ,

$$d\mathbf{x}(t) = \mathbf{b}(\mathbf{x})dt, \tag{3.2}$$

and the solution  $\hat{\mathbf{x}}(t)$  of this ODE (3.2) with initial condition  $\hat{\mathbf{x}}(0)$  has a invariant solution as a stable limit cycle  $\Gamma$ . Recall that the corresponding FPE of Eq. (3.1) is

$$\frac{\partial p_\varepsilon}{\partial t} = -\nabla \cdot \mathbf{J}[p_\varepsilon], \quad \mathbf{J}[p_\varepsilon] \equiv \mathbf{b}(\mathbf{x})p_\varepsilon - \varepsilon\mathbf{D}\nabla p_\varepsilon. \tag{3.3}$$

To analyze stochastic limit cycles defined by Eqs. (3.1)–(3.3) requires asymptotic analysis involving two limits ( $\varepsilon \rightarrow 0$  and  $t \rightarrow \infty$ ). Therefore, in the first part of Sect. 3.1, we provide a brief review of the previous work on asymptotic analysis of Eqs. (3.4) and (3.5). We shall note that Eq. (3.4) is a HJE for  $\varphi(\mathbf{x}, t)$  and Eq. (3.5) is a continuity equation for  $\omega(\mathbf{x}, t)$ . This pair of equations was known in the previous work by applying WKB method to the semi-classical limit of Schrödinger equations [58,59]. In the second part of Sect. 3.1, we show a new result for stochastic limit cycles (Theorem 3.1) by asymptotic analysis of Lemmas 2.1 and 2.4 from a finite time to the infinite time limit.

### 3.1 The Process Asymptotic to a Time-Invariant Set

In Sect. 2.2, we relate the time-dependent large deviation rate function and the prefactor in the WKB method with the stochastic trajectories of randomly perturbed dynamics systems. By this correspondence, we further inspect asymptotic behaviors of the time-dependent large deviation rate function and the prefactor as time goes to infinity. By plugging the WKB ansatz (2.22) into the FPE (2.20) with equating likeorder terms, we obtain two partial differential equations of  $\varphi(\mathbf{x}, t)$  and  $\omega(\mathbf{x}, t)$  respectively,

$$\frac{\partial \varphi(\mathbf{x}, t)}{\partial t} = -\nabla \varphi(\mathbf{x}, t) \cdot \boldsymbol{\gamma}(\mathbf{x}, t) \tag{3.4}$$

$$\frac{\partial \omega(\mathbf{x}, t)}{\partial t} = -\nabla \cdot (\boldsymbol{\gamma}(\mathbf{x}, t)\omega(\mathbf{x}, t)) - \mathbf{D}\nabla \varphi(\mathbf{x}, t) \cdot \nabla \omega(\mathbf{x}, t), \tag{3.5}$$

where  $\boldsymbol{\gamma}(\mathbf{x}, t) = \mathbf{D}\nabla \varphi(\mathbf{x}, t) + \mathbf{b}(\mathbf{x})$ . We will later discuss the meaning of  $\boldsymbol{\gamma}(\mathbf{x}, t)$ , which is closely related the concept of probability flux and the Onsager’s thermodynamic force.

Let us start from analysis of the solution of  $\varphi(\mathbf{x}, t)$  in Eq. (3.4), which is a Hamilton–Jacobi equation. The HJE derived in this place by equating the leading order term in the time-dependent WKB ansatz is the same one obtained in the previous derivation (Eqs. (2.42)–(2.44)) by minimizing of action among all possible smooth paths. However, when we further analyze time-invariant solutions of this HJE, we shall notice an essential difference between the WKB-type method and the trajectory-based method, which is due to the interchange of limits of  $t$  and  $\varepsilon$ .

Following the idea of WKB anstaz (2.22) for the time-dependent probability density, if the invariant probability exists, it could be written as the asymptotic form

$$\pi_\varepsilon(\mathbf{x}) = \hat{a}(\varepsilon) \exp \left[ -\frac{1}{\varepsilon} \hat{\varphi}(\mathbf{x}) + \ln \omega(\mathbf{x}) + O(\varepsilon) \right] \tag{3.6}$$

and it is equivalent to say that

$$\hat{\varphi}(\mathbf{x}) = -\lim_{\varepsilon \rightarrow 0} \lim_{t \rightarrow \infty} \ln p_\varepsilon(\mathbf{x}, t | \mathbf{x}_0, 0), \tag{3.7}$$

where  $\hat{\varphi}$  is independent of the initial condition  $\mathbf{x}_0$  and is well-defined in the whole space  $\mathbb{R}^n$ . On the other hand, from the standpoints of trajectories, the quasipotential of the system is defined by

$$\varphi(\mathbf{x}; \mathbf{x}_f) := \inf_{t>0} \inf_{\xi} \{S_{0,t}(\xi) : \xi_0 = \mathbf{x}_f, \xi_t = \mathbf{x}\}, \tag{3.8}$$

where  $\mathbf{x}_f$  is a fixed point of the deterministic dynamics  $\hat{\mathbf{x}}' = \mathbf{b}(\hat{\mathbf{x}})$ . This definition extends the definition of minimizing the action (2.42) from a fixed  $t$  to all  $t > 0$  and it has a corresponding

probabilistic representation [57]

$$\varphi(\mathbf{x}; \mathbf{x}_f) = - \lim_{t \rightarrow \infty} \lim_{\varepsilon \rightarrow 0} \ln p_\varepsilon(\mathbf{x}, t | \mathbf{x}_f, 0). \tag{3.9}$$

The quasipotential (3.8) is one of the invariant solutions of the HJE (3.4) [12]. It may contain non-differential parts since the HJE can have a non-smooth solution after a certain finite time by studying the characteristics of it [60]. We shall emphasize that the two potentials,  $\hat{\varphi}(\mathbf{x})$  and  $\varphi(\mathbf{x}; \mathbf{x}_f)$ , have the same shape only in the domain (denoted by  $\mathcal{D}$ ) where  $\hat{\varphi}(\mathbf{x}) > \hat{\varphi}(\mathbf{x}_f)$  and  $\hat{\varphi}(\mathbf{x})$  is continuously differentiable with  $\nabla \hat{\varphi}(\mathbf{x}) \neq 0$  by the Freidlin-Wentzell uniqueness theorem of the orthogonal decomposition of the drift function  $\mathbf{b}$  [12]. Since the quasipotential  $\varphi(\mathbf{x}; \mathbf{x}_f)$  is defined strictly by the trajectories of dynamics in Eq. (3.8), to equate  $\hat{\varphi}(\mathbf{x})$  and  $\varphi(\mathbf{x}; \mathbf{x}_f)$  on  $\mathcal{D}$ , we can justify the leading order term of the WKB “ansatz” (3.6) for the invariant probability at least on  $\mathcal{D}$ . Analogously,  $\mathbf{x}_f$  can be extended from a fixed point to an invariant set, so the above statement is also true for the systems with a stable limit cycle [12]. In the following work, we will restrict our analysis of dynamics contained in  $\mathcal{D}$  and assume  $\mathcal{D}$  is compact, and use one brief notation  $\varphi(\mathbf{x})$  to represent both of the potentials.

Based on the above justification of the time-invariant WKB ansatz (2.22), we can plug it into the stationary FPE (3.3) to get the system satisfied three equations [52]

$$\mathbf{b}(\mathbf{x}) = -\mathbf{D}\nabla\varphi(\mathbf{x}) + \boldsymbol{\gamma}(\mathbf{x}), \tag{3.10}$$

$$\nabla\varphi(\mathbf{x}) \cdot \boldsymbol{\gamma}(\mathbf{x}) = 0, \tag{3.11}$$

$$\nabla \cdot (\omega(\mathbf{x})\boldsymbol{\gamma}(\mathbf{x})) = -\nabla\varphi(\mathbf{x}) \cdot \mathbf{D}\nabla\omega(\mathbf{x}). \tag{3.12}$$

Note that the vector field  $\mathbf{b}(\mathbf{x})$  represents deterministic dynamics and can be decomposed to two terms  $\nabla\varphi(\mathbf{x}) \perp \boldsymbol{\gamma}(\mathbf{x})$  which is consistent with the FW’s orthogonal decomposition [12]. In comparison to the gradient flow  $\nabla\varphi(\mathbf{x})$ , dynamics following the vector field  $\boldsymbol{\gamma}$  represents the part of circular motion of  $\mathbf{b}$ . By the previous works [26,28–32], we have the following three propositions of  $\varphi(\mathbf{x})$ :

1. We have justified the leading order term  $\varphi(\mathbf{x})$  in the stationary WKB ansatz. But it is based on the existence of the stationary probability distribution  $\pi_\varepsilon$ . The existence of  $\pi_\varepsilon$  for stochastic stable limit cycles has been proved in the work of Holland [26].
2. On a limit cycle,  $\varphi(\mathbf{x})$  and  $\nabla\varphi(\mathbf{x})$  are always zero. The landscape of  $\varphi(\mathbf{x})$  has a Mexican hat shape and the bottom of the Mexican hat ring characterizes the deterministic trajectory of the cycle [30]; And the second derivative of  $\varphi(\mathbf{x})$  tangential to the cycle is also zero, which means the Gaussian fluctuations along the direction tangential to the cycle is eventually smeared out [28,29,32].
3. Along the cycle, the matrix  $\nabla\nabla\varphi(\mathbf{x})$  is positive semi-definite [28,31]. Specifically, the smallest eigenvalue of  $\nabla\nabla\varphi(\mathbf{x})$  is zero on the cycle and the corresponding eigenvector is tangential to the cycle (by the proposition (3.1)). The rest of eigenvalues are positive, i.e., the Gaussian fluctuations perpendicular to the cycle are outward and damped out by the dissipation toward the limit cycle.

The above is the setup of  $\varphi(\mathbf{x}, t)$ . Let us continue on analysis of the solution of  $\omega(\mathbf{x}, t)$  in Eq. (3.5). It can be rearranged as

$$\frac{\partial\omega(\mathbf{x}, t)}{\partial t} + \nabla \cdot (\mathbf{b}(\mathbf{x})\omega(\mathbf{x}, t)) = -2\mathbf{D}\nabla\varphi(\mathbf{x}, t)\nabla\omega(\mathbf{x}, t) - \mathbf{D}\nabla\nabla\varphi(\mathbf{x}, t)\omega(\mathbf{x}, t), \tag{3.13}$$

where  $\mathbf{D}\nabla\nabla\varphi(\mathbf{x}, t)$  is the the Frobenius product of the matrix  $\mathbf{D}$  and the matrix  $\nabla\nabla\varphi(\mathbf{x}, t)$ . We can identify Eq. (3.13) is a *continuity equation* that describes the transport following the vector field  $\hat{\mathbf{x}}' = \mathbf{b}(\hat{\mathbf{x}})$  with a density-dependent sink (source) term on the right hand side.

Therefore, the solution of Eq. (3.13) gives us a measure of a large number particle system “without” noise by the Eulerian description of dynamics  $\dot{\hat{\mathbf{x}}} = \mathbf{b}(\hat{\mathbf{x}})$ . The original effect of noise,  $\mathbf{D}$ , appears in the sink (source) term in this continuity equation of  $\omega$ . On the other hand, if we follow the dynamics of  $\dot{\hat{\mathbf{x}}} = \mathbf{b}(\hat{\mathbf{x}})$ , we have the dynamics of  $\omega$  by the Lagrangian description

$$\begin{aligned} \frac{d\omega(\hat{\mathbf{x}}(t), t)}{dt} &= \frac{\partial\omega(\hat{\mathbf{x}}(t), t)}{\partial t} + \nabla\omega(\hat{\mathbf{x}}(t), t) \frac{d\hat{\mathbf{x}}(t)}{dt} \\ &= \nabla \cdot \mathbf{b}(\mathbf{x})\omega(\hat{\mathbf{x}}(t), t) - \mathbf{D}\nabla\nabla\varphi(\hat{\mathbf{x}}(t), t)\omega(\hat{\mathbf{x}}(t), t) \\ &= -\nabla \cdot \boldsymbol{\gamma}(\mathbf{x}(t), t)\omega(\hat{\mathbf{x}}(t), t). \end{aligned} \tag{3.14}$$

To compare Eq. (3.13) with Eq. (3.12), we have that  $\omega(\mathbf{x})$  in the WKB ansatz (3.6) is one of invariant solutions of Eq. (3.13). However, in distinction to the HJE of  $\varphi(\mathbf{x}, t)$ , the uniqueness and the smoothness of invariant solutions of the PDE (3.13) are not discussed in the present work.

Based on the above setup, we have the following theorem about the curvature of invariant large deviation rate function  $\nabla\nabla\varphi(\mathbf{x})$  and the logarithm of the prefactor  $\ln \omega(\mathbf{x})$  along the dynamics  $\mathbf{x}^*(t)$  on the limit cycle  $\Gamma$ . For this theorem, we require three assumptions of regular conditions:

1. The functions  $\varphi(\mathbf{x}, t)$  and  $\ln \omega(\mathbf{x}, t)$  are smooth enough with respect to  $\mathbf{x}$ , and the derivatives uniformly converge on  $\mathcal{D}$  as  $t \rightarrow \infty$ , in which the compact domain  $\mathcal{D}$  is defined above.
2. The drift function  $\mathbf{b}(\mathbf{x})$  and its Jacobian  $\mathbf{A}(\mathbf{x})$  are continuous on  $\mathcal{D}$ .
3. For all  $t \geq 0$ , the covariance matrix  $\boldsymbol{\Sigma}(t)$  in Lemma 2.4 is nonsingular, i.e, the Gaussian fluctuations of each direction is nonzero. Therefore, the inverse of Eq. (2.33) in Lemma 2.4 is well-defined:  $[\boldsymbol{\Sigma}(t)]^{-1} = \nabla\nabla\varphi(\hat{\mathbf{x}}(t), t)$  for all  $t \geq 0$ .

**Theorem 3.1** *Let  $\mathbf{x}^*(t)$  be a deterministic trajectory with  $\mathbf{x}^*(0) \in \Gamma$ ,  $[\boldsymbol{\Sigma}^*(t)]^{-1} := \nabla\nabla\varphi(\mathbf{x}^*(t))$ , and  $\omega^*(t) := \omega(\mathbf{x}^*(t))$ . For all  $t > 0$ ,*

$$\frac{d[\boldsymbol{\Sigma}^*(t)]^{-1}}{dt} = -[\boldsymbol{\Sigma}^*]^{-1} \mathbf{A}(\mathbf{x}^*) - \mathbf{A}(\mathbf{x}^*)^T [\boldsymbol{\Sigma}^*]^{-1} - 2[\boldsymbol{\Sigma}^*]^{-1} \mathbf{D} [\boldsymbol{\Sigma}^*]^{-1}, \tag{3.15}$$

$$\frac{d \ln \omega^*(t)}{dt} = -\nabla \cdot \mathbf{b}(\mathbf{x}^*) - \mathbf{D} [\boldsymbol{\Sigma}^*]^{-1} = -\nabla \cdot \boldsymbol{\gamma}(\mathbf{x}^*). \tag{3.16}$$

Furthermore, the smallest eigenvalues of the matrix  $[\boldsymbol{\Sigma}^*(t)]^{-1}$  is zero with the eigenvector tangential to  $\Gamma$  and the other eigenvalues are positive with the eigenvectors in the hyperplane perpendicular to  $\Gamma$ .

**Proof** In this proof, the norm  $\|\cdot\|$  represents the supremum norm. By our setup of the diffusion process (3.1), the macroscopic deterministic trajectory  $\hat{\mathbf{x}}(t)$ , with  $\hat{\mathbf{x}}(0) \in \mathcal{D}$ , converges to the stable limit cycle  $\Gamma$ , so we have

$$\lim_{t \rightarrow \infty} \min_{\mathbf{x}^* \in \Gamma} \|\hat{\mathbf{x}}(t) - \mathbf{x}^*\| = 0. \tag{3.17}$$

By the assumption (3.1) and  $\lim_{t \rightarrow \infty} \varphi(\mathbf{x}, t) = \varphi(\mathbf{x})$  in the setup, we have that

$$\lim_{t \rightarrow \infty} \nabla\nabla\varphi(\mathbf{x}, t) = \nabla\nabla\varphi(\mathbf{x}) \text{ uniformly on } \mathcal{D}. \tag{3.18}$$



Therefore, for any  $\varepsilon > 0$ , there exists  $T(\varepsilon) > 0$  such that for every  $t > T(\varepsilon)$ , there is a point  $\mathbf{x}^*(t) \in \Gamma$  with its corresponding initial point  $\mathbf{x}^*(0) \in \Gamma$  and

$$\|\nabla\nabla\varphi(\hat{\mathbf{x}}(t), t) - \nabla\nabla\varphi(\mathbf{x}^*(t))\| = O(\varepsilon), \tag{3.19}$$

in which we use the triangle inequality

$$\|\nabla\nabla\varphi(\hat{\mathbf{x}}(t), t) - \nabla\nabla\varphi(\mathbf{x}^*(t))\| \leq \|\nabla\nabla\varphi(\hat{\mathbf{x}}(t), t) - \nabla\nabla\varphi(\hat{\mathbf{x}}(t))\| + \|\nabla\nabla\varphi(\hat{\mathbf{x}}(t)) - \nabla\nabla\varphi(\mathbf{x}^*(t))\| \tag{3.20}$$

with Eq. (3.18) for the first term and Eq. (3.17) for the second term on the right side of the inequality. Apply  $[\Sigma^*(t)]^{-1} := \nabla\nabla\varphi(\mathbf{x}^*(t))$  defined in the theorem, and  $[\Sigma(t)]^{-1} = \nabla\nabla\varphi(\hat{\mathbf{x}}(t), t)$  given in Lemma 2.4 with the assumption (3.1), to Eq. (3.19), we further have that

$$\left\| [\Sigma^*(t)]^{-1} - [\Sigma(t)]^{-1} \right\| = O(\varepsilon). \tag{3.21}$$

Following the same approach for the result (3.21) with the assumption (3.1) that the derivatives of  $\varphi$  uniformly converge on  $\mathcal{D}$  and the assumption (3.1) that  $\mathbf{b}$  is continuous on  $\mathcal{D}$ , we can also show that

$$\left\| \frac{d[\Sigma^*(t)]^{-1}}{dt} - \frac{d[\Sigma(t)]^{-1}}{dt} \right\| = O(\varepsilon). \tag{3.22}$$

Furthermore, by the Lyapunov equation of  $\Sigma(t)$  obtained in Lemma 2.1, we have that

$$\frac{d[\Sigma(t)]^{-1}}{dt} = [\Sigma(t)]^{-1} \frac{d[\Sigma(t)]}{dt} [\Sigma(t)]^{-1} = -[\Sigma]^{-1} \mathbf{A}(\hat{\mathbf{x}}) - \mathbf{A}(\hat{\mathbf{x}})^T [\Sigma]^{-1} - 2[\Sigma]^{-1} \mathbf{D} [\Sigma]^{-1}. \tag{3.23}$$

By the assumption (3.1) that  $\mathbf{A}$  is continuous on  $\mathcal{D}$ , combined with Eqs. (3.21), (3.22) and (3.23), we can show that

$$\frac{d[\Sigma^*(t)]^{-1}}{dt} = -[\Sigma^*]^{-1} \mathbf{A}(\mathbf{x}^*) - \mathbf{A}(\mathbf{x}^*)^T [\Sigma^*]^{-1} - 2[\Sigma^*]^{-1} \mathbf{D} [\Sigma^*]^{-1} + O(\varepsilon), \tag{3.24}$$

in which we use the triangular inequality. Equation (3.24) is true for any  $\varepsilon > 0$  and  $t > T(\varepsilon)$ , and furthermore, the functions  $[\Sigma^*(t)]^{-1}$  and  $\mathbf{A}(\mathbf{x}^*(t))$  are periodic by the definitions, so if Eq. (3.24) holds for  $t > T(\varepsilon)$ , by the phase shift, it should hold for all  $t > 0$ . Then we can take  $\varepsilon \rightarrow 0$  to obtain

$$\frac{d[\Sigma^*(t)]^{-1}}{dt} = -[\Sigma^*]^{-1} \mathbf{A}(\mathbf{x}^*) - \mathbf{A}(\mathbf{x}^*)^T [\Sigma^*]^{-1} - 2[\Sigma^*]^{-1} \mathbf{D} [\Sigma^*]^{-1}, \tag{3.25}$$

for all  $t > 0$  with the initial condition  $[\Sigma^*(0)]^{-1} = \nabla\nabla\varphi(\mathbf{x}^*(0))$ .

Next, we need to investigate the solution  $[\Sigma^*(t)]^{-1}$  given by the ODE (3.25). Let us introduce a coordinate transformation from the Cartesian coordinate to the curvilinear coordinate around the limit cycle by the change of basis

$$\mathbf{K}(t) = \mathbf{Q}(t)^{-1} [\Sigma^*(t)]^{-1} \mathbf{Q}(t) \quad \text{with} \quad \mathbf{Q}(t) = [\mathbf{e}_1(t) \ \mathbf{e}_2(t) \ \cdots \ \mathbf{e}_n(t)], \tag{3.26}$$

in which  $\mathbf{Q}(t)$  is a time-dependent orthonormal basis. In particular,  $\mathbf{e}_1(t) = \mathbf{b}(\mathbf{x}^*) / \|\mathbf{b}(\mathbf{x}^*)\|$  is the tangential unit vector on  $\Gamma$  and the span of the rest of vectors  $\{\mathbf{e}_2(t) \ \cdots \ \mathbf{e}_n(t)\}$

represents the hyperplane perpendicular to  $\Gamma$ . For every fixed time  $t$ ,  $\mathbf{Q}(t)$  can be obtain by the *Gram–Schmidt process* and  $\mathbf{Q}(t)$  is known as the *Frenet frame*. By the proposition (3.1), since  $\nabla\nabla\varphi(\mathbf{x}^*)$  is always zero on the direction tangential to  $\Gamma$ ,  $\mathbf{e}_1(t)$  is in the nullspace of  $\nabla\nabla\varphi(\mathbf{x}^*)$  for all  $t$ . With the fact  $[\Sigma^*(t)]^{-1} := \nabla\nabla\varphi(\mathbf{x}^*(t))$  is symmetric, for all  $1 \leq i \leq n$  and  $1 \leq j \leq n$ , we have that

$$[\mathbf{K}(t)]_{i1} = \mathbf{e}_i(t)^T [\Sigma^*(t)]^{-1} \mathbf{e}_1(t) \equiv 0 \quad \text{and} \quad [\mathbf{K}(t)]_{1j} = \mathbf{e}_1(t)^T [\Sigma^*(t)]^{-1} \mathbf{e}_j(t) \equiv 0, \tag{3.27}$$

thus the matrix  $\mathbf{K}(t)$  has both zero first column and zero first row. Therefore, we can define a submatrix  $\tilde{\mathbf{K}}(t)$  by deleting the first column and the first row of  $\mathbf{K}(t)$ , equipped with Eq. (3.24), we will obtain a  $n - 1$  by  $n - 1$  system of differential equations

$$\frac{d\tilde{\mathbf{K}}(t)}{dt} = -\tilde{\mathbf{K}}(t) \left[ \tilde{\mathbf{A}}(\mathbf{x}^*) - \tilde{\mathbf{S}}(t) \right] - \left[ \tilde{\mathbf{A}}(\mathbf{x}^*) - \tilde{\mathbf{S}}(t) \right]^T \tilde{\mathbf{K}}(t) - 2\tilde{\mathbf{K}}(t)\tilde{\mathbf{D}}\tilde{\mathbf{K}}(t), \tag{3.28}$$

in which the symbol  $\sim$  on top of each matrix represents the restriction of the original matrix in the hyperplane perpendicular to  $\Gamma$  and the additional term  $\tilde{\mathbf{S}}(t)$ ,  $\mathbf{S} = \mathbf{Q}^{-1}(t)\dot{\mathbf{Q}}(t)$ , is from the coordinate transformation. We can identify that Eq. (3.28) is a *periodic Riccati differential equation*. Under the assumptions that  $\Gamma$  is a stable limit cycle and  $\tilde{\mathbf{D}}$  is definite positive (the later follows from the definite positive  $\mathbf{D}$  in the setup (3.1)), the solution of the periodic Riccati differential equation (3.28) has to be positive definite and periodic with the same period of the limit cycle. Mathematical analysis of this type of periodic Riccati differential equations can be found in the works [33–36] and a comprehensive numerical analysis with several examples was provided in the work [31].

The proof for Eq. (3.16) of  $\ln \omega^*(t)$  follows the proof for Eq. (3.15) of  $[\Sigma^*(t)]^{-1}$ . Apply the same asymptotic analysis (Eqs. (3.17)–(3.25)) to the dynamics of  $\omega(\hat{\mathbf{x}}(t), t)$  in Eq. (3.14), we can show that  $\ln \omega^*$  satisfies

$$\frac{d \ln \omega^*(t)}{dt} = \nabla \cdot \mathbf{b}(\mathbf{x}^*) - \mathbf{D}\nabla\nabla\varphi(\mathbf{x}^*) = -\nabla \cdot \boldsymbol{\gamma}(\mathbf{x}^*), \tag{3.29}$$

for all  $t > 0$  with the initial condition  $\ln \omega^*(0) = \ln \omega(\mathbf{x}^*(0))$ .

To the best of our knowledge, the asymptotic analysis from Eqs. (3.17) to (3.25) is the first work rigorously shows the Lyapunov differential equation (3.25) for the curvature of large deviation rate function near a deterministic trajectory from a transient state to an invariant set. Additionally, by analyzing the solution given by the Lyapunov differential equation with the coordinate transformation, we confirm that the solution  $[\Sigma^*(t)]^{-1}$  is consistent with the features of  $\nabla\nabla\varphi(\mathbf{x})$  on  $\Gamma$  in the proposition (3.1) from the previous work. We will apply Theorem 3.1 to further obtain three characterizations of stochastic limit cycles: (i) probability flux near the cycle (Sect. 3.2); (ii) two special features of  $\boldsymbol{\gamma}$  on the limit cycle (Sect. 3.3); (iii) a local entropy balance equation on the cycle (Sect. 3.4).

### 3.2 Probability Flux of Near a Limit Cycle

Recall that the vector field  $\boldsymbol{\gamma}$  is defined by the orthogonal decomposition of the drift  $\mathbf{b}$  in Eqs. (3.10) and (3.11), which characterizes the direction of circular motion of the deterministic dynamics  $\dot{\hat{\mathbf{x}}} = \mathbf{b}(\hat{\mathbf{x}})$ . In addition to the orthogonal decomposition of the drift  $\mathbf{b}$ ,  $\boldsymbol{\gamma}$  can be

derived from the probability flux  $\mathbf{J}[p_\varepsilon(\mathbf{x}, t)]$  defined in the FPE (3.3):

$$\boldsymbol{\gamma}_\varepsilon(\mathbf{x}, t) := \frac{\mathbf{J}[p_\varepsilon(\mathbf{x}, t)]}{p_\varepsilon(\mathbf{x}, t)} = \mathbf{b}(\mathbf{x}) - \varepsilon \mathbf{D} \nabla \ln p_\varepsilon(\mathbf{x}, t), \tag{3.30}$$

and take  $\varepsilon \rightarrow 0$  before  $t \rightarrow \infty$ , we have that

$$\lim_{t \rightarrow \infty} \lim_{\varepsilon \rightarrow 0} \boldsymbol{\gamma}_\varepsilon(\mathbf{x}, t) = \lim_{t \rightarrow \infty} \boldsymbol{\gamma}(\mathbf{x}, t) = \boldsymbol{\gamma}(\mathbf{x}), \tag{3.31}$$

in which  $\boldsymbol{\gamma}(\mathbf{x}, t)$  is the same one defined in the HJE (3.4). For the stationary probability flux  $\mathbf{J}[\pi_\varepsilon(\mathbf{x})]$ ,

$$\boldsymbol{\gamma}_\varepsilon(\mathbf{x}) := \frac{\mathbf{J}[\pi_\varepsilon(\mathbf{x})]}{\pi_\varepsilon(\mathbf{x})} = \mathbf{b}(\mathbf{x}) - \varepsilon \mathbf{D} \nabla \ln \pi_\varepsilon(\mathbf{x}), \tag{3.32}$$

which is corresponding to the reverse order of limits

$$\lim_{\varepsilon \rightarrow 0} \lim_{t \rightarrow \infty} \boldsymbol{\gamma}_\varepsilon(\mathbf{x}, t) = \lim_{\varepsilon \rightarrow 0} \boldsymbol{\gamma}_\varepsilon(\mathbf{x}) = \boldsymbol{\gamma}(\mathbf{x}). \tag{3.33}$$

Note that  $\boldsymbol{\gamma}_\varepsilon(\mathbf{x})$  has been recognized as Onsager’s thermodynamics force [61] or “local velocity” of the probability flux. Here we don’t need to worry about the orders of limits if we just focus on the domain  $\mathcal{D}$  based on our discussion in Sect. 3.1.

Recall that  $\mathbf{A}(\mathbf{x})$  is the Jacobian matrix of  $\mathbf{b}(\mathbf{x})$  and  $[\boldsymbol{\Sigma}^*]^{-1} = \nabla \nabla \varphi(\mathbf{x}^*(t))$ . By the above setup, we have the following theorem for  $\boldsymbol{\gamma}_\varepsilon(\mathbf{x})$  and the stationary probability flux  $\mathbf{J}[\pi_\varepsilon(\mathbf{x})]$  near  $\Gamma$ .

**Theorem 3.2** For  $\mathbf{x}^* \in \Gamma$  and  $\|\mathbf{x} - \mathbf{x}^*\| = O(\sqrt{\varepsilon})$ ,

$$\boldsymbol{\gamma}_\varepsilon(\mathbf{x}) := \frac{\mathbf{J}[\pi_\varepsilon(\mathbf{x})]}{\pi_\varepsilon(\mathbf{x})} = \left[ \mathbf{b}(\mathbf{x}^*) + \left( \mathbf{A}(\mathbf{x}^*) + \mathbf{D} [\boldsymbol{\Sigma}^*]^{-1} \right) (\mathbf{x} - \mathbf{x}^*) \right] + O(\varepsilon), \tag{3.34}$$

where  $\mathbf{D} [\boldsymbol{\Sigma}^*]^{-1}$  is the Frobenius product, and  $[\boldsymbol{\Sigma}^*]^{-1}$  satisfies the equation

$$\frac{d [\boldsymbol{\Sigma}^*(t)]^{-1}}{dt} = - [\boldsymbol{\Sigma}^*]^{-1} \mathbf{A}(\mathbf{x}^*) - \mathbf{A}(\mathbf{x}^*)^T [\boldsymbol{\Sigma}^*]^{-1} - 2 [\boldsymbol{\Sigma}^*]^{-1} \mathbf{D} [\boldsymbol{\Sigma}^*]^{-1}. \tag{3.35}$$

**Proof** We first approximate  $\nabla \ln \pi_\varepsilon(\mathbf{x})$  near  $\Gamma$ , in which  $\pi_\varepsilon$  has the WKB expansion in Eq. (3.6). Let  $\|\mathbf{x} - \mathbf{x}^*\| = O(\sqrt{\varepsilon})$ ,

$$\begin{aligned} \nabla \ln \pi_\varepsilon(\mathbf{x}) &= \nabla \left( \ln \omega(\mathbf{x}) - \frac{\varphi(\mathbf{x})}{\varepsilon} \right) + O(\varepsilon) \\ &= \nabla \left( \ln \omega(\mathbf{x}) - \frac{(\mathbf{x} - \mathbf{x}^*)^T \nabla \nabla \varphi(\mathbf{x}^*) (\mathbf{x} - \mathbf{x}^*)}{2\varepsilon} \right) + O(1) \\ &= - \frac{\nabla \nabla \varphi(\mathbf{x}^*) (\mathbf{x} - \mathbf{x}^*)}{\varepsilon} + O(1), \end{aligned} \tag{3.36}$$

where we use  $\varphi(\mathbf{x}^*) \equiv 0$  and  $\nabla \varphi(\mathbf{x}^*) \equiv 0$ . Apply the approximation (3.36) to Eq. (3.32), we can further approximate  $\boldsymbol{\gamma}_\varepsilon(\mathbf{x})$  around  $\Gamma$

$$\begin{aligned} \boldsymbol{\gamma}_\varepsilon(\mathbf{x}) &:= \frac{\mathbf{J}[\pi_\varepsilon(\mathbf{x})]}{\pi_\varepsilon(\mathbf{x})} = \mathbf{b}(\mathbf{x}) - \varepsilon \mathbf{D} \nabla \ln \pi_\varepsilon(\mathbf{x}) \\ &= \mathbf{b}(\mathbf{x}^*) + \left( \mathbf{A}(\mathbf{x}^*) + \mathbf{D} [\boldsymbol{\Sigma}^*]^{-1} \right) (\mathbf{x} - \mathbf{x}^*) + O(\varepsilon). \end{aligned} \tag{3.37}$$

The dynamics (3.35) is from Eq. (3.15) in Theorem 3.1.

**Remark 3.1** In the particular case of linear dynamics  $\hat{\mathbf{x}}' = \mathbf{A}\hat{\mathbf{x}}$  ( $\mathbf{A}$  is a constant  $n \times n$  matrix) with a stable fixed point  $\mathbf{x}^* = 0$ , the stationary probability flux  $\mathbf{J}$  has the formula [39]

$$\mathbf{J}[\pi_\varepsilon(\mathbf{x})] = \pi_\varepsilon^{-1}(\mathbf{x}) [(\mathbf{A} + \mathbf{D}\Sigma^{-1}) \mathbf{x}], \tag{3.38}$$

where  $\Sigma^{-1}$  satisfies  $\Sigma^{-1}\mathbf{A} + \mathbf{A}^T \Sigma^{-1} + 2\Sigma^{-1}\mathbf{D}\Sigma^{-1} = 0$ . To compare Eqs. (3.34) with (3.38), Theorem 3.2 can be regarded as a local linear approximation of the stationary probability flux near each point  $\mathbf{x}^*$  on the limit cycle. This new result extends the case from a fix point to an invariant set.

### 3.3 Two Features of $\boldsymbol{\gamma}$ on the Limit Cycle

By Theorem 3.2, we have an approximate the probability flux near  $\Gamma$ , and furthermore, by the stationary FPE (3.3), there is another important property of the probability flux

$$\nabla \cdot \mathbf{J}[\pi_\varepsilon(\mathbf{x})] = 0, \quad \text{for all } \mathbf{x} \in \mathbb{R}^n. \tag{3.39}$$

Since this property holds in the whole space, we can apply it to an arbitrary neighborhood of the limit cycle. Having this property, we will obtain two special features of  $\boldsymbol{\gamma}$  on the limit cycle in this section.

The derivation of the system of equations (3.10)–(3.12) in the work [52] was by plugging the WKB ansatz (3.6) into the stationary FPE and equating like-order terms of

$$\nabla \cdot (\pi_\varepsilon(\mathbf{x})\boldsymbol{\gamma}_\varepsilon(\mathbf{x})) = \nabla \cdot \mathbf{J}[\pi_\varepsilon(\mathbf{x})] = 0, \quad \text{for all } \mathbf{x} \in \mathbb{R}^n. \tag{3.40}$$

Applying Eq. (3.12) in the system of equations to  $\Gamma$ , we obtain the first feature of  $\boldsymbol{\gamma}$  on  $\Gamma$

$$\nabla \cdot (\omega(\mathbf{x})\boldsymbol{\gamma}(\mathbf{x})) = -\nabla\varphi(\mathbf{x}) \cdot \mathbf{D}\nabla\omega(\mathbf{x}) = 0 \quad \text{for all } \mathbf{x} \in \Gamma, \tag{3.41}$$

where we use  $\nabla\varphi(\mathbf{x}) \equiv 0$  for  $\mathbf{x} \in \Gamma$ . We can recognize that the divergence-free stationary probability flux in  $\mathbb{R}^n$  is the key to get Eq. (3.40) so that we can further obtain the divergence-free  $\omega(\mathbf{x})\boldsymbol{\gamma}(\mathbf{x})$  on the limit cycle in Eq. (3.41). With a similar approach, not only the divergence of  $\omega(\mathbf{x})\boldsymbol{\gamma}(\mathbf{x})$ , we can also obtain the second feature,  $\|\omega(\mathbf{x})\boldsymbol{\gamma}(\mathbf{x})\|$ , on the limit cycle via the following theorem:

**Theorem 3.3** *Let  $1/v(\mathbf{x})$  be the product of the nonzero eigenvalues of the matrix  $\nabla\nabla\varphi(\mathbf{x})$ . Then*

$$\sqrt{v(\mathbf{x})} \times \|\omega(\mathbf{x})\boldsymbol{\gamma}(\mathbf{x})\| \tag{3.42}$$

*is constant on the limit cycle  $\Gamma$ . Furthermore, let  $g_\varepsilon(\mathbf{x})$  be the marginal density of  $\pi_\varepsilon(\mathbf{x})$  on the limit cycle  $\Gamma$ , then for  $\mathbf{x} \in \Gamma$ ,*

$$g_\varepsilon(\mathbf{x}) = \frac{\omega(\mathbf{x})\sqrt{v(\mathbf{x})}}{\int_\Gamma \omega(\mathbf{y})\sqrt{v(\mathbf{y})}d\mathbf{y}} + O(\varepsilon), \tag{3.43}$$

*and there exists a constant  $K$  such that  $g_\varepsilon(\mathbf{x})\|\boldsymbol{\gamma}(\mathbf{x})\| = K + O(\varepsilon)$ .*

From the preceding discussion in Sect. 3.1, since  $\varphi(\mathbf{x})$  is constant on  $\Gamma$ , the eigenvector of  $\nabla\nabla\varphi(\mathbf{x})$  corresponding to the only one zero eigenvalue is tangential to  $\Gamma$ . Therefore,  $v(\mathbf{x})$  in Theorem 3.3 defined on  $\Gamma$  represents the scaled variance in the hyperplane perpendicular to  $\Gamma$  (Recall that this hyperplane is defined by the span of the vectors  $\mathbf{e}_2, \dots, \mathbf{e}_n$  in the coordinate transformation (3.26)). By Eq. (3.41), we have that  $\omega(\mathbf{x})\boldsymbol{\gamma}(\mathbf{x})$  is divergence-free on the limit cycle. By Theorem 3.3, we further have that  $\|\omega(\mathbf{x})\boldsymbol{\gamma}(\mathbf{x})\|$  is reciprocal to the

scaled standard deviation perpendicular to the limit cycle. The later was mentioned in the previous work [28]. In the present work, we provide a mathematical proof in Appendix 2. The idea of proof is by using the Gauss’s theorem for a tube around the limit cycle. Since the Gauss’s theorem can only be applied to a small but finite tube, the divergence for the Gauss’s theorem we use in the proof is  $\nabla \cdot (\boldsymbol{\gamma}_\varepsilon(\mathbf{x})\pi_\varepsilon(\mathbf{x})) = \nabla \cdot \mathbf{J}[\pi_\varepsilon(\mathbf{x})] = 0$ , which holds for an arbitrary neighborhood of  $\Gamma$ .

### 3.4 A Local Entropy Balance Equation on the Limit Cycle

On the limit cycle, we have derived the local Gaussian fluctuations of dynamics represented by the covariance  $\boldsymbol{\Sigma}^*(t)$  which follows the periodic Lyapunov equation (3.15) in Theorem 3.1. In general, the entropy of a Gaussian distribution  $p$  with a covariance  $\boldsymbol{\Sigma}$  is

$$S = - \int p(\mathbf{x}) \ln p(\mathbf{x})d\mathbf{x} = \frac{1}{2} \ln [2\pi \det(\boldsymbol{\Sigma})]. \tag{3.44}$$

Therefore, in nonlinear stochastic systems, the “local” entropy (denoted by  $S_l$ ) due to the local Gaussian fluctuations has the rate defined by

$$\frac{dS_l(t)}{dt} := \frac{1}{2} \frac{d \ln \det(\boldsymbol{\Sigma}^*(t))}{dt}. \tag{3.45}$$

By the property that the determinant of a matrix equals to the product of its eigenvalues, the local rate of entropy change (3.45) has an equivalent definition,

$$\frac{dS_l(t)}{dt} := -\frac{1}{2} \frac{d \left( \sum_{k=1}^n \ln \lambda_k^*(t) \right)}{dt}, \tag{3.46}$$

where  $\lambda_k^*(t)$ ,  $1 \leq k \leq n$  are the eigenvalues of  $[\boldsymbol{\Sigma}^*(t)]^{-1}$ . Note that  $1/v(\mathbf{x}^*(t))$  defined in Theorem 3.3 equals to the product of all nonzero eigenvalues  $\lambda_2^*(t) \cdots \lambda_n^*(t)$  of the matrix  $[\boldsymbol{\Sigma}^*(t)]^{-1}$ .

By the above setup, we have the following theorem of a *local entropy balance equation* on the limit cycle  $\Gamma$  with three equivalent expressions.

**Theorem 3.4** *For  $\mathbf{x}^*(t) \in \Gamma$ , by the definition (3.45) of the local rate of entropy change, there exists a local entropy balance equation with three equivalent expressions,*

$$\frac{dS_l(t)}{dt} = \nabla \cdot \boldsymbol{\gamma}(\mathbf{x}^*(t)), \tag{3.47}$$

$$= -\frac{d \ln \omega(\mathbf{x}^*(t))}{dt}, \tag{3.48}$$

$$= \frac{d \ln \|\boldsymbol{\gamma}(\mathbf{x}^*(t))\|}{dt} + \frac{1}{2} \frac{d \ln v(\mathbf{x}^*(t))}{dt}. \tag{3.49}$$

**Proof** By Eq. (3.45), with the dynamics of  $[\boldsymbol{\Sigma}^*(t)]^{-1}$  (3.15), we can obtain

$$\frac{dS_l(t)}{dt} = \nabla \cdot \mathbf{b}(\mathbf{x}^*) + \mathbf{D}[\boldsymbol{\Sigma}^*]^{-1} = \nabla \cdot \boldsymbol{\gamma}(\mathbf{x}^*), \tag{3.50}$$

where  $\mathbf{D}[\boldsymbol{\Sigma}^*]^{-1}$  is the Frobenius product of the matrix  $\mathbf{D}$  and the matrix  $[\boldsymbol{\Sigma}^*]^{-1}$ . Furthermore, by Eq. (3.16) for the prefactor  $\omega$ , we can link Eq. (3.50) to the dynamics of  $\omega$ ,

$$\frac{dS_l(t)}{dt} = -\frac{d \ln \omega(\mathbf{x}^*(t))}{dt}. \tag{3.51}$$

So far, we have proved the first two expressions (3.47) and (3.48). The following proof is for the third expression (3.49): By Theorem 3.1, we know the smallest eigenvalue  $\lambda_1^*(t) \equiv 0$  with its eigenvector tangential to the limit cycle. Therefore, the first term on the right side of Eq. (3.46) requires a further analysis since

$$\frac{d \ln \lambda_1^*(t)}{dt} = \frac{1}{\lambda_1^*(t)} \frac{d \lambda_1^*(t)}{dt} = \infty \times 0. \tag{3.52}$$

To find an explicit formula of (3.52), we can use

$$\frac{d \left( \boldsymbol{\gamma}^T(\mathbf{x}^*(t)) [\boldsymbol{\Sigma}^*(t)]^{-1} \boldsymbol{\gamma}(\mathbf{x}^*(t)) \right)}{dt} = 0, \quad \text{for all } t > 0, \tag{3.53}$$

since  $[\boldsymbol{\Sigma}^*(t)]^{-1} \boldsymbol{\gamma}(\mathbf{x}^*(t)) \equiv \mathbf{0}$  by Theorem 3.1. By rearranging Eq. (3.53), we find a formula of Eq. (3.52),

$$\frac{d \ln \lambda_1^*(t)}{dt} = -2 \frac{d \ln \|\boldsymbol{\gamma}(\mathbf{x}^*(t))\|}{dt}, \tag{3.54}$$

where we use that  $\lambda_1^*(t)$  is the eigenvalue of  $[\boldsymbol{\Sigma}^*(t)]^{-1}$  with respect to the eigenvector  $\boldsymbol{\gamma}(\mathbf{x}^*(t))$ . Therefore, by Eqs. (3.46) and (3.54), and with the definition of  $1/v(\mathbf{x}^*(t)) := \prod_{k=2}^n \lambda_k^*(t)$ , the local rate of entropy change on  $\Gamma$  has another expression

$$\frac{dS_I(t)}{dt} = \frac{d \ln \|\boldsymbol{\gamma}(\mathbf{x}^*(t))\|}{dt} + \frac{1}{2} \frac{d \ln v(\mathbf{x}^*(t))}{dt}. \tag{3.55}$$

Each expression has a clear physical meaning:

1. The first expression (3.47): The divergence of a vector field characterizes the volume change of the flow following this vector field. Therefore, the local entropy change can be considered as a consequence of volume-expanding (entropy-increasing) or volume-contracting (entropy-decreasing) of the circular flow  $\mathbf{x}^*(t)' = \boldsymbol{\gamma}(\mathbf{x}^*)$ . This expression of the entropy balance is corresponding to the *microscopic entropy production rate* given by a large number particle system without noise [62,63].
2. The second expression (3.48): Let us compare the rate of free energy change [29] (the free energy is defined by the relative entropy  $F(t) = \int_{\mathbb{R}^n} p(\mathbf{x}, t) \log \left( \frac{p(\mathbf{x}, t)}{\pi(\mathbf{x})} \right) d\mathbf{x}$ ) with the local rate of entropy change on the limit cycle  $\Gamma$ :

$$\frac{dF(t)}{dt} = \frac{d\varphi(\mathbf{x}^*(t))}{dt} \equiv 0, \tag{3.56}$$

$$\frac{dS_I(t)}{dt} = - \frac{d \ln \omega(\mathbf{x}^*(t))}{dt}. \tag{3.57}$$

The former follows the change of the large-deviation rate function  $\varphi(\mathbf{x})$  on the deterministic trajectory, which is always zero on the limit cycle due to constant  $\varphi(\mathbf{x}^*(t))$ ; The later follows the change of  $-\ln \omega(\mathbf{x}^*(t))$ , where the prefactor  $\omega(\mathbf{x})$  is known as “degeneracy” in the classical statistical mechanical terminology [52], which is not constant on the limit cycle in general.

3. The third expression (3.49): The local entropy balance equation can be decomposed into two parts

$$\frac{dS_I(t)}{dt} = \underbrace{\frac{d \ln \|\boldsymbol{\gamma}(\mathbf{x}^*(t))\|}{dt}}_{\text{dissipative part}} + \underbrace{\frac{1}{2} \frac{d \ln v(\mathbf{x}^*(t))}{dt}}_{\text{fluctuation part}}. \tag{3.58}$$

The first part is yielded by the change of speed on the limit cycle, which is determined from the deterministic path of the dissipative dynamics  $\mathbf{x}^*(t)' = \boldsymbol{\gamma}(\mathbf{x}^*)$ ; The second part is constituted by the change of the Gaussian fluctuations perpendicular to  $\Gamma$ . The *fluctuation-dissipation theory of nonequilibrium systems* by Keizer [17] elucidated the relation between the fluctuations of a time-inhomogeneous Gaussian process and the associated dissipative deterministic path. Following this theory, Eq. (3.58) can be regarded as a *fluctuation-dissipation decomposition* of the local entropy balance equation on the limit cycle.

**Remark 3.2** By integrating the second expression (3.48), when the system reaches its steady state, we have an equation of local entropy near the limit cycle,

$$S_I(t) = -\ln \omega(\mathbf{x}^*(t)) + C, \tag{3.59}$$

for some constant  $C$ . By the equation of entropy (3.59), we know that in the long run, the entropy of system measured near the limit cycle in the scope of the CLT should be *periodic* with the same period of the cycle. On the other hand, the global entropy in the total system has to be constant in the long run because of the existence of stationary distribution.

**Remark 3.3** By the equivalence of the expressions (3.48) and (3.49) in Theorem 3.4, we have an alternative proof for the constant  $\sqrt{v(\mathbf{x})} \times \|\omega(\mathbf{x})\boldsymbol{\gamma}(\mathbf{x})\|$  on the limit cycle  $\Gamma$  in Theorem 3.3.

### 4 Related Issue: The Scaling Hypothesis of Diffusion Processes

As the success of the well-established scaling hypothesis in the continuous-time non-negative integer valued Markov population process  $\mathbf{n}_V(t)$  (it has a law of large number as the system’s size  $V \rightarrow \infty : V^{-1}\mathbf{n}_V(t) \rightarrow \mathbf{c}(t)$ , the concentration of all the species [5]), we shall justify the origin of  $\varepsilon$  in (3.1) with physical interpretations more than just a mathematical tool.

Let us begin with a diffusion process  $\mathbf{Y}(\tau) \in \mathbb{R}^n$  satisfied the following SDE

$$d\mathbf{Y}(\tau) = \mathbf{g}(\mathbf{Y})d\tau + [2\mathbf{D}(\mathbf{Y})]^{1/2}d\mathbf{B}(\tau), \tag{4.1}$$

where  $\mathbf{g} : \mathbb{R}^n \rightarrow \mathbb{R}^n$  stands for the drift of the process,  $\mathbf{D} : \mathbb{R}^n \rightarrow \mathbb{R}^n \times \mathbb{R}^n$  is the diffusion matrix, and  $\mathbf{B}(\tau)$  is the standard  $n$ -dimensional Brownian motion. Through choosing different scales,  $\mathbf{X} = \mathbf{Y}/\alpha, t = \tau/\beta$ , the SDE (4.1) can be rescaled as

$$d\mathbf{X}(t) = \frac{\beta}{\alpha}\mathbf{g}(\alpha\mathbf{X})dt + \frac{\sqrt{\beta}}{\alpha}[2\mathbf{D}(\alpha\mathbf{X})]^{1/2}d\mathbf{B}(t). \tag{4.2}$$

We assume a *space-time structure*  $\beta = \xi(\alpha)$  by a function  $\xi : \mathbb{R} \rightarrow \mathbb{R}$ , and define a small parameter  $\varepsilon$

$$\varepsilon := \xi(\alpha)/\alpha^2 \tag{4.3}$$

with an implicit solution  $\alpha^*(\varepsilon)$  of Eq. (4.3). Under this framework, the scaled SDE (4.2) becomes a sequence of SDEs parameterized by  $\varepsilon$

$$d\mathbf{X}(t) = \mathbf{b}_\varepsilon(\mathbf{X})dt + [2\varepsilon\mathbf{D}_\varepsilon(\mathbf{X})]^{1/2}d\mathbf{B}(t), \tag{4.4}$$

where

$$\mathbf{b}_\varepsilon(\mathbf{x}) := \frac{\xi(\alpha^*(\varepsilon))}{\alpha^*(\varepsilon)}\mathbf{g}(\alpha^*(\varepsilon)\mathbf{x}) \quad \text{and} \quad \mathbf{D}_\varepsilon(\mathbf{x}) := \mathbf{D}(\alpha^*(\varepsilon)\mathbf{x}). \tag{4.5}$$

In order to observe an emergent phenomenon as  $\varepsilon \rightarrow 0$ , we require certain conditions (i)  $\lim_{\varepsilon \rightarrow 0} \mathbf{b}_\varepsilon(\mathbf{x}) = \mathbf{b}(\mathbf{x})$  exists. (ii)  $\lim_{\varepsilon \rightarrow 0} \mathbf{D}_\varepsilon(\mathbf{x}) = \mathbf{D}(\mathbf{x})$  exists. (iii) The convergence of random processes solved of the SDEs (4.4) in different modes exists [12]. Then the limit gives us an emergent deterministic dynamics  $d\mathbf{x}(t)/dt = \mathbf{b}(\mathbf{x})$  as  $\varepsilon \rightarrow 0$ .

In connection to classical overdamped mechanical motions in a viscous fluid, Eq. (4.4) is widely called a Langevin equation, and in this case the small parameter  $\varepsilon$  has been identified as related to temperature of the system as well as the “scale” under which the mechanical motion is being observed [9]. In reality, the limit of  $\varepsilon$  being zero should be interpreted as particle motions in a “continuous medium at finite temperature” rather than “temperature asymptotic to zero”. We follow this physical intuition and so called *scaling hypothesis* [64] for the origin of  $\varepsilon$ . The following discussion offers an insight into the connection between our scaling hypothesis for diffusion processes and the scaling hypothesis for statistical physics of fields.

The space-time structure defined by the function  $\xi$  is rather general. To illustrate our hypothesis, we focus on a specific space-time structure  $\beta = \alpha^k$  and thus the small parameter is defined as  $\varepsilon := \alpha^{k-2}$ . In this example, when  $k < 2$ , deterministic dynamics emerges at the macroscopic scale ( $\alpha \rightarrow \infty$ ); when  $k > 2$ , deterministic dynamics emerges at the microscopic scale ( $\alpha \rightarrow 0$ ). The choice of  $k$  depends on the property of the underlying drift function  $\mathbf{g}$  and the diffusion function  $\mathbf{D}$ . Given  $\mathbf{g}(\mathbf{x}) = c\mathbf{x}^n$ ,  $c$  is a constant, and  $\mathbf{D}$  is a constant matrix, then the sequence of SDEs (4.4) becomes

$$d\mathbf{X}(t) = \varepsilon^{\frac{k-1+n}{k-1}} c\mathbf{X}^n dt + [2\varepsilon\mathbf{D}]^{\frac{1}{2}} d\mathbf{B}(t). \tag{4.6}$$

In order to fulfill the conditions of convergence, in this example, the order of space-time structure  $k$  must be determined by the order of the underlying drift function  $n$ , i.e.,  $k = 1 - n$ . Hence, the scaled drift function becomes  $\varepsilon$ -independent while the diffusion term is asymptotic to zero as  $\varepsilon \rightarrow 0$ , which gives rise to an emergent deterministic dynamics  $d\mathbf{x}(t) = c\mathbf{x}^n dt$ . In other words, as a scientific theory, when we are able to observe deterministic dynamics  $d\mathbf{x}(t) = c\mathbf{x}^n dt$  in a “macroscopic” experiment, with an underlying stochastic dynamics having the drift function  $\mathbf{g}(\mathbf{x}) = c\mathbf{x}^n$ ,  $n > -1$ , this experiment must be running by the right space-time structure  $\beta = \alpha^{1-n}$ .

In addition to the scale of space, by the space-time relation  $\beta = \alpha^{1-n}$ , for the macroscopic emergent deterministic dynamics ( $\alpha \rightarrow \infty$ ), there is a corresponding scale of time for emergent laws: As  $-1 < n < 1$ , the deterministic dynamics emerges in a long-time limit; As  $n > 1$ , it emerges in a short-time limit. So emergent dynamics could be observed at different combinations of space-time scales, which are determined by  $n$ . This scaling exponent is given by the drift function  $\mathbf{g}$  of the underlying diffusion process. Therefore, in our scaling hypothesis for diffusion processes, experimental observation of a power law for the space-time structure is determined by the underlying physics. So we name it *scaling hypothesis*, which upholds the principle of scaling hypothesis for statistical physics of fields [64].

Here we want to introduce two types of celebrated theories which inspired our scaling hypothesis and point out what the new results we can provide beyond those theories:

1. As we mentioned in Sect. 1.2, the sequence of SDEs (4.4) has been carefully studied in the text *random perturbations of dynamical systems* by Freidlin and Wentzell [12]. Our scaling hypothesis gives  $\varepsilon$  a physical meaning which was unclear.
2. The Kurtz’s first theorem [5] showed that the ODE model is an emergent model under the infinite volume limit of the discrete Markov chain model; And the Kurtz’s second theorem [65] about the CLT for Markov chains is a generalization of a simple random walk for Donsker’s invariance principle [66]. The main distinction between our scaling



hypothesis and the scaling used in the Kurtz’s theorems is that the former is for a sequence of *scaled stochastic differential equations* but the later is for a sequence of the sum over a *scaled Markov chain*.

### 5 Discussions and Applications

In Sect. 2, we provided two preliminaries, Eq. (2.8) in Lemma 2.1 and Eq. (2.33) in Lemma 2.4: The former gives us the dynamics of the covariance of a time-inhomogeneous Gaussian process and the later characterizes the local curvature of the time-dependent large deviation function near it infimum. They both have a nice property that the existence of stationary probability is not required, so it helps us understand transient behaviors of the systems whose stationary probability may not exist. For example, if a system has unstable macroscopic deterministic dynamics, we can still compute its transient local Gaussian fluctuations and curvature of the rate function near the deterministic trajectory.

In Sect. 3, by asymptotic analysis, we characterized the dynamics near a stable limit cycle, and we found that the prefactor  $\omega$  in the WKB ansatz plays an important role, which can be seen in Theorem 3.3 and Theorem 3.4. In contradistinction to the well-established theories [20,21] of the HJE (3.4) for the large deviation rate function  $\varphi$ , to the best of our knowledge, sophisticated mathematical analysis of the PDE (3.14) for  $\omega$  might be missing and it is worthy of attention in the future. For applications, the local entropy balance equation in Theorem 3.4 can help us seek a better understanding of thermodynamic behaviors of stochastic biological oscillators, e.g., (i) mammalian cell cycles under external noises [30], (ii) a modified Morris–Lecar conductance-based model of a neuron driven by extrinsic noise [67], and (iii) Rosenzweig–MacArthur model for predator-prey interactions with the effect of stochasticity [68].

In Sect. 4, the scaling hypothesis as a scientific theory, it allows us to apply the treatment to mathematical models of the complex systems whose “noise” does not have a clear origin as classical overdamped mechanical motions in a viscous fluid described by the Langevin equation. For example, the hypothesis could be used to the models of mechanical motions in biology with noise due to coarse graining. In addition to the justification of small parameter  $\varepsilon$  itself, this hypothesis may give us a clarification of the origin of  $\varepsilon$ -dependent drift function **b**. It is known that there are two types of integrals for SDEs:

$$d\mathbf{X}(t) = \mathbf{b}_I(\mathbf{X})dt + [2\varepsilon\mathbf{D}(\mathbf{X})]^{1/2} d\mathbf{B}(t) \quad (\text{It\^o interpretation}), \tag{5.1}$$

$$d\mathbf{X}(t) = \mathbf{b}_S(\mathbf{X})dt + [2\varepsilon\mathbf{D}(\mathbf{X})]^{1/2} \circ d\mathbf{B}(t) \quad (\text{Stratonovich interpretation}). \tag{5.2}$$

The former is commonly used in mathematical analysis and financial mathematics and the later is mostly applied in physics and engineering. Note that

$$\mathbf{b}_I(\mathbf{x}) = \mathbf{b}_S(\mathbf{x}) + \varepsilon\nabla \cdot \mathbf{D}(\mathbf{x}). \tag{5.3}$$

Follow the scaling hypothesis, the existence of the extra  $\varepsilon$ -order term in Eq. (5.3) could be a corollary of the existence of higher-order terms in the underlying drift function before scaling. This hypothesis provides us a link between the two types of integral. It might help us to unravel the mystery of It\^o - Stratonovich dilemma [9,18,69] in the future.

**Acknowledgements** We thank Lowell Thompson and Ying-Jen Yang for many helpful discussions.

## Compliance with ethical standards

**Conflicts of interest** The authors declare that they have no conflict of interest.

## Appendix

### Proofs of Lemmas 2.2 and 2.3

#### Proof of Lemma 2.2

**Proof** For the special case of  $\mathbf{x} \in \mathbb{R}$ ,  $\Theta = 3$ ,  $\Lambda = 15$ , and  $\Xi = [h''(\mathbf{x}^*)]^{-1}$ . Then Eq. (2.25)

$$\eta(\mathbf{x}^*) = \frac{f''(\mathbf{x}^*)}{2h''(\mathbf{x}^*)} - \left[ \frac{f'(\mathbf{x}^*)h'''(\mathbf{x}^*)}{2[h''(\mathbf{x}^*)]^2} + \frac{f(\mathbf{x}^*)h''''(\mathbf{x}^*)}{8[h''(\mathbf{x}^*)]^2} \right] + \frac{5f(\mathbf{x}^*)[h'''(\mathbf{x}^*)]^2}{24[h''(\mathbf{x}^*)]^3}.$$

This result can be found on P. 273 of [53], Eq. (6.4.35).

For the general case,

$$\begin{aligned} & \int_{\mathbb{R}^n} f(\mathbf{x}) e^{-\frac{h(\mathbf{x})}{\varepsilon}} d\mathbf{x} \\ &= \int_{\mathbb{R}^n} \left[ f(\mathbf{x}^*) + (\mathbf{x} - \mathbf{x}^*) \cdot \nabla f(\mathbf{x}^*) + \frac{(\mathbf{x} - \mathbf{x}^*)^T \nabla \nabla f(\mathbf{x}^*) (\mathbf{x} - \mathbf{x}^*)}{2} + \dots \right] \\ & \times \exp \left[ -\frac{h(\mathbf{x}^*)}{\varepsilon} - \frac{(\mathbf{x} - \mathbf{x}^*)_i h'_{ij}(\mathbf{x}^*) (\mathbf{x} - \mathbf{x}^*)_j}{2\varepsilon} - \frac{h''_{ijk}(\mathbf{x}^*) (\mathbf{x} - \mathbf{x}^*)_i (\mathbf{x} - \mathbf{x}^*)_j (\mathbf{x} - \mathbf{x}^*)_k}{6\varepsilon} \right. \\ & \left. - \frac{h'''_{ijkl}(\mathbf{x}^*) (\mathbf{x} - \mathbf{x}^*)_i (\mathbf{x} - \mathbf{x}^*)_j (\mathbf{x} - \mathbf{x}^*)_k (\mathbf{x} - \mathbf{x}^*)_l}{24\varepsilon} + \dots \right] d\mathbf{x} \\ &= e^{-\frac{h(\mathbf{x}^*)}{\varepsilon}} \int_{\mathbb{R}^N} \left[ f(\mathbf{x}^*) + (\mathbf{x} - \mathbf{x}^*) \cdot \nabla f(\mathbf{x}^*) + \frac{(\mathbf{x} - \mathbf{x}^*)^T \nabla \nabla f(\mathbf{x}^*) (\mathbf{x} - \mathbf{x}^*)}{2} + \dots \right] \\ & \times \left[ 1 - \frac{h'_{ijk}(\mathbf{x}^*)}{6\varepsilon} (\mathbf{x} - \mathbf{x}^*)_i (\mathbf{x} - \mathbf{x}^*)_j (\mathbf{x} - \mathbf{x}^*)_k - \frac{h''_{ijkl}(\mathbf{x}^*)}{24\varepsilon} (\mathbf{x} - \mathbf{x}^*)_i (\mathbf{x} - \mathbf{x}^*)_j \right. \\ & \times (\mathbf{x} - \mathbf{x}^*)_k (\mathbf{x} - \mathbf{x}^*)_l + \left. \frac{[h'''_{ijk}(\mathbf{x}^*) (\mathbf{x} - \mathbf{x}^*)_i (\mathbf{x} - \mathbf{x}^*)_j (\mathbf{x} - \mathbf{x}^*)_k]^2}{72\varepsilon^2} + \dots \right] \\ & \times e^{-\frac{(\mathbf{x} - \mathbf{x}^*)^T \nabla \nabla h(\mathbf{x}^*) (\mathbf{x} - \mathbf{x}^*)}{2\varepsilon}} d\mathbf{x} \\ &= e^{-\frac{h(\mathbf{x}^*)}{\varepsilon}} \int_{\mathbb{R}^n} e^{-\frac{(\mathbf{x} - \mathbf{x}^*)^T \nabla \nabla h(\mathbf{x}^*) (\mathbf{x} - \mathbf{x}^*)}{2\varepsilon}} \left[ f(\mathbf{x}^*) + (\mathbf{x} - \mathbf{x}^*) \cdot \nabla f(\mathbf{x}^*) \right. \\ & + \frac{(\mathbf{x} - \mathbf{x}^*)^T \nabla \nabla f(\mathbf{x}^*) (\mathbf{x} - \mathbf{x}^*)}{2} - \frac{f(\mathbf{x}^*) h'_{ijk}(\mathbf{x}^*)}{6\varepsilon} (\mathbf{x} - \mathbf{x}^*)_i (\mathbf{x} - \mathbf{x}^*)_j (\mathbf{x} - \mathbf{x}^*)_k \\ & - \left( \frac{f'_i(\mathbf{x}^*) h'''_{jkl}(\mathbf{x}^*)}{6\varepsilon} + \frac{f(\mathbf{x}^*) h''''_{ijkl}(\mathbf{x}^*)}{24\varepsilon} \right) (\mathbf{x} - \mathbf{x}^*)_i (\mathbf{x} - \mathbf{x}^*)_j (\mathbf{x} - \mathbf{x}^*)_k (\mathbf{x} - \mathbf{x}^*)_l \\ & \left. + \frac{f(\mathbf{x}^*) [h'''_{ijk}(\mathbf{x}^*) (\mathbf{x} - \mathbf{x}^*)_i (\mathbf{x} - \mathbf{x}^*)_j (\mathbf{x} - \mathbf{x}^*)_k]^2}{72\varepsilon^2} + \dots \right] d\mathbf{x} \\ &= \sqrt{\frac{(2\pi\varepsilon)^N}{\det[\nabla \nabla h(\mathbf{x}^*)]}} e^{-\frac{h(\mathbf{x}^*)}{\varepsilon}} \left\{ f(\mathbf{x}^*) + \frac{\varepsilon f''_{ij}(\mathbf{x}^*)}{2} \Xi_{ij}(\mathbf{x}^*) \right\} \end{aligned}$$

$$\begin{aligned}
 & -\varepsilon \left[ \frac{f'_i(\mathbf{x}^*)h'''_{jkl}(\mathbf{x}^*)}{6} + \frac{f(\mathbf{x}^*)h''''_{ijk\ell}(\mathbf{x}^*)}{24} \right] \varepsilon^{\frac{1}{2}}_{i\mu} \varepsilon^{\frac{1}{2}}_{j\nu} \varepsilon^{\frac{1}{2}}_{k\rho} \varepsilon^{\frac{1}{2}}_{\ell\kappa} \Theta_{\mu\nu\rho\kappa} \\
 & + \varepsilon \left( \frac{f(\mathbf{x}^*)[h''''_{ijk}(\mathbf{x}^*)]^2}{72} \varepsilon^{-\frac{1}{2}}_{i\mu'} \varepsilon^{-\frac{1}{2}}_{j\nu'} \varepsilon^{-\frac{1}{2}}_{k\rho} \varepsilon^{-\frac{1}{2}}_{k\rho'} \varepsilon^{-\frac{1}{2}}_{\ell\kappa} \Lambda_{\mu\mu' \nu\nu' \rho\rho'} + \dots \right). \tag{6.1}
 \end{aligned}$$

A multivariate normal distribution with covariance matrix  $\Sigma$ , which is positive definite thus  $\Sigma = \Sigma^{\frac{1}{2}} \Sigma^T$  [70], has

$$\begin{aligned}
 & \frac{1}{[(2\pi\varepsilon)^N \det(\Sigma)]^{\frac{1}{2}}} \int_{\mathbb{R}^n} f''_{ij}(\mathbf{0}) x_i x_j \exp \left[ -\frac{1}{2\varepsilon} \mathbf{x}^T \Sigma^{-1} \mathbf{x} \right] d\mathbf{x} \\
 & = \frac{\varepsilon \varepsilon^{\frac{1}{2}}_{i\nu} \varepsilon^{\frac{1}{2}}_{j\mu} f''_{ij}(\mathbf{0})}{(2\pi)^{N/2}} \int_{\mathbb{R}^n} y_\nu y_\mu \exp \left[ -\frac{\mathbf{y}^T \mathbf{y}}{2} \right] d\mathbf{y} = \varepsilon f''_{ij}(\mathbf{0}) \Sigma_{ij},
 \end{aligned}$$

the Frobenius product of the Hessian matrix and covariant matrix  $\Sigma$ ,

$$\begin{aligned}
 & [(2\pi\varepsilon)^n \det(\Sigma)]^{-\frac{1}{2}} \int_{\mathbb{R}^n} f'''_{ijk}(\mathbf{0}) x_i x_j x_k \exp \left[ -\frac{1}{2\varepsilon} \mathbf{x}^T \Sigma \mathbf{x} \right] d\mathbf{x} = 0, \\
 & [(2\pi\varepsilon)^n \det(\Sigma)]^{-\frac{1}{2}} \int_{\mathbb{R}^n} f''''_{ijkl}(\mathbf{0}) x_i x_j x_k x_\ell \exp \left[ -\frac{1}{2\varepsilon} \mathbf{x}^T \Sigma \mathbf{x} \right] d\mathbf{x} \\
 & = \varepsilon^2 f''''_{ijkl}(\mathbf{0}) \varepsilon^{\frac{1}{2}}_{i\mu} \varepsilon^{\frac{1}{2}}_{j\nu} \varepsilon^{\frac{1}{2}}_{k\rho} \varepsilon^{\frac{1}{2}}_{\ell\kappa} \Theta_{\mu\nu\rho\kappa}, \\
 & [(2\pi\varepsilon)^n \det(\Sigma)]^{-\frac{1}{2}} \int_{\mathbb{R}^n} f'''_{ijk}(\mathbf{0}) x_i^2 x_j^2 x_k^2 \exp \left[ -\frac{1}{2\varepsilon} \mathbf{x}^T \Sigma \mathbf{x} \right] d\mathbf{x} \\
 & = \varepsilon^3 (2\pi)^{-\frac{n}{2}} f'''_{ijk}(\mathbf{0}) \varepsilon^{-\frac{1}{2}}_{i\mu} \varepsilon^{-\frac{1}{2}}_{i\mu'} \varepsilon^{-\frac{1}{2}}_{j\nu} \varepsilon^{-\frac{1}{2}}_{j\nu'} \varepsilon^{-\frac{1}{2}}_{k\rho} \varepsilon^{-\frac{1}{2}}_{k\rho'} \int_{\mathbb{R}^n} y_\mu y_{\mu'} y_\nu y_{\nu'} y_\rho y_{\rho'} \exp \left[ -\frac{\mathbf{y}^T \mathbf{y}}{2} \right] d\mathbf{x} \\
 & = \varepsilon^3 f''''_{ijk}(\mathbf{0}) \varepsilon^{-\frac{1}{2}}_{i\mu} \varepsilon^{-\frac{1}{2}}_{i\mu'} \varepsilon^{-\frac{1}{2}}_{j\nu} \varepsilon^{-\frac{1}{2}}_{j\nu'} \varepsilon^{-\frac{1}{2}}_{k\rho} \varepsilon^{-\frac{1}{2}}_{k\rho'} \Lambda_{\mu\mu' \nu\nu' \rho\rho'}.
 \end{aligned}$$

Applying (2.23) to both numerator and denominator of the lhs of (2.24),

$$\begin{aligned}
 & f(\mathbf{x}^*) + \varepsilon \left\{ \frac{f''_{ij}(\mathbf{x}^*) \Sigma_{ij}}{2} - \left[ \frac{f'_i(\mathbf{x}^*)h''''_{jkl}(\mathbf{x}^*)}{6} + \frac{f(\mathbf{x}^*)h''''_{ijk\ell}(\mathbf{x}^*)}{24} \right] \varepsilon^{\frac{1}{2}}_{i\mu} \varepsilon^{\frac{1}{2}}_{j\nu} \varepsilon^{\frac{1}{2}}_{k\rho} \varepsilon^{\frac{1}{2}}_{\ell\kappa} \Theta_{\mu\nu\rho\kappa} \right. \\
 & \quad \left. + \frac{f(\mathbf{x}^*)[h''''_{ijk}(\mathbf{x}^*)]^2}{72} \varepsilon^{-\frac{1}{2}}_{i\mu'} \varepsilon^{-\frac{1}{2}}_{j\nu'} \varepsilon^{-\frac{1}{2}}_{k\rho} \varepsilon^{-\frac{1}{2}}_{k\rho'} \varepsilon^{-\frac{1}{2}}_{\ell\kappa} \Lambda_{\mu\mu' \nu\nu' \rho\rho'} \right\} + O(\varepsilon^2) \\
 & \frac{1 + \varepsilon \left\{ -\left[ \frac{h''''_{ijk\ell}(\mathbf{x}^*)}{24} \right] \varepsilon^{\frac{1}{2}}_{i\mu} \varepsilon^{\frac{1}{2}}_{j\nu} \varepsilon^{\frac{1}{2}}_{k\rho} \varepsilon^{\frac{1}{2}}_{\ell\kappa} \Theta_{\mu\nu\rho\kappa} \right.}{\left. + \frac{[h''''_{ijk}(\mathbf{x}^*)]^2}{72} \varepsilon^{-\frac{1}{2}}_{i\mu'} \varepsilon^{-\frac{1}{2}}_{j\nu'} \varepsilon^{-\frac{1}{2}}_{k\rho} \varepsilon^{-\frac{1}{2}}_{k\rho'} \varepsilon^{-\frac{1}{2}}_{\ell\kappa} \Lambda_{\mu\mu' \nu\nu' \rho\rho'} \right\} + O(\varepsilon^2)}{=} \\
 & = f(\mathbf{x}^*) + \varepsilon \left[ \frac{f''_{ij}(\mathbf{x}^*) \Sigma_{ij}}{2} - \frac{f'_i(\mathbf{x}^*)h''''_{jkl}(\mathbf{x}^*) \varepsilon^{\frac{1}{2}}_{i\mu} \varepsilon^{\frac{1}{2}}_{j\nu} \varepsilon^{\frac{1}{2}}_{k\rho} \varepsilon^{\frac{1}{2}}_{\ell\kappa} \Theta_{\mu\nu\rho\kappa}}{6} \right] + O(\varepsilon^2).
 \end{aligned}$$

**Proof of Lemma 2.3**

We only provide the proof for the case  $\mathbf{x} \in \mathbb{R}^1$ , which is denoted by  $x$ . For higher dimensions, results are the same by using the notations from Lemma 2.2.

**Proof** Let the global minimum of  $[h(x) - \varepsilon \ln g(x)]$  be at  $\tilde{x}^* = x^* + \Delta x(\varepsilon)$ . Clearly,  $\Delta x \rightarrow 0$  as  $\varepsilon \rightarrow 0$ . In fact,

$$\begin{aligned} \left[ h'(x) - \varepsilon \left( \frac{g'(x)}{g(x)} \right) \right]_{x=\tilde{x}^*+\Delta x} &= 0, \\ h'(x^*) + h''(x^*)\Delta x - \varepsilon \left( \frac{g'(x^*)}{g(x^*)} \right) &= 0, \\ \Delta x &= \varepsilon \left( \frac{g'(x^*)}{g(x^*)h''(x^*)} \right) + O(\varepsilon^2). \end{aligned}$$

Now we apply the Eq. (2.24) in Lemma 2.2:

$$\begin{aligned} \frac{\int_{-\infty}^{\infty} f(x)g(x)e^{-\frac{h(x)}{\varepsilon}} dx}{\int_{-\infty}^{\infty} g(x)e^{-\frac{h(x)}{\varepsilon}} dx} &= \frac{\int_{-\infty}^{\infty} f(x)e^{-\frac{h(x)-\varepsilon \ln g(x)}{\varepsilon}} dx}{\int_{-\infty}^{\infty} e^{-\frac{h(x)-\varepsilon \ln g(x)}{\varepsilon}} dx} \\ &= f(\tilde{x}^*) + \varepsilon \left[ \frac{f''(\tilde{x}^*)}{2h''(\tilde{x}^*)} - \frac{f'(\tilde{x}^*)h'''(\tilde{x}^*)}{2[h''(\tilde{x}^*)]^2} \right] + O(\varepsilon^2), \\ &= f(x^*) + \varepsilon \left( \underbrace{\frac{f'(x^*)g'(x^*)}{g(x^*)h''(x^*)}}_{f'(x^*)\Delta x \text{ due to } g(x)} + \underbrace{\frac{f''(x^*)}{2h''(x^*)}}_{\text{due to } f(x)} - \underbrace{\frac{f'(x^*)h'''(x^*)}{2[h''(x^*)]^2}}_{\text{due to non-quadratic } h(x)} \right) + O(\varepsilon^2). \end{aligned} \tag{6.2}$$

For the terms on the order of  $\varepsilon$ , replacing  $\tilde{x}^*$  by  $x^*$  only affects the order  $\varepsilon^2$  term.

**Proof of Lemma 6.1**

Here we provide an additional Lemma, which is not used in the present work, but it is useful for related fields.

**Lemma 6.1**

$$\frac{\int_{-\infty}^{\infty} f^2(x)g(x)e^{-\frac{h(x)}{\varepsilon}} dx}{\int_{-\infty}^{\infty} g(x)e^{-\frac{h(x)}{\varepsilon}} dx} - \left[ \frac{\int_{-\infty}^{\infty} f(x)g(x)e^{-\frac{h(x)}{\varepsilon}} dx}{\int_{-\infty}^{\infty} g(x)e^{-\frac{h(x)}{\varepsilon}} dx} \right]^2 = \varepsilon \left( \frac{f'^2(x^*)}{h''(x^*)} \right) + O(\varepsilon^2). \tag{6.3}$$

**Proof**

$$\begin{aligned} \frac{\int_{-\infty}^{\infty} f^2(x)g(x)e^{-\frac{h(x)}{\varepsilon}} dx}{\int_{-\infty}^{\infty} g(x)e^{-\frac{h(x)}{\varepsilon}} dx} - \left[ \frac{\int_{-\infty}^{\infty} f(x)g(x)e^{-\frac{h(x)}{\varepsilon}} dx}{\int_{-\infty}^{\infty} g(x)e^{-\frac{h(x)}{\varepsilon}} dx} \right]^2 \\ = \varepsilon \left( \frac{2f(x^*)f'(x^*)g'(x^*)}{g(x^*)h''(x^*)} + \frac{2f'^2(x^*) + 2f(x^*)f''(x^*)}{2h''(x^*)} - \frac{2f(x^*)f'(x^*)h'''(x^*)}{2[h''(x^*)]^2} \right) \\ - 2\varepsilon f(x^*) \left( \frac{f'(x^*)g'(x^*)}{g(x^*)h''(x^*)} + \frac{f''(x^*)}{2h''(x^*)} - \frac{f'(x^*)h'''(x^*)}{2[h''(x^*)]^2} \right) + O(\varepsilon^2) \\ = \varepsilon \left( \frac{f'^2(x^*)}{h''(x^*)} \right) + O(\varepsilon^2). \end{aligned} \tag{6.4}$$

**Proof of Theorem 3.3**

**Proof** Given  $\mathbf{x}_1 \in \Gamma$ , let  $\mathcal{P}_1$  be a plane containing  $\mathbf{x}_1$  and perpendicular to the vector  $\boldsymbol{\gamma}(\mathbf{x}_1)$ . Given another point  $\mathbf{x}_2 \in \Gamma$ , let  $\mathcal{P}_2$  be a plane perpendicular to the vector  $\boldsymbol{\gamma}(\mathbf{x}_2)$ . Let  $\mathcal{S}_1 \subset \mathcal{P}_1$  and  $\mathcal{S}_2 \subset \mathcal{P}_2$  be compact sets such that

$$\max_{\mathbf{x}, \mathbf{y} \in \mathcal{S}_1} \|\mathbf{x} - \mathbf{y}\| = \max_{\mathbf{x}, \mathbf{y} \in \mathcal{S}_2} \|\mathbf{x} - \mathbf{y}\| = \delta > 0. \tag{6.5}$$

We then can define a tube  $\Phi(\delta)$  with two side boundaries  $\mathcal{S}_1$  and  $\mathcal{S}_2$  and  $\Gamma \subset \Phi(\delta)$ .

Recall that  $\boldsymbol{\gamma}_\varepsilon(\mathbf{x}) = \pi_\varepsilon(\mathbf{x})^{-1} \mathbf{J}[\pi_\varepsilon(\mathbf{x})]$ . By the stationary Fokker-Planck equation, we have that

$$\nabla \cdot (\boldsymbol{\gamma}_\varepsilon(\mathbf{x})\pi_\varepsilon(\mathbf{x})) = 0, \quad \text{for all } \mathbf{x} \in \mathbb{R}^n. \tag{6.6}$$

Furthermore, by the Gauss’s theorem,

$$\int_{\mathcal{S}} (\boldsymbol{\gamma}_\varepsilon(\mathbf{x})\pi_\varepsilon(\mathbf{x}) \cdot \mathbf{n})d\mathcal{S} = \int_{\mathcal{V}} (\nabla \cdot (\boldsymbol{\gamma}_\varepsilon(\mathbf{x})\pi_\varepsilon(\mathbf{x})))d\mathcal{V} = 0, \tag{6.7}$$

where  $\mathcal{S}$  in the surface of  $\Phi(\delta)$ ,  $\mathbf{n}$  is the outward normal vector to  $\mathcal{S}$ , and  $\mathcal{V}$  is the volume of  $\Phi(\delta)$ .

For the left hand side of Eq. (6.7), it can be written as a sum of three terms

$$\begin{aligned} \int_{\mathcal{S}} (\boldsymbol{\gamma}_\varepsilon(\mathbf{x})\pi_\varepsilon(\mathbf{x}) \cdot \mathbf{n})d\mathcal{S} &= \int_{\mathcal{S}_1} (\boldsymbol{\gamma}_\varepsilon(\mathbf{x})\pi_\varepsilon(\mathbf{x}) \cdot \mathbf{n})d\mathcal{S}_1 + \int_{\mathcal{S}_2} (\boldsymbol{\gamma}_\varepsilon(\mathbf{x})\pi_\varepsilon(\mathbf{x}) \cdot \mathbf{n})d\mathcal{S}_2 \\ &\quad + \int_{\mathcal{S}_3} (\boldsymbol{\gamma}_\varepsilon(\mathbf{x})\pi_\varepsilon(\mathbf{x}) \cdot \mathbf{n})d\mathcal{S}_3, \end{aligned} \tag{6.8}$$

in which  $\mathcal{S}_3$  is the lateral surface of  $\Phi(\delta)$  and  $\mathcal{S}_3 \cap \Gamma = \emptyset$ .

For every  $\mathbf{y} \in \Gamma$ , we denote  $\mathcal{S}_\mathbf{y} := \Phi(\delta) \cap \mathcal{P}_\mathbf{y}$ ,  $\mathcal{P}_\mathbf{y}$  is the plane perpendicular to  $\boldsymbol{\gamma}(\mathbf{y})$ . By Lemma 2.2,

$$\frac{\int_{\mathcal{S}_\mathbf{y}} f(\mathbf{x})e^{-\frac{\varphi(\mathbf{x})}{\varepsilon}} d\mathcal{S}_\mathbf{y}}{\int_{\mathcal{S}_\mathbf{y}} e^{-\frac{\varphi(\mathbf{x})}{\varepsilon}} d\mathcal{S}_\mathbf{y}} = f(\mathbf{y}) + O(\varepsilon), \tag{6.9}$$

for any continuous and bounded function  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ . Furthermore, by the definition of function  $v$  in Lemma 3.3, we can approximate the ratio of two integrals

$$\frac{\int_{\mathcal{S}_1} e^{-\frac{\varphi(\mathbf{x})}{\varepsilon}} d\mathcal{S}_1}{\int_{\mathcal{S}_1} e^{-\frac{\varphi(\mathbf{x})}{\varepsilon}} d\mathcal{S}_1} = \frac{\sqrt{2\pi\varepsilon v(\mathbf{y})}e^{-\frac{\varphi(\mathbf{y})}{\varepsilon}}}{\sqrt{2\pi\varepsilon v(\mathbf{x})}e^{-\frac{\varphi(\mathbf{x}_1)}{\varepsilon}}} + O(\varepsilon) = \frac{\sqrt{v(\mathbf{y})}}{\sqrt{v(\mathbf{x}_1)}} + O(\varepsilon), \tag{6.10}$$

in which we use Laplace’s method in the first equality and  $\varphi \equiv 0$  on  $\Gamma$  in the second equality. To choose  $f(\mathbf{x}) = \omega(\mathbf{x})$  for Eq. (6.9), combined with the result of (6.10), we can obtain

$$\frac{\int_{\mathcal{S}_1} \omega(\mathbf{x})e^{-\frac{\varphi(\mathbf{x})}{\varepsilon}} d\mathcal{S}_1}{\int_{\mathcal{S}_1} e^{-\frac{\varphi(\mathbf{x})}{\varepsilon}} d\mathcal{S}_1} = \frac{\omega(\mathbf{y})\sqrt{v(\mathbf{y})}}{\sqrt{v(\mathbf{x}_1)}} + O(\varepsilon). \tag{6.11}$$

Since the function  $e^{-\frac{\varphi(\mathbf{x})}{\varepsilon}}$  is concentrated near  $\Gamma$ , we can further approximate the normalization factor  $\int_{\mathbb{R}^n} \omega(\mathbf{x}) e^{-\frac{\varphi(\mathbf{x})}{\varepsilon}} d\mathbf{x}$  as follows

$$\frac{\int_{\mathbb{R}^n} \omega(\mathbf{x}) e^{-\frac{\varphi(\mathbf{x})}{\varepsilon}} d\mathbf{x}}{\int_{S_1} e^{-\frac{\varphi(\mathbf{x})}{\varepsilon}} dS_1} = \frac{\int_{\Gamma} \int_{S_y} \omega(\mathbf{x}) e^{-\frac{\varphi(\mathbf{x})}{\varepsilon}} dS_y dy}{\int_{S_1} e^{-\frac{\varphi(\mathbf{x})}{\varepsilon}} dS_1} + O(\varepsilon). \tag{6.12}$$

By (6.11) and (6.12), we have that

$$\frac{\int_{\mathbb{R}^n} \omega(\mathbf{x}) e^{-\frac{\varphi(\mathbf{x})}{\varepsilon}} d\mathbf{x}}{\int_{S_1} e^{-\frac{\varphi(\mathbf{x})}{\varepsilon}} dS_1} = \frac{\int_{\Gamma} \omega(\mathbf{y}) \sqrt{v(\mathbf{y})} dy}{\sqrt{v(\mathbf{x}_1)}} + O(\varepsilon). \tag{6.13}$$

By the WKB expansion of  $\pi_\varepsilon$  in Eq. (3.6), with Eq. (6.13), the first term on the right hand side of Eq. (6.8) can be written as

$$\begin{aligned} \int_{S_1} (\boldsymbol{\gamma}_\varepsilon(\mathbf{x}) \pi_\varepsilon(\mathbf{x}) \cdot \mathbf{n}) dS_1 &= \frac{\int_{S_1} (\boldsymbol{\gamma}_\varepsilon(\mathbf{x}) \cdot \mathbf{n}) \omega(\mathbf{x}) e^{-\frac{\varphi(\mathbf{x})}{\varepsilon}} dS_1}{\int_{\mathbb{R}^n} \omega(\mathbf{x}) e^{-\frac{\varphi(\mathbf{x})}{\varepsilon}} d\mathbf{x}} \\ &= \left( \frac{\sqrt{v(\mathbf{x}_1)}}{\int_{\Gamma} \omega(\mathbf{y}) \sqrt{v(\mathbf{y})} dy} \right) \left( \frac{\int_{S_1} (\boldsymbol{\gamma}_\varepsilon(\mathbf{x}) \cdot \mathbf{n}) \omega(\mathbf{x}) e^{-\frac{\varphi(\mathbf{x})}{\varepsilon}} dS_1}{\int_{S_1} e^{-\frac{\varphi(\mathbf{x})}{\varepsilon}} dS_1} \right) + O(\varepsilon). \end{aligned} \tag{6.14}$$

Note that  $\boldsymbol{\gamma}_\varepsilon(\mathbf{x}) \rightarrow \boldsymbol{\gamma}(\mathbf{x})$ . Without loss of generality, we assume that  $\boldsymbol{\gamma}(\mathbf{x}_1)$  is inflow and  $\boldsymbol{\gamma}(\mathbf{x}_2)$  is outflow of  $\Phi(\delta)$ . To choose  $f(\mathbf{x}) = (\boldsymbol{\gamma}_\varepsilon(\mathbf{x}) \cdot \mathbf{n}) \omega(\mathbf{x})$  for Eq. (6.9), combined with Eq. (6.14), we then obtain

$$\int_{S_1} (\boldsymbol{\gamma}_\varepsilon(\mathbf{x}) \pi_\varepsilon(\mathbf{x}) \cdot \mathbf{n}) dS_1 \rightarrow -C \sqrt{v(\mathbf{x}_1)} \omega(\mathbf{x}_1) \|\boldsymbol{\gamma}(\mathbf{x}_1)\| \text{ as } \varepsilon \rightarrow 0, \tag{6.15}$$

in which the constant  $C = 1 / \int_{\Gamma} \omega(\mathbf{y}) \sqrt{v(\mathbf{y})} dy$ . By the same approach, the second term on the right hand side of Eq. (6.8) has a convergence

$$\int_{S_2} (\boldsymbol{\gamma}_\varepsilon(\mathbf{x}) \pi_\varepsilon(\mathbf{x}) \cdot \mathbf{n}) dS_2 \rightarrow C \sqrt{v(\mathbf{x}_2)} \omega(\mathbf{x}_2) \|\boldsymbol{\gamma}(\mathbf{x}_2)\| \text{ as } \varepsilon \rightarrow 0. \tag{6.16}$$

Since  $S_3 \cap \Gamma = \emptyset$ , the third term

$$\int_{S_3} (\boldsymbol{\gamma}_\varepsilon(\mathbf{x}) \pi_\varepsilon(\mathbf{x}) \cdot \mathbf{n}) dS_3 \rightarrow 0 \text{ as } \varepsilon \rightarrow 0. \tag{6.17}$$

To apply the results (6.15), (6.16), and (6.17) to the equations (6.7) and (6.8), we can show that

$$\left| C \sqrt{v(\mathbf{x}_1)} \omega(\mathbf{x}_1) \|\boldsymbol{\gamma}(\mathbf{x}_1)\| - C \sqrt{v(\mathbf{x}_2)} \omega(\mathbf{x}_2) \|\boldsymbol{\gamma}(\mathbf{x}_2)\| \right| = 0. \tag{6.18}$$

Since Eq. (6.18) holds for every pair of two points on  $\Gamma$ ,  $\sqrt{v(\mathbf{x})} \omega(\mathbf{x}) \|\boldsymbol{\gamma}(\mathbf{x})\|$  is constant on  $\Gamma$ .

For  $\mathbf{x} \in \Gamma$ , the marginal density can be approximated by

$$g_\varepsilon(\mathbf{x}) = \frac{\int_{\mathbb{R}^n \setminus \Gamma} \omega(\mathbf{y}) e^{-\frac{\varphi(\mathbf{y})}{\varepsilon}} d\mathbf{y}}{\int_{\mathbb{R}^n} \omega(\mathbf{y}) e^{-\frac{\varphi(\mathbf{y})}{\varepsilon}} d\mathbf{y}} = \frac{\int_{S_x} \omega(\mathbf{y}) e^{-\frac{\varphi(\mathbf{y})}{\varepsilon}} dS_x}{\int_{\mathbb{R}^n} \omega(\mathbf{y}) e^{-\frac{\varphi(\mathbf{y})}{\varepsilon}} d\mathbf{y}} + O(\varepsilon) = \frac{\omega(\mathbf{x}) \sqrt{v(\mathbf{x})}}{\int_{\Gamma} \omega(\mathbf{y}) \sqrt{v(\mathbf{y})} dy} + O(\varepsilon), \tag{6.19}$$

which follows the steps from Eqs. (6.9) to (6.14) with choosing  $\boldsymbol{\gamma}_\varepsilon(\mathbf{x}) \cdot \mathbf{n} = 1$ . Furthermore, since  $\sqrt{v(\mathbf{x})}\omega(\mathbf{x})\|\boldsymbol{\gamma}(\mathbf{x})\|$  is constant on  $\Gamma$ , with the result (6.19), there exists a constant  $K$  such that

$$g_\varepsilon(\mathbf{x})\|\boldsymbol{\gamma}(\mathbf{x})\| = K + O(\varepsilon). \quad (6.20)$$

## References

- Hopfield, J.J.: Physics, computation, and why biology looks so different. *J. Theor. Biol.* **171**(1), 53–60 (1994)
- Peletier, M.A., van Santen, R.A., Steur, E.: *Complexity Science: An Introduction*. World Scientific, Singapore (2019)
- Anderson, P.W.: More is different. *Science* **177**(4047), 393–396 (1972)
- Berg, H.C.: *Random Walks in Biology*. Princeton University Press, Princeton (1993)
- Kurtz, T.G.: The relationship between stochastic and deterministic models for chemical reactions. *J. Chem. Phys.* **57**(7), 2976–2978 (1972)
- Kramers, H.A.: Brownian motion in a field of force and the diffusion model of chemical reactions. *Physica* **7**(4), 284–304 (1940)
- Moyal, J.E.: Stochastic processes and statistical physics. *J. R. Stat. Soc. B* **11**(2), 150–210 (1949)
- Hu, G.: Stationary solution of master equations in the large-system-size limit. *Phys. Rev. A* **36**(12), 5782 (1987)
- van Kampen, N.G.: *Stochastic Processes in Physics and Chemistry*. Elsevier, Amsterdam (1992)
- Vellela, M., Qian, H.: Stochastic dynamics and non-equilibrium thermodynamics of a bistable chemical system: the Schlögl model revisited. *J. R. Soc. Interface* **6**(39), 925–940 (2009)
- Feller, W.: The general diffusion operator and positivity preserving semi-groups in one dimension. *Ann. Math.* **60**, 417–436 (1954)
- Freidlin, M.I., Wentzell, A.D.: *Random Perturbations of Dynamical Systems*. Springer, Berlin (1998)
- Donsker, M.D., Varadhan, S.R.S.: Asymptotic evaluation of certain Markov process expectations for large time, I. *Commun. Pure Appl. Math.* **28**(1), 1–47 (1975)
- Donsker, M.D., Varadhan, S.R.S.: Asymptotic evaluation of certain Markov process expectations for large time, II. *Commun. Pure Appl. Math.* **28**(2), 279–301 (1975)
- Donsker, M.D., Varadhan, S.R.S.: Asymptotic evaluation of certain Markov process expectations for large time—III. *Commun. Pure Appl. Math.* **29**(4), 389–461 (1976)
- Donsker, M.D., Varadhan, S.R.S.: Asymptotic evaluation of certain Markov process expectations for large time. IV. *Commun. Pure Appl. Math.* **36**(2), 183–212 (1983)
- Keizer, J.: *Statistical Thermodynamics of Nonequilibrium Processes*. Springer-Verlag, Berlin (1987)
- Gardiner, C.: *Stochastic Methods*. Springer, Berlin (2009)
- Fleming, W.H.: Exit probabilities and optimal stochastic control. *Appl. Math. Optim.* **4**(1), 329–346 (1977)
- Evans, L.C., Ishii, H.: A PDE approach to some asymptotic problems concerning random differential equations with small noise intensities. *Ann. Inst. Henri Poincaré (C) Non Linear Anal.* **2**(1), 1–20 (1985)
- Feng, J., Kurtz, T.G.: *Large Deviations for Stochastic Processes*. American Mathematical Society, Providence (2006)
- Kubo, R., Matsuo, K., Kitahara, K.: Fluctuation and relaxation of macrovariables. *J. Stat. Phys.* **9**(1), 51–96 (1973)
- Graham, R., Tél, T.: On the weak-noise limit of Fokker–Planck models. *J. Stat. Phys.* **35**(5–6), 729–748 (1984)
- Donsker, M.D., Varadhan, S.R.S.: Large deviations for stationary Gaussian processes. *Commun. Math. Phys.* **97**(1–2), 187–210 (1985)
- Bryc, W., Dembo, A.: Large deviations for quadratic functionals of Gaussian processes. *J. Theor. Probab.* **10**(2), 307–332 (1997)
- Holland, C.J.: Stochastically perturbed limit cycles. *J. Appl. Probab.* **15**(2), 311–320 (1978)
- Dykman, M., Chu, X., Ross, J.: Stationary probability distribution near stable limit cycles far from Hopf bifurcation points. *Phys. Rev. E* **48**(3), 1646 (1993)
- Vance, W., Ross, J.: Fluctuations near limit cycles in chemical reaction systems. *J. Chem. Phys.* **105**(2), 479–487 (1996)
- Ge, H., Qian, H.: Landscapes of non-gradient dynamics without detailed balance: stable limit cycles and multiple attractors. *Chaos* **22**(2), 023140 (2012)

30. Li, C., Wang, J.: Landscape and flux reveal a new global view and physical quantification of mammalian cell cycle. *Proc. Natl. Acad. Sci.* **111**(39), 14130–14135 (2014)
31. Lin, L., Yu, H., Zhou, X.: Quasi-potential calculation and minimum action method for limit cycle. *J. Nonlinear Sci.* **29**(3), 961–991 (2019)
32. Kurrer, C., Schulten, K.: Effect of noise and perturbations on limit cycle systems. *Physica D* **50**(3), 311–320 (1991)
33. Bittanti, S., Bolzern, P., Colaneri, P.: Stability analysis of linear periodic systems via the Lyapunov equation. *IFAC Proc.* **17**(2), 213–216 (1984)
34. Pastor, A., Hernández, V.: Differential periodic Riccati equations: existence and uniqueness of nonnegative definite solutions. *Math. Control Signals Syst.* **6**(4), 341–362 (1993)
35. Chen, Y., Chen, S.: Existence of positive definite solution to periodic Riccati differential equation. *Proceedings of the 3rd World Congress on Intelligent Control and Automation (Cat. No. 00EX393)* **4**, 2820–2823 (2000)
36. Zhou, B., Duan, G.R.: Periodic Lyapunov equation based approaches to the stabilization of continuous-time periodic linear systems. *IEEE Trans. Autom. Control* **57**(8), 2139–2146 (2011)
37. Qian, H.: Mathematical formalism for isothermal linear irreversibility. *Proc. R. Soc. Lond. A* **457**(2011), 1645–1655 (2001)
38. Kwon, C., Ao, P., Thouless, D.J.: Structure of stochastic dynamics near fixed points. *Proc. Natl. Acad. Sci. USA* **102**(37), 13029–13033 (2005)
39. Qian, H.: Nonlinear stochastic dynamics of mesoscopic homogeneous biochemical reaction systems—an analytical theory. *Nonlinearity* **24**(6), R19 (2011)
40. Wang, J., Xu, L., Wang, E.: Potential landscape and flux framework of nonequilibrium networks: robustness, dissipation, and coherence of biochemical oscillations. *Proc. Natl. Acad. Sci. USA* **105**(34), 12271–12276 (2008)
41. Wang, J., Xu, L., Wang, E.: Robustness, dissipations and coherence of the oscillation of circadian clock: potential landscape and flux perspectives. *PMC Biophys.* **1**(1), 7 (2008)
42. Wang, J., Zhang, K., Wang, E.: Kinetic paths, time scale, and underlying landscapes: A path integral framework to study global natures of nonequilibrium systems and networks. *J. Chem. Phys.* **133**(12), 09B613 (2010)
43. Xu, L., Zhang, F., Wang, E., Wang, J.: The potential and flux landscape, lyapunov function and non-equilibrium thermodynamics for dynamic systems and networks with an application to signal-induced  $Ca^{2+}$  oscillation. *Nonlinearity* **26**(2), R69 (2012)
44. Feng, H., Han, B., Wang, J.: Landscape and global stability of nonadiabatic and adiabatic oscillations in a gene network. *Biophys. J.* **102**(5), 1001–1010 (2012)
45. Feng, H., Zhang, K., Wang, J.: Non-equilibrium transition state rate theory. *Chem. Sci.* **5**(10), 3761–3769 (2014)
46. Zhang, F., Xu, L., Zhang, K., Wang, E., Wang, J.: The potential and flux landscape theory of evolution. *J. Chem. Phys.* **137**(6), 065102 (2012)
47. Zhang, F., Xu, L., Wang, J.: The extinction differential induced virulence macroevolution. *Chem. Phys. Lett.* **599**, 38–43 (2014)
48. Xu, L., Zhang, F., Zhang, K., Wang, E., Wang, J.: The potential and flux landscape theory of ecology. *PLoS ONE* **9**(1), e86746 (2014)
49. Tong, Y.L.: *The Multivariate Normal Distribution*. Springer Science & Business Media, Berlin (2012)
50. Jiang, D.Q., Qian, M., Qian, M.P.: *Mathematical Theory of Nonequilibrium Steady States: On the Frontier of Probability and Dynamical Systems*. Springer-Verlag, Berlin Heidelberg (2004)
51. Nakanishi, H., Sakaue, T., Wakou, J.: Hamilton-Jacobi method for molecular distribution function in a chemical oscillator. *J. Chem. Phys.* **139**(21), 12B602 (2013)
52. Qian, H., Cheng, Y.C., Yang, Y.J.: Kinematic basis of emergent energetics of complex dynamics. *EPL (Europhysics Letters)* **131**(5), 50002 (2020)
53. Bender, K.M., Orszag, S.A.: *Advanced Mathematical Methods for Scientists and Engineers: I Asymptotic methods and perturbation theory*. Springer Science & Business Media, Berlin (2013)
54. Zeitouni, O., Dembo, A.: *Large Deviations Techniques and Applications*. Springer Science & Business Media, Berlin (2009)
55. Bryc, W.: A remark on the connection between the large deviation principle and the central limit theorem. *Stat. Probab. Lett.* **18**(4), 253–256 (1993)
56. Touchette, H.: The large deviation approach to statistical mechanics. *Phys. Rep.* **478**(1–3), 1–69 (2009)
57. Zhou, P., Li, T.: Construction of the landscape for multi-stable systems: potential landscape, quasi-potential, A-type integral and beyond. *J. Chem. Phys.* **144**(9), 094109 (2016)
58. Grenier, E.: Semiclassical limit of the nonlinear Schrödinger equation in small time. *Proc. Am. Math. Soc.* **126**(2), 523–530 (1998)



59. Benci, V., D'Aprile, T.: The semiclassical limit of the nonlinear Schrödinger equation in a radial potential. *J. Differ. Equ.* **184**(1), 109–138 (2002)
60. Evans, L.C.: *Partial Differential Equations Graduate Studies in Mathematics*. American Mathematical Society, Providence (1998)
61. Qian, H.: Mesoscopic nonequilibrium thermodynamics of single macromolecules and dynamic entropy-energy compensation. *Phys. Rev. E* **65**(1), 016102 (2001)
62. Dobbertin, R.: On functional relations between reduced distribution functions and entropy production by non-hamiltonian perturbations. *Phys. Scr.* **14**(3), 85 (1976)
63. Steeb, W.H.: Generalized Liouville equation, entropy, and dynamic systems containing limit cycles. *Physica A* **95**(1), 181–190 (1979)
64. Kardar, M.: *Statistical Physics of Fields*. Cambridge University Press, Cambridge (2007)
65. Kurtz, T.G.: The central limit theorem for Markov chains. *Ann. Probab.* **10**, 557–560 (1981)
66. Donsker, M.D.: *An Invariance Principle for Certain Probability Limit Theorems*. AMS, Providence (1951)
67. Bressloff, P.C., MacLaurin, J.N.: A variational method for analyzing stochastic limit cycle oscillators. *SIAM J. Appl. Dyn. Syst.* **17**(3), 2205–2233 (2018)
68. Mandler, M., Falk, J., Drossel, B.: Analysis of stochastic bifurcations with phase portraits. *PLoS ONE* **13**(4), e196256 (2018)
69. Ao, P.: Potential in stochastic differential equations: novel construction. *J. Phys. A* **37**(3), L25 (2004)
70. Bleistein, N.A., Handelsman, R.A.: *Asymptotic Expansions of Integrals*. Courier Corporation, Chelmsford (1986)

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.