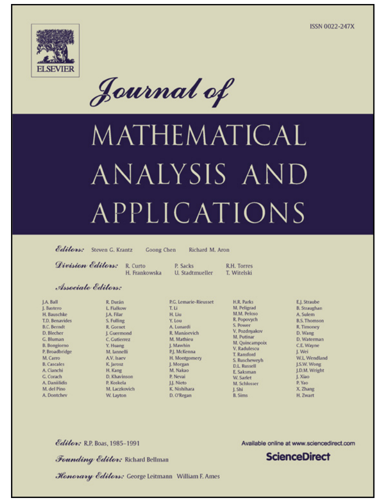
Accepted Manuscript



Identification of unstable fixed points for randomly perturbed dynamical systems with multistability

Xian Chen, Chen Jia

PII: S0022-247X(16)30384-5
DOI: http://dx.doi.org/10.1016/j.jmaa.2016.07.060
Reference: YJMAA 20622

To appear in: *Journal of Mathematical Analysis and Applications*

Received date: 26 October 2015

Please cite this article in press as: X. Chen, C. Jia, Identification of unstable fixed points for randomly perturbed dynamical systems with multistability, *J. Math. Anal. Appl.* (2016), http://dx.doi.org/10.1016/j.jmaa.2016.07.060

Identification of unstable fixed points for randomly perturbed dynamical systems with multistability

Xian Chen[1], Chen Jia[2,3,*]

[1]School of Mathematical Sciences, Peking University, Beijing 100871, P.R. China

[2]Beijing Computational Science Research Center, Beijing 100094, P.R. China

[3]Department of Mathematical Sciences, The University of Texas at Dallas, Richardson, Texas 75080, U.S.A.

[*]Author to whom correspondence should be addressed.

E-mail: chenxian@amss.ac.cn (X. Chen), jiac@utdallas.edu (C. Jia)

Abstract

Multistability, especially bistability, is one of the most important nonlinear phenomena in deterministic and stochastic dynamics. The identification of unstable fixed points for randomly perturbed dynamical systems with multistability has drawn increasing attention in recent years. In this paper, we provide a rigorous mathematical theory of the previously proposed data-driven method to identify the unstable fixed points of multistable systems. Specifically, we define a family of statistics which can be estimated by practical time-series data and prove that the local maxima of this family of statistics will converge to the unstable fixed points asymptotically. During the proof of the above result, we obtain two mathematical by-products which are interesting in their own right. We prove that the downhill timescale for randomly perturbed dynamical systems is $\log(1/\epsilon)$, different from the uphill timescale of $e^{V/\epsilon}$ for some $V > 0$ predicted by the Freidlin-Wentzell theory. Moreover, we also obtain an L^p maximum inequality for randomly perturbed dynamical systems and a class of diffusion processes.

Keywords: bistability, diffusion process, downhill timescale, maximum inequality

AMS Subject Classifications: 60H10, 60J60, 34D10, 37N25

1 Introduction

A number of deterministic and stochastic dynamical systems possess multiple stable or metastable equilibrium states. This phenomenon is widely referred to as multistability, which is one of the most important nonlinear phenomena in deterministic and stochastic dynamics [16]. Systems and devices with multistability, especially bistability, have been found or used in a wide range of scientific fields, including but not limited to mechanics, electronics, optics, thermodynamics, chemistry, biology, ecology, and meteorology. In the recent two decades, multistability has been extensively studied in biology. It has become increasingly clear that multistability is the key to understanding various basic cellular functions and the onset of complex diseases [20].

Due to the stochastic effects, a multistable system in natural sciences is usually modeled by the following randomly perturbed dynamical system:

$$dX_t^\epsilon = b(X_t^\epsilon)dt + \sqrt{\epsilon}\sigma(X_t^\epsilon)dB_t, \tag{1.1}$$

where $B = (B_t)_{t\geq 0}$ is a standard Brownian motion. For simplicity, we only consider the one-dimensional case in this paper. A multistable system can be clearly described in terms of its potential, which is defined as

$$U(x) = -\int_0^x \frac{2b(y)}{\sigma^2(y)} dy. \tag{1.2}$$

The potential of a multistable system has multiple local minima and any two adjacent local minima are separated by a local maximum (see Figure 1(a)). In the language of dynamical systems, the local minima and local maxima of the potential are the stable and unstable fixed points of the deterministic counterpart $\dot{x} = b(x)$ of the randomly perturbed dynamical system (1.1), respectively. Let s_i be all the stable fixed points and let u_i be all the unstable fixed points of the dynamical system $\dot{x} = b(x)$. Then s_i and u_i can be generally arranged as (see Figure 1(b)):

$$-\infty < s_1 < u_1 < s_2 < \cdots < s_{k-1} < u_{k-1} < s_k < \infty.$$

In recent years, the identification of unstable fixed points for multistable systems has attracted increasing attention. Recent studies on complex diseases have shown that any disease progression can be divided into a normal state, a pre-disease state, and a disease state [3]. The normal and disease states correspond to the stable fixed points of a multistable system and the pre-disease state corresponds to the unstable fixed point between them. Once the expression level of the disease-related gene in a person is close to the unstable fixed point, we have good reasons to believe that this person is in a pre-disease state and is at high risk of disease progression. This suggests that the identification of unstable fixed points for multistable systems is closely related to the early diagnosis of complex diseases.

Now that the unstable fixed points of multistable systems are of great importance, it is natural to ask whether we can detect them in an effective way by using the experimental data. Recently, several research groups have proposed different methods to solve this problem [3, 4, 10, 12]. In biological experiments, it often occurs that a large number of multistable systems with the same distribution can be observed or measured at several discrete times $t_1, t_2, \cdots, t_n$ with time interval $h(\epsilon) = t_{m+1} - t_m$. For example, the expression levels of some pivotal genes of a large amount of cells within an isogenic population can be measured at several discrete times. Intuitively, if the measurement at time t_m is around the stable fixed point s_i, then the measurement at time t_{m+1} should be also around s_i. However, if the measurement at time t_m is around the unstable fixed point u_i, then the measurement at time t_{m+1} should become rather scattered. Based on this idea, Dai et al. [4] and Jia et al. [10] define the following variance function:

$$D^\epsilon(x) = \text{Var}(X^\epsilon_{t_{m+1}} | X^\epsilon_{t_m} = x) = \mathbb{E}_x \left| X^\epsilon_{h(\epsilon)} - \mathbb{E}_x X^\epsilon_{h(\epsilon)} \right|^2. \tag{1.3}$$

According to the above intuitive ideas, the variance function should be very small around each stable fixed point and should be very large around each unstable fixed point. This suggests that the unstable fixed points should be detected by seeking the local maxima of the variance function.

Although the above method has been applied to detect the unstable fixed points of multistable biological systems based on practical time-series data, there is still a lack of a rigourous mathematical theory of this method. In this paper, we generalize the variance function proposed earlier to the case of any pth moment with $p > 0$:

$$V^{\epsilon,p}(x) = \mathbb{E}_x \left| X^\epsilon_{h(\epsilon)} - \mathbb{E}_x X^\epsilon_{h(\epsilon)} \right|^p.$$

In the case of $p = 2$, the function $V^{\epsilon,p}(x)$ is exactly the same as the variance function $D^{\epsilon}(x)$. Let m_i^{ϵ} be the maximum point of $V^{\epsilon,p}(x)$ between two adjacent stable fixed points, s_i and s_{i+1}. We prove that when the time interval $h(\epsilon)$ satisfies appropriate conditions, we have

$$\lim_{\epsilon\to 0} m_i^{\epsilon} = u_i,$$

where u_i is the unstable fixed point between s_i and s_{i+1}. This shows that when ϵ is not very large, we can indeed identity the unstable fixed points by seeking the local maxima of $V^{\epsilon,p}(x)$ for any $p > 0$. This is the first main result of this paper.

In order to prove the above result, we must solve two mathematical problems which are interesting in their own right. The first problem is the downhill timescale for randomly perturbed dynamical systems. Roughly speaking, the downhill time of a multistable system is the time for the system to arrive at the stable fixed points along the potential gradient, while the uphill time of a multistable system is the time for the system to arrive at the unstable fixed points against the potential gradient. It is a classical result of the Freidlin-Wentzell theory that the uphill timescale for multistable systems is $e^{V/\epsilon}$ for some $V > 0$ [7]. In this paper, we prove that the downhill timescale for randomly perturbed dynamical systems is $\log(1/\epsilon)$. This result, which is the second main result of this paper, is closely related to the previous studies on the escape of a randomly perturbed dynamical system from unstable fixed points or limit cycles [1, 5, 6, 13, 18].

The second problem is the L^p maximum inequality for randomly perturbed dynamical systems. In fact, the L^1 maximum inequalities for one-dimensional diffusion processes have been studied by Peskir et al. by using Lenglart's domination principle [8, 15]. However, the L^p maximum inequalities for diffusion processes with $p \neq 1$ turn out to be rather difficult, even for the Ornstein-Uhlenbeck (OU) process. As an attempt, Yan et al. [21, 22] have studied the L^p maximum inequalities for a class of diffusion processes. Although their ideas are fairly nice, their detailed proofs are questionable (see Remark 4.3). In this paper, we provide a complete proof of the L^p maximum inequality for diffusion processes with convex potential and use it to prove an L^p maximum inequality for randomly perturbed dynamical systems. This is the third main result of this paper.

The content of this paper is organized as follows. In Section 2, we develop a mathematical theory of the previously proposed method to identify the unstable fixed points of multistable systems. In Section 3, we study the downhill timescale for randomly perturbed dynamical systems. In Section 4, we study the L^p maximum inequalities for randomly perturbed dynamical systems and a class of diffusion processes.

2 Identification of unstable fixed points for randomly perturbed dynamical systems

2.1 Model

For any $\epsilon > 0$, let $X^{\epsilon} = (X_t^{\epsilon})_{t\geq 0}$ be a one-dimensional time-homogeneous diffusion process solving the stochastic differential equation

$$dX_t^{\epsilon} = b(X_t^{\epsilon})dt + \sqrt{\epsilon}\sigma(X_t^{\epsilon})dB_t, \tag{2.1}$$

where $B = (B_t)_{t\geq 0}$ is a standard Brownian motion defined on some filtered space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, P)$ satisfying the usual conditions. When $\epsilon = 0$, the stochastic differential equation (2.1) degenerates to the dynamical system $\dot{x} = b(x)$. When ϵ is small, the diffusion process X^ϵ can be viewed as a random perturbation to the dynamical system $\dot{x} = b(x)$. Therefore, X^ϵ is widely referred to as a randomly perturbed dynamical system.

In this paper, we consider the case when the dynamical system $\dot{x} = b(x)$ has a finite number of fixed points which are either stable or unstable. This implies that the number of zeros of $b(x)$ is finite and the sign of $b(x)$ changes at each zero of $b(x)$. Thus there exists a unstable fixed point between any two adjacent stable fixed points and a stable fixed point between any two adjacent unstable fixed points. We further assume that the smallest and largest fixed points are stable. Under these assumptions, all the stable and unstable fixed points can be arranged as

$$-\infty < s_1 < u_1 < s_2 < \cdots < s_{k-1} < u_{k-1} < s_k < \infty,$$

where s_i are all the stable fixed points and u_i are all the unstable fixed points of the dynamical system $\dot{x} = b(x)$ (see Figure 1(b)). When $k = 2$, the dynamical system $\dot{x} = b(x)$ has two stable fixed points and thus X^ϵ is called bistable; when $k \geq 2$, the dynamical system $\dot{x} = b(x)$ has multiple stable fixed points and thus X^ϵ is called multistable.

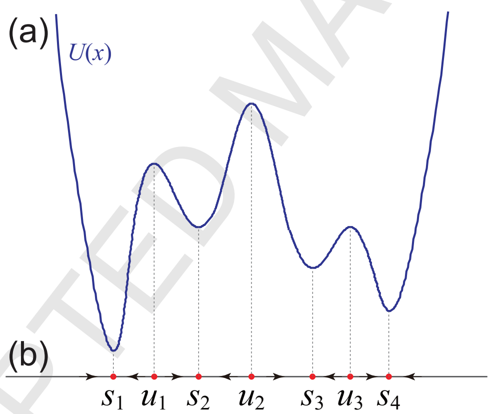


Figure 1. (a) The potential of a multiscale system. (b) The phase portrait of the deterministic counterpart of the multiscale system.

For further reference, recall that the potential of X^ϵ is defined in (1.2). Moreover, the infinitesimal generator of X^ϵ is given by

$$L^\epsilon = b(x)\frac{d}{dx} + \frac{\epsilon}{2}\sigma^2(x)\frac{d^2}{dx^2}$$

and the scale function of X^ϵ is given by

$$s^\epsilon(x) = \int_0^x e^{U(y)/\epsilon}dy.$$

It is easy to see that all the stable fixed points s_i are the local minima of the potential $U(x)$ and all the unstable fixed points u_i are the local maxima of the potential $U(x)$ (see Figure 1(a)). Intuitively, a multistable system has multiple local minima of the potential and any two adjacent local minima are separated by a local maximum, which can be visualized as a potential barrier between them.

Throughout this paper, we make the following technical assumptions.

Assumption 2.1. (i) The drift coefficient $b(x)$ and the diffusion coefficient $\sigma(x)$ are locally Lipschitz continuous, that is, for any $N > 0$, there exists a constant K_N such that for any $|x| \leq N$ and $|y| \leq N$,

$$|b(x) - b(y)| + |\sigma(x) - \sigma(y)| \leq K_N |x - y|.$$

(ii) The diffusion coefficient $\sigma(x)$ is uniformly elliptic and bounded, that is, there exist constants $\gamma_1 > 0$ and $\gamma_2 > 0$ such that $\gamma_1 \leq \sigma(x) \leq \gamma_2$ for any $x \in \mathbb{R}$.

Remark 2.2. If $b(x)$ and $\sigma(x)$ are locally Lipschitz continuous, the stochastic differential equation (2.1) has a unique strong solution X^ϵ with explosion time η^ϵ [9, Page 178]. Since the potential $U(x)$ has a finite number of local minima, it is easy to see that $U(x)$ is bounded from below. This implies that the scale function $s^\epsilon(x)$ satisfies $s^\epsilon(\infty) = \infty$ and $s^\epsilon(-\infty) = -\infty$, which guarantee that $\eta^\epsilon = \infty$, a.s. [9, Page 447, Theorem 3.1] and X^ϵ is recurrent [17, Page 311, Exercise 3.21]. Thus under Assumption 2.1, the stochastic differential equation (2.1) has a unique strong solution over the whole time axis which is also recurrent.

Assumption 2.3. The potential $U(x)$ satisfies:
(i) For any $1 \leq i \leq k$, $U(x)$ is convex in a neighborhood of s_i and $U''(s_i) > 0$;
(ii) For any $1 \leq i \leq k-1$, $U(x)$ is concave in a neighborhood of u_i and $U''(u_i) < 0$.

Under Assumption 2.1, $U(x)$ must be a C^1 function and may not be a C^2 function. Thus $U''(s_i) > 0$ in general cannot ensure $U(x)$ to be convex in a neighborhood of s_i. However, if $b(x)$ and $\sigma(x)$ are both C^1 functions, then $U(x)$ is a C^2 function and thus $U''(s_i) > 0$ is enough to ensure $U(x)$ to be convex in a neighborhood of s_i.

2.2 Main results of this section

In biological experiments, it often occurs that a large number of multistable systems with the same distribution can be measured at several discrete times $t_1, t_2, \cdots, t_n$ with time interval $h(\epsilon) = t_{m+1} - t_m$. In this paper, we define the following function $V^{\epsilon,p}(x)$ for any $p > 0$:

$$V^{\epsilon,p}(x) = \mathbb{E}_x \left| X^\epsilon_{h(\epsilon)} - \mathbb{E}_x X^\epsilon_{h(\epsilon)} \right|^p. \tag{2.2}$$

The following theorem provides a theoretical basis for the method to identify the unstable fixed points of multistable systems proposed in previous papers [4, 10].

Theorem 2.4. For any $1 \leq i \leq k-1$, let m_i^ϵ be the maximum point of $V^{\epsilon,p}(x)$ within the interval $[s_i, s_{i+1}]$. that is, $V^{\epsilon,p}(m_i^\epsilon) = \max\{V^{\epsilon,p}(x) : x \in [s_i, s_{i+1}]\}$. Assume that the time interval $h(\epsilon)$ satisfies the following two conditions:

$$\lim_{\epsilon \to 0} h(\epsilon)^{-1} \log(1/\epsilon) = 0, \tag{2.3}$$

$$\lim_{\epsilon \to 0} \epsilon \log h(\epsilon) = 0. \tag{2.4}$$

Then for any $1 \le i \le k-1$,

$$\lim_{\epsilon \to 0} m_i^\epsilon = u_i.$$

Remark 2.5. The above theorem shows that when the time interval $h(\epsilon)$ is appropriately chosen, the local maxima m_i^ϵ of $V^{\epsilon,p}(x)$ will be very close to the unstable fixed points u_i when ϵ is small. This implies that we can indeed detect the unstable fixed points of X^ϵ by seeking the local maxima of $V^{\epsilon,p}(x)$. This idea provides a data-driven method to identify the unstable fixed points of multistable systems without resorting to the details of the specific model such as the forms of the drift and diffusion coefficients.

For any $1 \le i \le k-1$, let $\tau_i^\epsilon = \inf\{t \ge 0 : X_t^\epsilon \notin (s_i, s_{i+1})\}$. Then τ_i^ϵ is the downhill time for X^ϵ when the initial value is between the stable fixed points s_i and s_{i+1}. We further define an auxiliary function $\tilde{V}^{\epsilon,p}(x)$ as

$$\tilde{V}^{\epsilon,p}(x) = \mathbb{E}_x |X_{\tau_i^\epsilon}^\epsilon - \mathbb{E}_x X_{\tau_i^\epsilon}^\epsilon|^p. \tag{2.5}$$

The following lemma gives an upper bound of the downhill timescale.

Lemma 2.6. For any $1 \le i \le k-1$,

$$\sup_{x \in [s_i, s_{i+1}]} \mathbb{E}_x \tau_i^\epsilon = O\left(\log(1/\epsilon)\right). \tag{2.6}$$

Proof. This lemma is a direct corollary of Theorem 3.1 proved in Section 3.1. □

The following lemma gives a condition under which X^ϵ cannot go too far within the time interval $h(\epsilon)$ when starting from the stable fixed point s_i.

Lemma 2.7. Assume that the time interval $h(\epsilon)$ satisfies (2.4). Then for any $1 \le i \le k$ and $p > 0$,

$$\lim_{\epsilon \to 0} \mathbb{E}_{s_i} \sup_{0 \le t \le h(\epsilon)} |X_t^\epsilon - s_i|^p = 0.$$

Proof. This lemma will be proved in Section 4.1. □

With the above two lemmas, we can prove the following approximation theorem.

Theorem 2.8. Assume that the time interval $h(\epsilon)$ satisfies (2.3) and (2.4). Then for any $1 \le i \le k-1$ and $p > 0$,

$$\lim_{\epsilon \to 0} \sup_{x \in [s_i, s_{i+1}]} |V^{\epsilon,p}(x) - \tilde{V}^{\epsilon,p}(x)| = 0. \tag{2.7}$$

Proof. For any fixed $1 \le i \le k-1$ and $x \in [s_i, s_{i+1}]$, let $\xi_{\epsilon,x} = |X_{h(\epsilon)}^\epsilon - \mathbb{E}_x X_{h(\epsilon)}^\epsilon|$ and $\eta_{\epsilon,x} = |X_{\tau_i^\epsilon}^\epsilon - \mathbb{E}_x X_{\tau_i^\epsilon}^\epsilon|$. Note that for any $p > 0$,

$$|V^{\epsilon,p}(x) - \tilde{V}^{\epsilon,p}(x)| = \left|\mathbb{E}_x \xi_{\epsilon,x}^p - \mathbb{E}_x \eta_{\epsilon,x}^p\right| \le \mathbb{E}_x \left|\xi_{\epsilon,x}^p - \eta_{\epsilon,x}^p\right|.$$

Recall that for any $a \ge 0$, $b \ge 0$, and $p > 0$,

$$(a+b)^p \le \max\{2^{p-1}, 1\}(a^p + b^p). \tag{2.8}$$

If $p < 1$, it follows from (2.8) that

$$|\xi^p_{\epsilon,x} - \eta^p_{\epsilon,x}| \le |\xi_{\epsilon,x} - \eta_{\epsilon,x}|^p. \tag{2.9}$$

If $p \ge 1$, it is readily checked from (2.8) that

$$|\xi^p_{\epsilon,x} - \eta^p_{\epsilon,x}| \le (|\xi_{\epsilon,x} - \eta_{\epsilon,x}| + \eta_{\epsilon,x})^p - \eta^p_{\epsilon,x}.$$

Denote by $\lfloor p \rfloor$ the integer satisfying $p - 1 \le \lfloor p \rfloor < p$. Direct computation shows that for any $p \ge 1$,

$$\begin{aligned}
|\xi^p_{\epsilon,x} - \eta^p_{\epsilon,x}| &\le (|\xi_{\epsilon,x} - \eta_{\epsilon,x}| + \eta_{\epsilon,x})^{\lfloor p \rfloor} (|\xi_{\epsilon,x} - \eta_{\epsilon,x}| + \eta_{\epsilon,x})^{p-\lfloor p \rfloor} - \eta^p_{\epsilon,x} \\
&\le \sum_{m=0}^{\lfloor p \rfloor} C^m_{\lfloor p \rfloor} |\xi_{\epsilon,x} - \eta_{\epsilon,x}|^m \eta^{\lfloor p \rfloor - m}_{\epsilon,x} \left(|\xi_{\epsilon,x} - \eta_{\epsilon,x}|^{p-\lfloor p \rfloor} + \eta^{p-\lfloor p \rfloor}_{\epsilon,x} \right) - \eta^p_{\epsilon,x} \\
&= \sum_{m=0}^{\lfloor p \rfloor} C^m_{\lfloor p \rfloor} |\xi_{\epsilon,x} - \eta_{\epsilon,x}|^{m+p-\lfloor p \rfloor} \eta^{\lfloor p \rfloor - m}_{\epsilon,x} + \sum_{m=1}^{\lfloor p \rfloor} C^m_{\lfloor p \rfloor} |\xi_{\epsilon,x} - \eta_{\epsilon,x}|^m \eta^{p-m}_{\epsilon,x}.
\end{aligned} \tag{2.10}$$

Since $\eta_{\epsilon,x} \le s_{i+1} - s_i$ and $p - \lfloor p \rfloor > 0$, (2.9) and (2.10) imply that for any $p > 0$, $|V^{\epsilon,p}(x) - \tilde{V}^{\epsilon,p}(x)|$ can be controlled by a finite linear combination of $\{\mathbb{E}_x |\xi_{\epsilon,x} - \eta_{\epsilon,x}|^q : q > 0\}$. For any $q > 0$,

$$\begin{aligned}
\mathbb{E}_x |\xi_{\epsilon,x} - \eta_{\epsilon,x}|^q &= \mathbb{E}_x \left| \left| X^\epsilon_{h(\epsilon)} - \mathbb{E}_x X^\epsilon_{h(\epsilon)} \right| - |X^\epsilon_{\tau^\epsilon_i} - \mathbb{E}_x X^\epsilon_{\tau^\epsilon_i}| \right|^q \\
&\le \mathbb{E}_x \left[|X^\epsilon_{h(\epsilon)} - X^\epsilon_{\tau^\epsilon_i}| + |\mathbb{E}_x X^\epsilon_{h(\epsilon)} - \mathbb{E}_x X^\epsilon_{\tau^\epsilon_i}| \right]^q \\
&\le \max\{2^{q-1}, 1\} \left[\mathbb{E}_x |X^\epsilon_{h(\epsilon)} - X^\epsilon_{\tau^\epsilon_i}|^q + \left(\mathbb{E}_x |X^\epsilon_{h(\epsilon)} - X^\epsilon_{\tau^\epsilon_i}| \right)^q \right].
\end{aligned}$$

Thus for any $p > 0$, $|V^{\epsilon,p}(x) - \tilde{V}^{\epsilon,p}(x)|$ can be controlled by a finite polynomial of $\{\mathbb{E}_x |X^\epsilon_{h(\epsilon)} - X^\epsilon_{\tau^\epsilon_i}|^q : q > 0\}$. By Chebyshev's inequality and the strong Markov property of X^ϵ at τ^ϵ_i, for any $q > 0$,

$$\begin{aligned}
\mathbb{E}_x |X^\epsilon_{h(\epsilon)} - X^\epsilon_{\tau^\epsilon_i}|^q &= \mathbb{E}_x |X^\epsilon_{h(\epsilon)} - X^\epsilon_{\tau^\epsilon_i}|^q I_{\{\tau^\epsilon_i > h(\epsilon)\}} + \mathbb{E}_x |X^\epsilon_{h(\epsilon)} - X^\epsilon_{\tau^\epsilon_i}|^q I_{\{\tau^\epsilon_i \le h(\epsilon)\}} \\
&\le (s_{i+1} - s_i)^q \mathbb{P}_x(\tau^\epsilon_i > h(\epsilon)) + \mathbb{E}_x [\sup_{0 \le t \le h(\epsilon)} |X^\epsilon_{\tau^\epsilon_i + t} - X^\epsilon_{\tau^\epsilon_i}|^q] \\
&\le (s_{i+1} - s_i)^q \frac{\mathbb{E}_x \tau^\epsilon_i}{h(\epsilon)} + \mathbb{E}_x \mathbb{E}_{X^\epsilon_{\tau^\epsilon_i}} [\sup_{0 \le t \le h(\epsilon)} |X^\epsilon_t - X^\epsilon_0|^q] \\
&\le (s_{i+1} - s_i)^q \frac{\mathbb{E}_x \tau^\epsilon_i}{h(\epsilon)} + \mathbb{E}_{s_i} [\sup_{0 \le t \le h(\epsilon)} |X^\epsilon_t - s_i|^q] + \mathbb{E}_{s_{i+1}} [\sup_{0 \le t \le h(\epsilon)} |X^\epsilon_t - s_{i+1}|^q].
\end{aligned}$$

It follows from Lemma 2.6 that

$$\lim_{\epsilon \to 0} \sup_{x \in [s_i, s_{i+1}]} \frac{\mathbb{E}_x \tau^\epsilon_i}{h(\epsilon)} = \lim_{\epsilon \to 0} h(\epsilon)^{-1} O(\log(1/\epsilon)) = 0.$$

This fact, together with Lemma 2.7, shows that

$$\lim_{\epsilon \to 0} \sup_{x \in [s_i, s_{i+1}]} \mathbb{E}_x |X^\epsilon_{h(\epsilon)} - X^\epsilon_{\tau^\epsilon_i}|^q = 0.$$

This completes the proof of this theorem. □

We are now in a position to prove Theorem 2.4.

Proof of Theorem 2.4. For any $1 \le i \le k-1$, let $\tilde{m}_i^\epsilon$ be the maximum point of $\tilde{V}^{\epsilon,p}(x)$ within the interval $[s_i, s_{i+1}]$. Direct computation shows that

$$\tilde{V}^{\epsilon,p}(x) = (s_{i+1}-s_i)^p \left[\mathbb{P}_x(X^\epsilon_{\tau_i^\epsilon} = s_i)\mathbb{P}_x(X^\epsilon_{\tau_i^\epsilon} = s_{i+1})^p + \mathbb{P}_x(X^\epsilon_{\tau_i^\epsilon} = s_i)^p\mathbb{P}_x(X^\epsilon_{\tau_i^\epsilon} = s_{i+1})\right].$$

For any $a \in [0,1]$, let $g(a) = a(1-a)^p + (1-a)a^p$. It is easy to verify that $g(a)$ is strictly increasing on $[0,r]$ and is strictly decreasing on $[1-r,1]$, where

$$r = \frac{1\wedge p}{1+p} \le \frac{1}{2}.$$

Therefore, $\tilde{m}_i^\epsilon$ should satisfy

$$\tilde{V}^{\epsilon,p}(\tilde{m}_i^\epsilon) \ge (s_{i+1}-s_i)^p g(r). \tag{2.11}$$

It is a classical result [11, Page 344, (5.61)] that

$$\mathbb{P}_x(X^\epsilon_{\tau_i^\epsilon} = s_{i+1}) = \frac{s^\epsilon(x) - s^\epsilon(s_i)}{s^\epsilon(s_{i+1}) - s^\epsilon(s_i)} = \frac{\int_{s_i}^x e^{U(y)/\epsilon}dy}{\int_{s_i}^{s_{i+1}} e^{U(y)/\epsilon}dy}. \tag{2.12}$$

Note that $U_i''(u_i) < 0$ and u_i is the unique maximum point of $U(x)$ between s_i and s_{i+1}. By Laplace's method [19, Page 277], it is easy to check that

$$\lim_{\epsilon\to 0}\int_{s_i}^{s_{i+1}} e^{U(y)/\epsilon}dy \Big/ \left[e^{U(u_i)/\epsilon}\sqrt{-2\pi\epsilon/U''(u_i)}\right] = 1.$$

This shows that when ϵ is sufficiently small,

$$\int_{s_i}^{s_{i+1}} e^{U(y)/\epsilon}dy \ge \frac{1}{2}e^{U(u_i)/\epsilon}\sqrt{-2\pi\epsilon/U''(u_i)}.$$

Thus for any $\eta > 0$, whenever $x < u_i - \eta$, we have

$$\mathbb{P}_x(X^\epsilon_{\tau_i^\epsilon} = s_{i+1}) = \frac{\int_{s_i}^x e^{U(y)/\epsilon}dy}{\int_{s_i}^{s_{i+1}} e^{U(y)/\epsilon}dy} \le \frac{2(u_i-\eta-s_i)e^{U(u_i-\eta)/\epsilon}}{e^{U(u_i)/\epsilon}\sqrt{-2\pi\epsilon/U''(u_i)}} \to 0, \quad \text{as } \epsilon \to 0.$$

Thus when ϵ is sufficiently small, $\mathbb{P}_x(X^\epsilon_{\tau_i^\epsilon} = s_{i+1}) < r/2$ for any $x < u_i - \eta$. Similarly, when ϵ is sufficiently small, $\mathbb{P}_x(X^\epsilon_{\tau_i^\epsilon} = s_{i+1}) > 1 - r/2$ for any $x > u_i + \eta$. This shows that for any $|x - u_i| > \eta$,

$$\tilde{V}^{\epsilon,p}(x) \le g(r/2)(s_{i+1}-s_i)^p.$$

By Theorem 2.8, when ϵ is sufficiently small, we have

$$\sup_{x\in[s_i,s_{i+1}]} |V^{\epsilon,p}(x) - \tilde{V}^{\epsilon,p}(x)| < \frac{1}{2}(s_{i+1}-s_i)^p(g(r) - g(r/2)). \tag{2.13}$$

Thus for any $|x - u_i| > \eta$,

$$V^{\epsilon,p}(x) \le \tilde{V}^{\epsilon,p}(x) + \sup_{x\in[s_i,s_{i+1}]} |V^{\epsilon,p}(x) - \tilde{V}^{\epsilon,p}(x)| < \frac{1}{2}(s_{i+1}-s_i)^p(g(r) + g(r/2)).$$

On the other hand, it follows from (2.11) and (2.13) that

$$V^{\epsilon,p}(\tilde{m}_i^\epsilon) \ge \tilde{V}^{\epsilon,p}(\tilde{m}_i^\epsilon) - \sup_{x\in[s_i,s_{i+1}]} |V^{\epsilon,p}(x) - \tilde{V}^{\epsilon,p}(x)| > \frac{1}{2}(s_{i+1}-s_i)^p\left(g(r) + g(r/2)\right).$$

This implies that $V^{\epsilon,p}(\tilde{m}_i^\epsilon) > V^{\epsilon,p}(x)$ whenever $|x - u_i| > \eta$. Thus when ϵ is sufficiently small, we have $|m_i^\epsilon - u_i| \le \eta$. By the arbitrariness of η, we obtain the desired result. $\square$

We have seen that the keys to the proof of Theorem 2.4 are Lemmas 2.6 and 2.7, which are interesting in their own right. Lemma 2.6 gives an upper bound of the downhill timescale for randomly perturbed dynamical systems and Lemma 2.7 is closely related to the L^p maximum inequality for randomly perturbed dynamical systems. These two lemmas will be proved in the following two sections.

2.3 A remark on the choice of the time interval $h(\epsilon)$

There is an important question that has not been answered satisfactorily in the previous studies: how should we choose the time interval $h(\epsilon)$ between two successive measurements? In this paper, we prove that if the interval $h(\epsilon)$ satisfies the two conditions given in (2.3) and (2.4), then we can detect the unstable fixed points of multistable systems by seeking the local maxima of $V^{\epsilon,p}(x)$. However, this raises the following question: what are the intuitive implications of these two conditions?

Intuitively, if we hope $V^{\epsilon,p}(x)$ to be large around the unstable fixed points u_i, then the interval $h(\epsilon)$ should be long enough to make the system arrive at the stable fixed points s_i or s_{i+1} within the interval $h(\epsilon)$. This shows that $h(\epsilon)$ should have a larger timescale than the downhill time τ_i^ϵ. This intuitive idea, together with Lemma 2.6, shows that $h(\epsilon)$ should be chosen to satisfy the condition (2.3).

On the other hand, if we hope $V^{\epsilon,p}(x)$ to be small around the stable fixed points s_i, then the interval $h(\epsilon)$ should be short enough to ensure the system not to make transitions between different stable fixed points within the interval $h(\epsilon)$. According to the Freidlin-Wentzell theory [7], X^ϵ will transition between different stable fixed points at the timescale of $e^{V/\epsilon}$ for some $V > 0$. This suggests that $h(\epsilon)$ should be chosen to satisfy the condition (2.4).

The above discussion shows that the two conditions given in (2.3) and (2.4) coincide with our intuitive ideas perfectly. Based on numerical simulations, Jia et al. [10] suggested that the timescale of $h(\epsilon)$ may be chosen as $1/\epsilon$. It is obvious that any polynomial timescale of $1/\epsilon$ must satisfy these two conditions and thus the above mentioned method to identify the unstable fixed points of multistable systems can be successfully applied.

3 Downhill timescale for randomly perturbed dynamical systems

3.1 Main results of this section

In this section, we shall study the downhill timescale for randomly perturbed dynamical systems. The following theorem gives an upper bound of the downhill time for X^ϵ when the initial value is between two adjacent stable fixed points.

Theorem 3.1. For any $1 \leq i \leq k-1$, let $\tau_i^\epsilon = \inf\{t \geq 0 : X_t^\epsilon \notin (s_i, s_{i+1})\}$. Then for any $1 \leq i \leq k-1$,

$$\limsup_{\epsilon \to 0} \frac{\sup_{x\in[s_i,s_{i+1}]} \mathbb{E}_x \tau_i^\epsilon}{\log(1/\epsilon)} \leq \frac{2}{\gamma_1^2}\left[\frac{1}{U''(s_i)} - \frac{2}{U''(u_i)} + \frac{1}{U''(s_{i+1})}\right],$$

where γ_1 is the constant described in Assumption 2.1.

Proof. This theorem will be proved in Section 3.2. □

The reader may ask what is the downhill timescale for X^ϵ when the initial value is outside the interval $[s_1, s_k]$. To answer this question, we make the following assumption, which is equivalent to saying that X^ϵ has an invariant distribution when ϵ is sufficiently small.

Assumption 3.2. There exists a constant $\gamma_3 > 0$ such that

$$\int_{\mathbb{R}} e^{-\gamma_3 U(y)} dy < \infty.$$

Let $\tau_x^\epsilon = \inf\{t \geq 0 : X_t^\epsilon = x\}$ be the first passage time of x by X^ϵ. Then $\tau_{s_1}^\epsilon$ is the downhill time for X^ϵ when the initial value $x \leq s_1$ and $\tau_{s_k}^\epsilon$ is the downhill time for X^ϵ when the initial value $x \geq s_k$. The following theorem gives an upper bound of the downhill timescale for X^ϵ when the initial value is outside the interval $[s_1, s_k]$.

Theorem 3.3. Under Assumptions 3.2, for any $x \geq s_k$,

$$\limsup_{\epsilon \to 0} \frac{\mathbb{E}_x \tau_{s_k}^\epsilon}{\log(1/\epsilon)} \leq \frac{1}{\gamma_1^2 U''(s_k)}, \tag{3.1}$$

and for any $x \leq s_1$,

$$\limsup_{\epsilon \to 0} \frac{\mathbb{E}_x \tau_{s_1}^\epsilon}{\log(1/\epsilon)} \leq \frac{1}{\gamma_1^2 U''(s_1)}. \tag{3.2}$$

Proof. This theorem will be proved in Section 3.3. □

Remark 3.4. The above two theorems show that wherever the initial value is located, the downhill timescale for X^ϵ has an upper bound of $\log(1/\epsilon)$. In fact, it can be proved that under stronger conditions, $\log(1/\epsilon)$ is also a lower bound. This shows that in general, $\log(1/\epsilon)$ is the correct downhill timescale for randomly perturbed dynamical systems.

3.2 Proof of Theorem 3.1

In order to prove Theorem 3.1, we need a lemma.

Lemma 3.5. For any $1 \leq i \leq k-1$,

$$\sup_{x \in [s_i, s_{i+1}]} \mathbb{E}_x \tau_i^\epsilon \leq \frac{4}{\epsilon \gamma_1^2} \left[\int_{s_i}^{u_i} e^{U(y)/\epsilon} dy \int_y^{u_i} e^{-U(z)/\epsilon} dz + \int_{u_i}^{s_{i+1}} e^{U(y)/\epsilon} dy \int_{u_i}^{y} e^{-U(z)/\epsilon} dz \right].$$

Proof. For any $1 \leq i \leq k-1$ and $x \in [s_i, s_{i+1}]$, it is a classical result [17, Page 305, Corollary 3.8] that

$$\begin{aligned} \mathbb{E}_x \tau_i^\epsilon = & \left[\frac{1}{\epsilon} \int_{s_i}^{x} e^{U(y)/\epsilon} dy \int_y^x \frac{2}{\sigma^2(z)} e^{-U(z)/\epsilon} dz \right] \mathbb{P}_x(X_{\tau_i^\epsilon}^\epsilon = s_i) \\ & + \left[\frac{1}{\epsilon} \int_x^{s_{i+1}} e^{U(y)/\epsilon} dy \int_x^y \frac{2}{\sigma^2(z)} e^{-U(z)/\epsilon} dz \right] \mathbb{P}_x(X_{\tau_i^\epsilon}^\epsilon = s_{i+1}). \end{aligned}$$

It thus follows from Assumption 2.1 that

$$\mathbb{E}_x\tau_i^\epsilon \le \frac{2}{\epsilon\gamma_1^2}\left[\int_{s_i}^x e^{U(y)/\epsilon}dy\int_y^x e^{-U(z)/\epsilon}dz\right]\mathbb{P}_x(X^\epsilon_{\tau_i^\epsilon}=s_i)$$
$$+\frac{2}{\epsilon\gamma_1^2}\left[\int_x^{s_{i+1}} e^{U(y)/\epsilon}dy\int_x^y e^{-U(z)/\epsilon}dz\right]\mathbb{P}_x(X^\epsilon_{\tau_i^\epsilon}=s_{i+1}).$$

For any $x\in[s_i,u_i]$,

$$\begin{aligned}
&\int_x^{s_{i+1}} e^{U(y)/\epsilon}dy\int_x^y e^{-U(z)/\epsilon}dz\\
&=\int_x^{u_i} e^{U(y)/\epsilon}dy\int_x^y e^{-U(z)/\epsilon}dz+\int_{u_i}^{s_{i+1}} e^{U(y)/\epsilon}dy\left[\int_x^{u_i} e^{-U(z)/\epsilon}dz+\int_{u_i}^y e^{-U(z)/\epsilon}dz\right]\\
&\le\int_x^{s_{i+1}} e^{U(y)/\epsilon}dy\int_x^{u_i} e^{-U(z)/\epsilon}dz+\int_{u_i}^{s_{i+1}} e^{U(y)/\epsilon}dy\int_{u_i}^y e^{-U(z)/\epsilon}dz
\end{aligned}$$

Thus it follows from (2.12) that

$$\begin{aligned}
&\left[\int_x^{s_{i+1}} e^{U(y)/\epsilon}dy\int_x^y e^{-U(z)/\epsilon}dz\right]\mathbb{P}_x(X^\epsilon_{\tau_i^\epsilon}=s_{i+1})\\
&\le\left[\int_x^{s_{i+1}} e^{U(y)/\epsilon}dy\int_x^{u_i} e^{-U(z)/\epsilon}dz\right]\frac{\int_{s_i}^x e^{U(y)/\epsilon}dy}{\int_{s_i}^{s_{i+1}} e^{U(y)/\epsilon}dy}+\int_{u_i}^{s_{i+1}} e^{U(y)/\epsilon}dy\int_{u_i}^y e^{-U(z)/\epsilon}dz\\
&\le\int_{s_i}^x e^{U(y)/\epsilon}dy\int_x^{u_i} e^{-U(z)/\epsilon}dz+\int_{u_i}^{s_{i+1}} e^{U(y)/\epsilon}dy\int_{u_i}^y e^{-U(z)/\epsilon}dz\\
&\le\int_{s_i}^{u_i} e^{U(y)/\epsilon}dy\int_y^{u_i} e^{-U(z)/\epsilon}dz+\int_{u_i}^{s_{i+1}} e^{U(y)/\epsilon}dy\int_{u_i}^y e^{-U(z)/\epsilon}dz.
\end{aligned}$$

Note that

$$\left[\int_{s_i}^x e^{U(y)/\epsilon}dy\int_y^x e^{-U(z)/\epsilon}dz\right]\mathbb{P}_x(X^\epsilon_{\tau_i^\epsilon}=s_i)\le\int_{s_i}^{u_i} e^{U(y)/\epsilon}dy\int_y^{u_i} e^{-U(z)/\epsilon}dz.$$

Thus for any $x\in[s_i,u_i]$,

$$\mathbb{E}_x\tau_i^\epsilon\le\frac{4}{\epsilon\gamma_1^2}\int_{s_i}^{u_i} e^{U(y)/\epsilon}dy\int_y^{u_i} e^{-U(z)/\epsilon}dz+\frac{2}{\epsilon\gamma_1^2}\int_{u_i}^{s_{i+1}} e^{U(y)/\epsilon}dy\int_{u_i}^y e^{-U(z)/\epsilon}dz. \tag{3.3}$$

Similarly, for any $x\in[u_i,s_{i+1}]$, one can prove that

$$\mathbb{E}_x\tau_i^\epsilon\le\frac{2}{\epsilon\gamma_1^2}\int_{s_i}^{u_i} e^{U(y)/\epsilon}dy\int_y^{u_i} e^{-U(z)/\epsilon}dz+\frac{4}{\epsilon\gamma_1^2}\int_{u_i}^{s_{i+1}} e^{U(y)/\epsilon}dy\int_{u_i}^y e^{-U(z)/\epsilon}dz. \tag{3.4}$$

Combining (3.3) and (3.4), we obtain the result of this lemma. □

We are now in a position to finish the proof of Theorem 3.1.

Proof of Theorem 3.1. By Assumption 2.3, $U(x)$ is convex in a neighborhood of s_i and is concave in a neighborhood of u_i. Thus there exists $b_i>s_i$ and $c_i<u_i$, such that $U(x)$ is convex on $[s_i,b_i]$ and is

concave on $[c_i, u_i]$. Note that

$$\begin{aligned}
&\frac{1}{\epsilon}\int_{s_i}^{u_i} e^{U(y)/\epsilon}dy\int_y^{u_i} e^{-U(z)/\epsilon}dz\\
=&\frac{1}{\epsilon}\int_{s_i}^{b_i} e^{U(y)/\epsilon}dy\int_y^{b_i} e^{-U(z)/\epsilon}dz+\frac{1}{\epsilon}\int_{s_i}^{b_i} e^{U(y)/\epsilon}dy\int_{b_i}^{u_i} e^{-U(z)/\epsilon}dz\\
&+\frac{1}{\epsilon}\int_{b_i}^{c_i} e^{U(y)/\epsilon}dy\int_y^{u_i} e^{-U(z)/\epsilon}dz+\frac{1}{\epsilon}\int_{c_i}^{u_i} e^{U(y)/\epsilon}dy\int_y^{u_i} e^{-U(z)/\epsilon}dz\\
:=&\,\mathrm{I}+\mathrm{II}+\mathrm{III}+\mathrm{IV}
\end{aligned}$$

We shall next estimate I, II, III, and IV, respectively. To estimate I, let

$$g_i^\epsilon(x)=\frac{1}{\epsilon}\int_x^{b_i} e^{U(y)/\epsilon}dy\int_y^{b_i} e^{-U(z)/\epsilon}dz.$$

Note that the value of $g_i^\epsilon(x)$ does not change under any translation of $U(x)$. Without loss of generality, we assume that $U(s_i)=0$. By Assumption 2.3, there exists $\delta>0$ such that $U(y)\le U''(s_i)(y-s_i)^2$ for any $y\in[s_i, s_i+\delta]$. By Laplace's method [19, Page 277], it is easy to check that

$$\lim_{\epsilon\to 0}\int_{s_i}^{b_i} e^{-U(z)/\epsilon}dz\Big/\sqrt{\frac{\pi\epsilon}{2U''(s_i)}}=1.$$

This shows that when ϵ is sufficiently small, we have

$$\int_{s_i}^{b_i} e^{-U(z)/\epsilon}dz\le\sqrt{2\pi\epsilon/U''(s_i)}.$$

Thus when ϵ is sufficiently small, we have

$$\begin{aligned}
g_i^\epsilon(s_i)-g_i^\epsilon(s_i+\sqrt{\epsilon})&=\frac{1}{\epsilon}\int_{s_i}^{s_i+\sqrt{\epsilon}} e^{U(y)/\epsilon}dy\int_y^{b_i} e^{-U(z)/\epsilon}dz\\
&\le\left[\frac{1}{\epsilon}\int_{s_i}^{s_i+\sqrt{\epsilon}} e^{U''(s_i)(y-s_i)^2/\epsilon}dy\right]\sqrt{2\pi\epsilon/U''(s_i)}\\
&=\frac{\sqrt{\pi}}{U''(s_i)}\int_0^{\sqrt{2U''(s_i)}} e^{y^2/2}dy.
\end{aligned}$$

Since $U(x)$ is convex on $[s_i, b_i]$, we have $U'(z)/U'(y)\ge 1$ for any $s_i\le y\le z\le b_i$. Therefore,

$$\begin{aligned}
g_i^\epsilon(s_i+\sqrt{\epsilon})&\le\frac{1}{\epsilon}\int_{s_i+\sqrt{\epsilon}}^{b_i} e^{U(y)/\epsilon}dy\int_y^{b_i} e^{-U(z)/\epsilon}\frac{U'(z)}{U'(y)}dz\\
&\le\int_{s_i+\sqrt{\epsilon}}^{b_i} e^{U(y)/\epsilon}\frac{1}{U'(y)}e^{-U(y)/\epsilon}dy=\int_{s_i+\sqrt{\epsilon}}^{b_i}\frac{1}{U'(y)}dy.
\end{aligned}$$

Applying L'Hospital's rule and Assumption 2.3(i), we obtain that

$$\limsup_{\epsilon\to 0}\frac{g_i^\epsilon(s_i+\sqrt{\epsilon})}{\log(1/\epsilon)}\le\lim_{\epsilon\to 0}\frac{\int_{s_i+\sqrt{\epsilon}}^{b_i}\frac{1}{U'(y)}dy}{\log(1/\epsilon)}=\lim_{\epsilon\to 0}\frac{\sqrt{\epsilon}}{2U'(s_i+\sqrt{\epsilon})}=\frac{1}{2U''(s_i)}. \tag{3.5}$$

Therefore,

$$\limsup_{\epsilon\to 0}\frac{\mathrm{I}}{\log(1/\epsilon)}=\limsup_{\epsilon\to 0}\frac{g_i^\epsilon(s_i)-g_i^\epsilon(s_i+\sqrt{\epsilon})+g_i^\epsilon(s_i+\sqrt{\epsilon})}{\log(1/\epsilon)}\leq\frac{1}{2U''(s_i)}. \tag{3.6}$$

Note that $U'(x)$ is continuous and $U'(x)>0$ for any $x\in[b_i,c_i]$. This shows that

$$r:=\min_{x\in[b_i,c_i]}U'(x)>0.$$

Direct computation shows that

$$\begin{aligned}
\mathrm{II}&\leq\frac{1}{\epsilon}(b_i-s_i)e^{U(b_i)/\epsilon}\left[\int_{b_i}^{c_i}e^{-U(z)/\epsilon}dz+\int_{c_i}^{u_i}e^{-U(c_i)/\epsilon}dz\right]\\
&\leq\frac{1}{\epsilon}(b_i-s_i)\left[\int_{b_i}^{c_i}e^{-(U(z)-U(b_i))/\epsilon}dz+(u_i-c_i)e^{-(U(c_i)-U(b_i))/\epsilon}\right]\\
&\leq\frac{1}{\epsilon}(b_i-s_i)\int_{b_i}^{c_i}e^{-r(z-b_i)/\epsilon}dz+\frac{1}{\epsilon}(b_i-s_i)(u_i-c_i)e^{-(U(c_i)-U(b_i))/\epsilon}\\
&\leq\frac{b_i-s_i}{r}+\frac{1}{\epsilon}(b_i-s_i)(u_i-c_i)e^{-(U(c_i)-U(b_i))/\epsilon}.
\end{aligned}$$

In analogy to the estimation of II, we have

$$\begin{aligned}
\mathrm{III}&\leq\frac{1}{\epsilon}\int_{b_i}^{c_i}e^{U(y)/\epsilon}dy\left[\int_y^{c_i}e^{-U(z)/\epsilon}dz+\int_{c_i}^{u_i}e^{-U(c_i)/\epsilon}dz\right]\\
&\leq\frac{1}{\epsilon}\int_{b_i}^{c_i}dy\int_y^{c_i}e^{-r(z-y)/\epsilon}dz+\frac{1}{\epsilon}(u_i-c_i)\int_{b_i}^{c_i}e^{-r(c_i-y)/\epsilon}dy\leq\frac{c_i-b_i}{r}+\frac{u_i-c_i}{r}.
\end{aligned}$$

To estimate IV, let $\tilde{U}(x)=-U(x)$. By Fubini's theorem, we have

$$\mathrm{IV}=\frac{1}{\epsilon}\int_{c_i}^{u_i}e^{\tilde{U}(z)/\epsilon}dz\int_{c_i}^{z}e^{-\tilde{U}(y)/\epsilon}dy.$$

By Assumption 2.3, it is easy to see that u_i is a local minimum point of $\tilde{U}(x)$, $\tilde{U}(x)$ is convex on $[c_i,u_i]$, and $\tilde{U}''(u_i)>0$. In analogy to the estimation of I, it is easy to check that

$$\limsup_{\epsilon\to 0}\frac{\mathrm{IV}}{\log(1/\epsilon)}\leq-\frac{1}{2U''(u_i)}.$$

Therefore,

$$\begin{aligned}
&\limsup_{\epsilon\to 0}\frac{1}{\epsilon\log(1/\epsilon)}\int_{s_i}^{u_i}e^{U(y)/\epsilon}dy\int_y^{u_i}e^{-U(z)/\epsilon}dz\\
=&\limsup_{\epsilon\to 0}\frac{\mathrm{I}+\mathrm{II}+\mathrm{III}+\mathrm{IV}}{\log(1/\epsilon)}\leq\frac{1}{2U''(s_i)}-\frac{1}{2U''(u_i)}.
\end{aligned}$$

Similarly, one can show that

$$\limsup_{\epsilon\to 0}\frac{1}{\epsilon\log(1/\epsilon)}\int_{u_i}^{s_{i+1}}e^{U(y)/\epsilon}dy\int_{u_i}^{y}e^{-U(z)/\epsilon}dz\leq\frac{1}{2U''(s_{i+1})}-\frac{1}{2U''(u_i)}.$$

Thus by Lemma 3.5, we obtain the desired result. □

3.3 Proof of Theorem 3.3

In order to prove Theorem 3.3, we need a lemma.

Lemma 3.6. Let $\tau_\alpha^\epsilon = \inf\{t \geq 0 : X_t^\epsilon = \alpha\}$. Then under Assumption 3.2, for any $x \geq \alpha \geq s_k$,

$$\mathbb{E}_x \tau_\alpha^\epsilon = \frac{1}{\epsilon} \int_\alpha^x e^{U(y)/\epsilon} dy \int_y^\infty \frac{2}{\sigma^2(z)} e^{-U(z)/\epsilon} dz. \tag{3.7}$$

Proof. It is a classical result [11, Pages 343-344, (5.55) and (5.59)] that for any $\beta \geq x$,

$$\begin{aligned}
&\mathbb{E}_x \tau_\alpha^\epsilon \wedge \tau_\beta^\epsilon \\
=& \left[\frac{1}{\epsilon} \int_\alpha^\beta e^{U(y)/\epsilon} dy \int_\alpha^y \frac{2}{\sigma^2(z)} e^{-U(z)/\epsilon} dz\right] \frac{\int_\alpha^x e^{U(y)/\epsilon} dy}{\int_\alpha^\beta e^{U(y)/\epsilon} dy} - \frac{1}{\epsilon} \int_\alpha^x e^{U(y)/\epsilon} dy \int_\alpha^y \frac{2}{\sigma^2(z)} e^{-U(z)/\epsilon} dz \\
=& \frac{1}{\epsilon} \int_\alpha^x e^{U(y)/\epsilon} dy \int_y^\infty \frac{2}{\sigma^2(z)} e^{-U(z)/\epsilon} dz - \frac{1}{\epsilon} \int_\alpha^x e^{U(y)/\epsilon} dy \left[\frac{\int_\alpha^\beta e^{U(y)/\epsilon} dy \int_y^\infty \frac{2}{\sigma^2(z)} e^{-U(z)/\epsilon} dz}{\int_\alpha^\beta e^{U(y)/\epsilon} dy}\right].
\end{aligned}$$

Applying Assumptions 2.1(ii), 3.2, and L'Hospital's rule, we have

$$\limsup_{\beta \to \infty} \frac{\int_\alpha^\beta e^{U(y)/\epsilon} dy \int_y^\infty \frac{2}{\sigma^2(z)} e^{-U(z)/\epsilon} dz}{\int_\alpha^\beta e^{U(y)/\epsilon} dy} \leq \lim_{\beta \to \infty} \frac{2}{\gamma_1^2} \int_\beta^\infty e^{-U(y)/\epsilon} dy = 0.$$

It follows from Remark 2.2 that X^ϵ is non-explosive. This implies that $\lim_{\beta \to \infty} \tau_\beta^\epsilon = \infty$, a.s. Thus

$$\mathbb{E}_x \tau_\alpha^\epsilon = \lim_{\beta \to \infty} \mathbb{E}_x \tau_\alpha^\epsilon \wedge \tau_\beta^\epsilon = \frac{1}{\epsilon} \int_\alpha^x e^{U(y)/\epsilon} dy \int_y^\infty \frac{2}{\sigma^2(z)} e^{-U(z)/\epsilon} dz,$$

which gives the desired result. □

We are now in a position to finish the proof of Theorem 3.3.

Proof of Theorem 3.3. The proofs of (3.1) and (3.2) are totally the same. Thus we only prove (3.1). By Assumption 2.3, $U(x)$ is convex in a neighborhood of s_k. Thus there exists $b_k > s_k$, such that $U(x)$ is convex on $[s_k, b_k]$. It is easy to see that $\mathbb{E}_x \tau_{s_k}^\epsilon$ is an increasing function of x. Without loss of generality, we assume that $x \geq b_k$ and $U(s_k) = 0$. By Assumptions 2.1(ii) and Lemma 3.6, we have

$$\mathbb{E}_x \tau_{s_k}^\epsilon \leq \frac{2}{\gamma_1^2} \left[\frac{1}{\epsilon} \int_{s_k}^x e^{U(y)/\epsilon} dy \int_y^\infty e^{-U(z)/\epsilon} dz\right].$$

Note that

$$\begin{aligned}
&\frac{1}{\epsilon} \int_{s_k}^x e^{U(y)/\epsilon} dy \int_y^\infty e^{-U(z)/\epsilon} dz \\
=& \frac{1}{\epsilon} \int_{s_k}^{b_k} e^{U(y)/\epsilon} dy \int_y^{b_k} e^{-U(z)/\epsilon} dz + \frac{1}{\epsilon} \int_{s_k}^{b_k} e^{U(y)/\epsilon} dy \int_{b_k}^\infty e^{-U(z)/\epsilon} dz \\
&+ \frac{1}{\epsilon} \int_{b_k}^x e^{U(y)/\epsilon} dy \int_y^x e^{-U(z)/\epsilon} dz + \frac{1}{\epsilon} \int_{b_k}^x e^{U(y)/\epsilon} dy \int_x^\infty e^{-U(z)/\epsilon} dz \\
:=& \mathrm{I} + \mathrm{II} + \mathrm{III} + \mathrm{IV}.
\end{aligned}$$

We shall next estimate I, II, III, and IV, respectively. In analogy to the proof of (3.6) in Theorem 3.1, it is easy to check that

$$\limsup_{\epsilon\to 0}\frac{\mathrm{I}}{\log(1/\epsilon)}\le\frac{1}{2U''(s_k)}.$$

To estimate II, we arbitrarily choose $0<\delta<b_k-s_k$. Note that $U'(x)$ is continuous and $U'(y)>0$ for any $y\in[s_k+\delta,x]$. This shows that

$$r:=\min_{y\in[s_k+\delta,x]}U'(y)>0.$$

By Assumption 3.2, when ϵ is sufficiently small,

$$\begin{aligned}
\mathrm{II}&=\frac{1}{\epsilon}\int_{s_k}^{b_k}dy\int_{b_k}^{\infty}e^{-\gamma_3(U(z)-U(y))-(1/\epsilon-\gamma_3)(U(z)-U(y))}dz\\
&\le\frac{1}{\epsilon}e^{\gamma_3U(b_k)}\int_{b_k}^{\infty}e^{-\gamma_3U(z)}dz\int_{s_k}^{b_k}e^{-(1/\epsilon-\gamma_3)(U(b_k)-U(y))}dy\\
&\le\frac{1}{\epsilon}e^{\gamma_3U(b_k)}\int_{b_k}^{\infty}e^{-\gamma_3U(z)}dz\left[\int_{s_k}^{s_k+\delta}e^{-(U(b_k)-U(y))/2\epsilon}dy+\int_{s_k+\delta}^{b_k}e^{-(U(b_k)-U(y))/2\epsilon}dy\right]\\
&\le\frac{1}{\epsilon}e^{\gamma_3U(b_k)}\int_{b_k}^{\infty}e^{-\gamma_3U(z)}dz\left[\delta e^{-(U(b_k)-U(s_k+\delta))/2\epsilon}+\int_{s_k+\delta}^{b_k}e^{-r(b_k-y)/2\epsilon}dy\right]\\
&\le e^{\gamma_3U(b_k)}\int_{b_k}^{\infty}e^{-\gamma_3U(z)}dz\left[\frac{\delta}{\epsilon}e^{-(U(b_k)-U(s_k+\delta))/2\epsilon}+\frac{2}{r}\right].
\end{aligned}$$

To estimate III, note that

$$\mathrm{III}=\frac{1}{\epsilon}\int_{b_k}^{x}dy\int_{y}^{x}e^{-(U(z)-U(y))/\epsilon}dz\le\frac{1}{\epsilon}\int_{b_k}^{x}dy\int_{y}^{x}e^{-r(z-y)/\epsilon}dz\le\frac{x-b_k}{r}.$$

By Assumption 3.2, when ϵ is sufficiently small,

$$\begin{aligned}
\mathrm{IV}&=\frac{1}{\epsilon}\int_{b_k}^{x}dy\int_{x}^{\infty}e^{-\gamma_3(U(z)-U(y))-(1/\epsilon-\gamma_3)(U(z)-U(y))}dz\\
&\le\frac{1}{\epsilon}e^{\gamma_3U(x)}\int_{x}^{\infty}e^{-\gamma_3U(z)}dz\int_{b_k}^{x}e^{-(U(x)-U(y))/2\epsilon}dy\le\frac{2}{r}e^{\gamma_3U(x)}\int_{x}^{\infty}e^{-\gamma_3U(z)}dz.
\end{aligned}$$

Thus by Lemma 3.6, it is easy to check that

$$\limsup_{\epsilon\to 0}\frac{\mathbb{E}_x\tau_{s_k}^{\epsilon}}{\log(1/\epsilon)}\le\limsup_{\epsilon\to 0}\frac{2}{\gamma_1^2}\cdot\frac{\mathrm{I}+\mathrm{II}+\mathrm{III}+\mathrm{IV}}{\log(1/\epsilon)}\le\frac{1}{\gamma_1^2U''(s_k)},$$

which gives the desired result. □

4 L^p maximum inequality for randomly perturbed dynamical systems

4.1 Main results of this section

In this section, we shall study the L^p maximum inequality for randomly perturbed dynamical systems. The following theorem gives an L^p maximum inequality for diffusion processes with convex potential.

Theorem 4.1. Let $X = (X_t)_{t\geq 0}$ be a diffusion process solving the stochastic differential equation

$$dX_t = b(X_t)dt + dB_t, \quad X_0 = 0,$$

where $b(x)$ is locally Lipschitz continuous and monotonically decreasing. Let $F(x)$ be the function defined by

$$F(x) = 2\int_0^x e^{U(y)}dy \int_0^y e^{-U(z)}dz,$$

where $U(x) = -2\int_0^x b(y)dy$ is the potential of X. Assume that the following condition is satisfied:

$$\sup_{x>0}\left[\frac{F(x)}{x}\int_x^\infty \frac{dz}{F(z)}\right] < \infty \tag{4.1}$$

and let $H(x) = F^{-1}(x)$ denote the inverse of $F(x)$ for $x \geq 0$. Then for any $p > 0$, there exists a constant $c_p > 0$ such that for any stopping time τ,

$$\mathbb{E}[\sup_{0\leq t\leq\tau} X_t]^p \leq c_p \mathbb{E}H(\tau)^p.$$

Proof. This theorem will be proved in Section 4.2. □

Applying the above theorem, we can obtain an L^p maximum inequality for the OU process.

Corollary 4.2. Let $V = (V_t)_{t\geq 0}$ be an OU process solving the stochastic differential equation

$$dV_t = -\alpha V_t dt + \beta dB_t, \quad V_0 = 0,$$

where $\alpha, \beta > 0$. Then for any $p > 0$, there exists a constant $c_p > 0$ such that for any stopping time τ,

$$\mathbb{E}[\sup_{0\leq t\leq\tau} |V_t|]^p \leq \frac{c_p\beta^p}{\alpha^{p/2}}\mathbb{E}\log^{p/2}(1+\alpha\tau).$$

Proof. Let $X_t = V_t/\beta$. Then X is the solution to the following stochastic differential equation:

$$dX_t = -\alpha X_t dt + dB_t, \quad X_0 = 0.$$

It is easy to check that the potential of X is $U(x) = \alpha x^2$. This shows that

$$F(x) = 2\int_0^x e^{U(y)}dy\int_0^y e^{-U(z)}dz = 2\int_0^x e^{\alpha y^2}dy\int_0^y e^{-\alpha z^2}dz.$$

A successive application of L'Hospital's rule then shows that (4.1) holds. By estimating the inverse $H(x) = F^{-1}(x)$, it is easy to check that

$$\frac{1}{\sqrt{\alpha}}\sqrt{\log(1+\alpha x)} \leq H(x) \leq \frac{\sqrt{2}}{\sqrt{\alpha}}\sqrt{\log(1+\alpha x)}.$$

It thus follows from Theorem 4.1 that for any $p > 0$, there exists a constant $c_p > 0$ such that for any stopping time τ,

$$\mathbb{E}[\sup_{0\leq t\leq\tau} X_t]^p \leq \frac{c_p}{\alpha^{p/2}}\mathbb{E}\log^{p/2}(1+\alpha\tau). \tag{4.2}$$

This shows that

$$\mathbb{E}[\sup_{0\leq t\leq\tau} V_t]^p \leq \frac{c_p\beta^p}{\alpha^{p/2}}\mathbb{E}\log^{p/2}(1+\alpha\tau). \tag{4.3}$$

By symmetry, it is easy to see that the above inequality also holds when V_t is replaced by $-V_t$. This completes the proof of this corollary. □

Remark 4.3. In fact, the L^1 maximum inequality for the OU process was first obtained by Graversen and Peskir [8]. Subsequently, Yan et al. [21] stated the L^p maximum inequality with $p \neq 1$ for the OU process. However, although their ideas are fairly nice, their detailed proofs are questionable because they mistakenly regarded the random time $TI_{\{S<T\}}$ as a stopping time, where S and T are two stopping times with $S \leq T$ (see [21, Page 6, Lines 2 and 10]). In this paper, we provide a complete proof of the L^p maximum inequality for the OU process.

Based on Theorem 4.1, we can obtain an L^p maximum inequality for randomly perturbed dynamical systems, which is stated in the following theorem.

Theorem 4.4. Assume that the time interval $h(\epsilon)$ satisfies (2.4). Then for any $1 \leq i \leq k$ and $p > 0$, there exist constants $\alpha_i > 0$, $\beta_i > 0$, and $K_p > 0$ such that

$$\mathbb{E}_{s_i}[\sup_{0\leq t\leq h(\epsilon)} (X_t^\epsilon - s_i)]^p \leq K_p[\epsilon \log(1+\alpha_i h(\epsilon))]^{p/2} + K_p e^{-\beta_i/\epsilon}\left[(\epsilon h(\epsilon))^{p/2} + 1\right] \tag{4.4}$$

when ϵ is sufficiently small.

Proof. This theorem will be proved in Section 4.3. □

The proof of Lemma 2.7 follows from the above theorem immediately.

Proof of Lemma 2.7. Fix $1 \leq i \leq k$ and $p > 0$. For any $\eta > 0$, when ϵ is sufficiently small, we have $h(\epsilon) \leq e^{\eta/\epsilon}$. Thus when ϵ is sufficiently small,

$$\epsilon \log(1+\alpha_i h(\epsilon)) \leq \epsilon \log(1+\alpha_i e^{\eta/\epsilon}) \leq \epsilon \log(e^{2\eta/\epsilon}) = 2\eta.$$

By the arbitrariness of η, we obtain that

$$\lim_{\epsilon\to 0} \epsilon \log(1+\alpha_i h(\epsilon)) = 0. \tag{4.5}$$

On the other hand, when ϵ is sufficiently small, we have

$$e^{-\beta_i/\epsilon} h(\epsilon)^{p/2} = \exp\left[-\frac{\beta_i}{\epsilon} + \frac{p}{2\epsilon} o(1)\right] \leq e^{-\beta_i/2\epsilon}.$$

This shows that

$$\lim_{\epsilon\to 0} e^{-\beta_i/\epsilon}(\epsilon h(\epsilon))^{p/2} = 0. \tag{4.6}$$

Thus (4.5) and (4.6), together with Theorem 4.4, show that

$$\lim_{\epsilon\to 0} \mathbb{E}_{s_i}[\sup_{0\leq t\leq h(\epsilon)} (X_t^\epsilon - s_i)]^p = 0. \tag{4.7}$$

By symmetry, it is easy to check that the above equality holds when $X_t^\epsilon - s_i$ is replaced by $s_i - X_t^\epsilon$. This completes the proof of this lemma. □

4.2 Proof of Theorem 4.1

In order to prove Theorem 4.1, we need two lemmas. The following lemma, which gives the L^1 maximum inequality for diffusion processes, can be found in [15, Theorem 2.5].

Lemma 4.5. Let $X = (X_t)_{t\geq 0}$ be a diffusion process solving the stochastic differential equation

$$dX_t = b(X_t)dt + dB_t, \quad X_0 = 0,$$

where $b(x)$ is continuous. Assume that the condition (4.1) holds. Then there exists a universal constant $c > 0$ such that for any stopping time τ,

$$\mathbb{E}[\sup_{0\leq t\leq \tau} X_t] \leq c\mathbb{E}H(\tau).$$

The following useful lemma can be found in [2, Lemma 4.1].

Lemma 4.6. Let $A = (A_t)_{t\geq 0}$ and $B = (B_t)_{t\geq 0}$ be two continuous, adapted, increasing processes with $A_0 = 0$ and $B_0 \geq 0$. Assume that there exist $q > 0$ and $c > 0$ such that for any pair of finite stopping times S and T with $S \leq T$,

$$\mathbb{E}[(A_T - A_S)^q] \leq c\|B_T\|_\infty^q \mathbb{P}(S < T).$$

Then for any $p > 0$, there exists a constant $c_p > 0$ such that for any stopping time τ,

$$\mathbb{E}A_\tau^p \leq c_p \mathbb{E}B_\tau^p.$$

Based on the above two lemmas, we are in a position to prove Theorem 4.1.

Proof of Theorem 4.1. For any pair of finite stopping times S and T with $S \leq T$, it is easy to check that

$$\sup_{0\leq t\leq T} X_t - \sup_{0\leq t\leq S} X_t = \sup_{0\leq t\leq S} X_t \vee \sup_{S\leq t\leq T} X_t - \sup_{0\leq t\leq S} X_t \leq \sup_{S\leq t\leq T} X_t$$

and that

$$\sup_{0\leq t\leq T} X_t - \sup_{0\leq t\leq S} X_t = \sup_{0\leq t\leq S} X_t \vee (X_S + \sup_{S\leq t\leq T}(X_t - X_S)) - \sup_{0\leq t\leq S} X_t \leq \sup_{S\leq t\leq T}(X_t - X_S).$$

Thus we have

$$\begin{aligned}\mathbb{E}[\sup_{0\leq t\leq T} X_t - \sup_{0\leq t\leq S} X_t] &\leq \mathbb{E}[\sup_{S\leq t\leq T} X_t \wedge \sup_{S\leq t\leq T}(X_t - X_S)] \\ &\leq \mathbb{E}[\sup_{S\leq t\leq T} X_t; X_S < 0] + \mathbb{E}[\sup_{S\leq t\leq T}(X_t - X_S); X_S \geq 0] := \mathrm{I} + \mathrm{II}.\end{aligned} \tag{4.8}$$

To estimate I, let $R = \inf\{t \geq S : X_t = 0\}$ and $Y_t = X_{R+t}$. Direct computation shows that

$$\begin{aligned}\mathrm{I} &= \mathbb{E}[\sup_{S\leq t\leq T} X_t; X_S < 0, T \leq R] + \mathbb{E}[\sup_{S\leq t\leq T} X_t; X_S < 0, T > R] \\ &\leq \mathbb{E}[\sup_{S\leq t\leq T} X_t; X_S < 0, T > R] = \mathbb{E}[\sup_{R\leq t\leq T} X_t; X_S < 0, T > R] \\ &\leq \mathbb{E}[\sup_{0\leq t\leq T-R} Y_t; T > R] = \mathbb{E}[\sup_{0\leq t\leq T\vee R-R} Y_t].\end{aligned}$$

Since R is a stopping time, $W_t = B_{R+t} - B_R$ is an $\{\mathcal{F}_{R+t}\}$-Brownian motion. It is easy to check that Y is the solution to the stochastic differential equation

$$dY_t = b(Y_t)dt + dW_t, \quad Y_0 = 0.$$

Note that Y is an $\{\mathcal{F}_{R+t}\}$-adapted process and $T \vee R - R$ is an $\{\mathcal{F}_{R+t}\}$-stopping time. By Lemma 4.5, we obtain that

$$\mathrm{I} \le c\mathbb{E}[H(T \vee R - R)] \le c\|H(T)\|_\infty \mathbb{P}(R < T) \le c\|H(T)\|_\infty \mathbb{P}(S < T). \tag{4.9}$$

To estimate II, let $Z_t = X_{S+t}$. Direct computation shows that

$$\begin{aligned}
\mathrm{II} &= \mathbb{E}[\sup_{0\le t\le T-S}(X_{S+t} - X_S); X_S \ge 0] \\
&= \int_{[0,\infty)} \mathbb{E}[\sup_{0\le t\le T-S}(X_{S+t} - X_S)|X_S = a]\mathbb{P}(X_S \in da) \\
&= \int_{[0,\infty)} \mathbb{E}_a^Z[\sup_{0\le t\le T-S}(Z_t - a)]\mathbb{P}(X_S \in da),
\end{aligned} \tag{4.10}$$

where $\mathbb{E}_a^Z(\cdot) = \mathbb{E}(\cdot|Z_0 = a)$. Since S is a stopping time, $\beta_t = B_{S+t} - B_S$ is an $\{\mathcal{F}_{S+t}\}$-Brownian motion independent of $\mathcal{F}_S$. This implies that under $\mathbb{P}_a^Z$, β_t is still an $\{\mathcal{F}_{S+t}\}$-Brownian motion. For any $a \in \mathbb{R}$, let $Z_t^a = Z_t - a$ and $b^a(x) = b(x + a)$. It is easy to check that under $\mathbb{P}_a^Z$, Z^a is the solution to the following stochastic integral equation:

$$Z_t^a = \int_0^t b^a(Z_s^a)ds + \beta_t.$$

Moreover, let L be the solution to the following stochastic integral equation under $\mathbb{P}_a^Z$:

$$L_t = \int_0^t b(L_s)ds + \beta_t.$$

Since $b(x)$ is monotonically decreasing, it is easy to see that $b^a(x) \le b(x)$ for any $a \ge 0$ and $x \in \mathbb{R}$. By the comparison theorem [9, Page 437, Theorem 1.1], we have $Z_t^a \le L_t$ for any $t \ge 0$ with probability one under $\mathbb{P}_a^Z$. Since L is an $\{\mathcal{F}_{S+t}\}$-adapted process and $T - S$ is an $\{\mathcal{F}_{S+t}\}$-stopping time, it follows from Lemma 4.5 that

$$\begin{aligned}
\mathbb{E}_a^Z[\sup_{0\le t\le T-S}(Z_t - a)] = \mathbb{E}_a^Z[\sup_{0\le t\le T-S} Z_t^a] &\le \mathbb{E}_a^Z[\sup_{0\le t\le T-S} L_t] \\
&\le c\mathbb{E}_a^Z[H(T - S)] \le c\|H(T)\|_\infty \mathbb{P}_a^Z(S < T).
\end{aligned}$$

Thus by (4.10) we have

$$\begin{aligned}
\mathrm{II} &\le \int_{[0,\infty)} c\|H(T)\|_\infty \mathbb{P}_a^Z(S < T)\mathbb{P}(X_S \in da) \\
&= c\|H(T)\|_\infty \mathbb{P}(S < T, X_S \ge 0) \le c\|H(T)\|_\infty \mathbb{P}(S < T).
\end{aligned} \tag{4.11}$$

Combining (4.8), (4.9), and (4.11), we obtain that

$$\mathbb{E}[\sup_{0\le t\le T} X_t - \sup_{0\le t\le S} X_t] \le 2c\|H(T)\|_\infty \mathbb{P}(S < T).$$

It thus follows from Lemma 4.6 that for any $p > 0$, there exists a constant $c_p > 0$ such that for any stopping time τ,

$$\mathbb{E}[\sup_{0\le t\le\tau} X_t]^p \le c_p \mathbb{E}H(\tau)^p,$$

which gives the desired result. □

4.3 Proof of Theorem 4.4

In order to prove Theorem 4.4, it is convenient to introduce the following definition.

Definition 4.7. Let $X = (X_t)_{t\ge0}$ be a diffusion process solving the stochastic differential equation

$$dX_t = b(X_t)dt + \beta dB_t, \quad X_0 = 0,$$

where $b(x)$ is locally Lipschitz continuous and $\beta > 0$. Then X is called a unilateral OU process if $b(x)$ is monotonically decreasing and there exists $\alpha \ge 0$ such that $b(x) = -\alpha x$ for $x \ge 0$.

The following lemma gives an L^p maximum inequality for unilateral OU processes.

Lemma 4.8. Let $X = (X_t)_{t\ge0}$ be the unilateral OU process described in Definition 4.7. Then for any $p > 0$, there exists a constant $c_p > 0$ such that for any stopping time τ,

$$\mathbb{E}[\sup_{0\le t\le\tau} X_t]^p \le c_p \left[\frac{\beta^p}{\alpha^{p/2}}\mathbb{E}\log^{p/2}(1+\alpha\tau)I_{\{\alpha>0\}} + \beta^p\mathbb{E}\tau^{p/2}I_{\{\alpha=0\}}\right]. \tag{4.12}$$

Proof. Let $Y_t = X_t/\beta$ and let $\tilde{b}(x) = b(\beta x)/\beta$. Then Y is the solution to the following stochastic differential equation:

$$dY_t = \tilde{b}(Y_t)dt + dB_t, \quad Y_0 = 0.$$

Let $\tilde{U}(x)$ denote the potential function of Y.

We first consider the case when $\alpha = 0$. In this case, we have $\tilde{b}(x) = \tilde{U}(x) = 0$ for any $x \ge 0$. Thus for any $x \ge 0$,

$$F(x) = 2\int_0^x e^{\tilde{U}(y)}dy\int_0^y e^{-\tilde{U}(z)}dz = x^2.$$

and $H(x) = F^{-1}(x) = \sqrt{x}$. It is easy to check that

$$\sup_{x>0}\left[\frac{F(x)}{x}\int_x^\infty \frac{dz}{F(z)}\right] = 1 < \infty.$$

It thus follows from Theorem 4.1 that for any $p > 0$, there exists a constant $c_p > 0$ such that for any stopping time τ,

$$\mathbb{E}[\sup_{0\le t\le\tau} Y_t]^p \le c_p\mathbb{E}\tau^{p/2}.$$

This shows that (4.12) holds when $\alpha = 0$.

We next consider the case when $\alpha > 0$. In this case, we have $\tilde{b}(x) = -\alpha x$ and $\tilde{U}(x) = \alpha x^2$ for any $x \ge 0$. In analogy to the proof of Corollary 4.2, it is easy to prove that for any $p > 0$, there exists a constant $c_p > 0$ such that for any stopping time τ,

$$\mathbb{E}[\sup_{0\le t\le\tau} Y_t]^p \le \frac{c_p}{\alpha^{p/2}}\mathbb{E}\log^{p/2}(1+\alpha\tau).$$

This shows that (4.12) holds when $\alpha > 0$. □

We are now in a position to give the proof of Theorem 4.4.

Proof of Theorem 4.4. We first give the proof of the theorem when $\sigma(x) \equiv 1$. By Assumption 2.3, for any $1 \le i \le k$, there exists $\delta > 0$ such that for any $x \in [s_i, s_i + \delta]$,

$$-2b(x) = U'(x) \ge \frac{1}{2} U''(s_i)(x - s_i).$$

This shows that $b(x) \le -\alpha_i(x - s_i)$ for any $x \in [s_i, s_i + \delta]$, where $\alpha_i = U''(s_i)/4$. Since $b(x)$ is locally Lipschitz continuous, there exists a locally Lipschitz continuous function $b_i(x)$ such that $b_i(x)$ is monotonically decreasing, $b_i(s_i) = 0$, and $b(x) \le b_i(x)$ for any $x \le s_i$. Let

$$f_i(x) = \begin{cases} -\alpha_i(x - s_i), & x > s_i, \\ b_i(x), & x \le s_i. \end{cases}$$

It is easy to see that $b(x) \le f_i(x)$ for any $x \le s_i + \delta$ and $f_i(x)$ is monotonically decreasing. For any $1 \le i \le k$, let $X^{i,\epsilon}$ be the unilateral OU process solving the stochastic differential equation

$$dX_t^{i,\epsilon} = f_i(X_t^{i,\epsilon})dt + \sqrt{\epsilon} dB_t, \quad X_0^{i,\epsilon} = s_i.$$

Thus we have

$$\begin{aligned} &\mathbb{E}_{s_i}[\sup_{0 \le t \le h(\epsilon)} (X_t^\epsilon - s_i)]^p \\ &= \mathbb{E}_{s_i}\left[[\sup_{0 \le t \le h(\epsilon)} (X_t^\epsilon - s_i)]^p; h(\epsilon) \le \tau_{s_i+\delta}^\epsilon\right] + \mathbb{E}_{s_i}\left[[\sup_{0 \le t \le h(\epsilon)} (X_t^\epsilon - s_i)]^p; h(\epsilon) > \tau_{s_i+\delta}^\epsilon\right] \\ &:= \mathrm{I} + \mathrm{II}. \end{aligned} \tag{4.13}$$

Since $b(x) \le f_i(x)$ for any $x \le s_i + \delta$, by the comparison theorem [9, Page 437, Theorem 1.1] and the local property of diffusion processes, we have

$$\mathrm{I} \le \mathbb{E}_{s_i}\left[[\sup_{0 \le t \le h(\epsilon)} (X_t^\epsilon - s_i) \wedge \delta]^p\right] = \mathbb{E}_{s_i}\left[[\sup_{0 \le t \le h(\epsilon)} (X_t^{i,\epsilon} - s_i) \wedge \delta]^p\right] \le \mathbb{E}_{s_i}[\sup_{0 \le t \le h(\epsilon)} (X_t^{i,\epsilon} - s_i)]^p.$$

It follows from Lemma 4.8 that for any $p > 0$, there exists a constant $c_p > 0$ such that

$$\mathrm{I} \le \mathbb{E}_{s_i}[\sup_{0 \le t \le h(\epsilon)} (X_t^{i,\epsilon} - s_i)]^p \le c_p[\epsilon \log(1 + \alpha_i h(\epsilon))]^{p/2}. \tag{4.14}$$

On the other hand, by the strong Markov property of X^ϵ at $\tau_{s_i+\delta}^\epsilon$, we have

$$\begin{aligned} \mathrm{II} &= \mathbb{E}_{s_i}\left[[\sup_{\tau_{s_i+\delta}^\epsilon \le t \le h(\epsilon)} (X_t^\epsilon - s_i)]^p; h(\epsilon) > \tau_{s_i+\delta}^\epsilon\right] \\ &\le \mathbb{E}_{s_i}\left[[\sup_{0 \le t \le h(\epsilon)} (X_{t+\tau_{s_i+\delta}^\epsilon}^\epsilon - s_i)]^p; h(\epsilon) > \tau_{s_i+\delta}^\epsilon\right] \\ &= \mathbb{E}_{s_i+\delta}[\sup_{0 \le t \le h(\epsilon)} (X_t^\epsilon - s_i)]^p \mathbb{P}_{s_i}(\tau_{s_i+\delta}^\epsilon < h(\epsilon)). \end{aligned} \tag{4.15}$$

When $1 \le i \le k-1$, without loss of generality, we assume that $\delta < s_k - s_{k-1}$. This shows that $s_i + \delta < s_k$. Thus we obtain that

$$\begin{aligned}
&\mathbb{E}_{s_i+\delta}[\sup_{0\le t\le h(\epsilon)}(X_t^\epsilon - s_i)]^p \\
&= \mathbb{E}_{s_i+\delta}\left[[\sup_{0\le t\le h(\epsilon)}(X_t^\epsilon - s_i)]^p; h(\epsilon) \le \tau_{s_k}^\epsilon\right] + \mathbb{E}_{s_i+\delta}\left[[\sup_{0\le t\le h(\epsilon)}(X_t^\epsilon - s_i)]^p; h(\epsilon) > \tau_{s_k}^\epsilon\right] \\
&\le (s_k - s_i)^p + \mathbb{E}_{s_k}\left[[\sup_{0\le t\le h(\epsilon)}(X_t^\epsilon - s_i)]^p\right] \\
&\le (s_k - s_i)^p + \max\{2^{p-1}, 1\}\left[\mathbb{E}_{s_k}[\sup_{0\le t\le h(\epsilon)}(X_t^\epsilon - s_k)]^p + (s_k - s_i)^p\right].
\end{aligned}$$

Recall that there exists a locally Lipschitz continuous function $b_k(x)$ such that $b_k(x)$ is monotonically decreasing, $b_k(s_k) = 0$, and $b(x) \le b_k(x)$ for any $x \le s_k$. Let

$$g(x) = \begin{cases} 0, & x > s_k, \\ b_k(x), & x \le s_k. \end{cases}$$

It is easy to see that $b(x) \le g(x)$ for any $x \in \mathbb{R}$. Let Y^ϵ be the unilateral OU process solving the stochastic differential equation

$$dY_t^\epsilon = g(Y_t^\epsilon)dt + \sqrt{\epsilon}dB_t, \quad Y_0^\epsilon = s_k.$$

By the comparison theorem and Lemma 4.8, for any $p > 0$, there exists a constant $c_p > 0$ such that

$$\mathbb{E}_{s_k}[\sup_{0\le t\le h(\epsilon)}(X_t^\epsilon - s_k)]^p \le \mathbb{E}[\sup_{0\le t\le h(\epsilon)}(Y_t^\epsilon - s_k)]^p \le c_p(\epsilon h(\epsilon))^{p/2}. \tag{4.16}$$

This shows that

$$\mathbb{E}_{s_i+\delta}[\sup_{0\le t\le h(\epsilon)}(X_t^\epsilon - s_i)]^p \le (s_k - s_i)^p + \max\{2^{p-1}, 1\}\left[c_p(\epsilon h(\epsilon))^{p/2} + (s_k - s_i)^p\right]. \tag{4.17}$$

When $i = k$, direct computation shows that

$$\mathbb{E}_{s_k+\delta}[\sup_{0\le t\le h(\epsilon)}(X_t^\epsilon - s_k)]^p \le \max\{2^{p-1}, 1\}\mathbb{E}_{s_k+\delta}\left[[\sup_{0\le t\le h(\epsilon)}(X_t^\epsilon - s_k - \delta)]^p + \delta^p\right].$$

In analogy to the proof of (4.16), we can prove that for any $x \ge s_k$,

$$\mathbb{E}_x[\sup_{0\le t\le h(\epsilon)}(X_t^\epsilon - x)]^p \le c_p(\epsilon h(\epsilon))^{p/2}. \tag{4.18}$$

Thus we have

$$\mathbb{E}_{s_k+\delta}[\sup_{0\le t\le h(\epsilon)}(X_t^\epsilon - s_k)]^p \le \max\{2^{p-1}, 1\}\left[c_p(\epsilon h(\epsilon))^{p/2} + \delta^p\right]. \tag{4.19}$$

It follows from (4.17) and (4.19) that there exists a constant $K_p \ge c_p$ such that for any $1 \le i \le k$,

$$\mathbb{E}_{s_i+\delta}[\sup_{0\le t\le h(\epsilon)}(X_t^\epsilon - s_i)]^p \le K_p\left[(\epsilon h(\epsilon))^{p/2} + 1\right]. \tag{4.20}$$

Furthermore, since we have assumed that $\sigma(x) \equiv 1$, it is a classical result of the Freidlin-Wentzell theory that for any $1 \leq i \leq k$, there exists a constant $\beta_i > 0$ such that

$$\mathbb{P}_{s_i}(\tau^\epsilon_{s_i+\delta} < e^{\beta_i/\epsilon}) \leq e^{-\beta_i/\epsilon}$$

when ϵ is sufficiently small [14, Page 116, Remark 2.41]. By (2.4), we have $h(\epsilon) \leq e^{\beta_i/\epsilon}$ when ϵ is sufficiently small. This shows that when ϵ is sufficiently small,

$$\mathbb{P}_{s_i}(\tau^\epsilon_{s_i+\delta} < h(\epsilon)) \leq e^{-\beta_i/\epsilon}. \tag{4.21}$$

Combining (4.13), (4.14), (4.15), (4.20), and (4.21), we finally obtain that

$$\mathbb{E}_{s_i}[\sup_{0\leq t\leq h(\epsilon)}(X^\epsilon_t - s_i)]^p \leq K_p[\epsilon\log(1+\alpha_i h(\epsilon))]^{p/2} + K_p e^{-\beta_i/\epsilon}\left[(\epsilon h(\epsilon))^{p/2}+1\right]. \tag{4.22}$$

We next give the proof of the theorem for general $\sigma(x)$. Let $M^\epsilon_t = \int_0^t \sigma(X^\epsilon_s)dB_s$. Then the quadratic variation process of M^ϵ is $[M^\epsilon]_t = \int_0^t \sigma^2(X^\epsilon_s)ds$. By Assumption 2.1(ii), $[M^\epsilon]_t$ is strictly increasing and $[M^\epsilon]_\infty = \infty$. Let $\zeta^\epsilon_t = [M^\epsilon]^{-1}_t$ be the inverse of $[M^\epsilon]_t$ and let $Y^\epsilon_t = X^\epsilon_{\zeta^\epsilon_t}$. It is easy to check that Y^ϵ is the solution to the following stochastic differential equation

$$dY^\epsilon_t = \frac{b(Y^\epsilon_t)}{\sigma^2(Y^\epsilon_t)}dt + \sqrt{\epsilon}dW^\epsilon_t, \quad Y^\epsilon_0 = 0,$$

where $W^\epsilon_t = M^\epsilon_{\zeta^\epsilon_t}$ is a standard Brownian motion. By Assumptions 2.1 and 2.3, it is easy to check that $b(x)/\sigma^2(x)$ is locally Lipschitz continuous. Moreover, it is easy to see that

$$\sup_{0\leq t\leq h(\epsilon)} X^\epsilon_t = \sup_{0\leq t\leq [M^\epsilon]_{h(\epsilon)}} Y^\epsilon_t \leq \sup_{0\leq t\leq \gamma_2^2 h(\epsilon)} Y^\epsilon_t.$$

It thus follows from (4.23) that for any $1 \leq i \leq k$ and $p > 0$, there exists a constant $K_p > 0$, $\alpha_i > 0$, and $\beta_i > 0$ such that

$$\mathbb{E}_{s_i}[\sup_{0\leq t\leq h(\epsilon)}(X^\epsilon_t - s_i)]^p \leq K_p[\epsilon\log(1+\alpha_i\gamma_2^2 h(\epsilon))]^{p/2} + K_p e^{-\beta_i/\epsilon}\left[\gamma_2^p(\epsilon h(\epsilon))^{p/2}+1\right], \tag{4.23}$$

which gives the desired result. □

Acknowledgements

The authors gratefully acknowledge M.P. Qian and D.Q. Jiang for stimulating discussions and the anonymous reviewers for their valuable suggestions which greatly improve the quality of this paper. This work is supported by China Postdoctoral Science Foundation (No. 2015M570007).

References

[1] B. C. Bag and C.-K. Hu. Escape through an unstable limit cycle driven by multiplicative colored non-gaussian and additive white gaussian noises. *Phys. Rev. E*, 75(4):042101, 2007.

[2] M. T. Barlow and M. Yor. Semi-martingale inequalities via the Garsia-Rodemich-Rumsey lemma, and applications to local times. *J. Funct. Anal.*, 49(2):198–229, 1982.

[3] L. Chen, R. Liu, Z.-P. Liu, M. Li, and K. Aihara. Detecting early-warning signals for sudden deterioration of complex diseases by dynamical network biomarkers. *Sci. Rep.*, 2:342, 2012.

[4] L. Dai, D. Vorselen, K. S. Korolev, and J. Gore. Generic indicators for loss of resilience before a tipping point leading to population collapse. *Science*, 336(6085):1175–1177, 2012.

[5] A. De Masi, A. Pellegrinotti, E. Presutti, and M. Vares. Spatial patterns when phases separate in an interacting particle system. *Ann. Probab.*, 41:334–371, 1994.

[6] A. De Masi, E. Presutti, and M. Vares. Escape from the unstable equilibrium in a random process with infinitely many interacting particles. *J. Stat. Phys.*, 44(3-4):645–696, 1986.

[7] M. I. Freidlin and A. D. Wentzell. *Random Perturbations of Dynamical Systems*. Springer, New York, 2nd edition, 1998.

[8] S. E. Graversen and G. Peskir. Maximal inequalities for the Ornstein-Uhlenbeck process. *P. Am. Math. Soc.*, pages 3035–3041, 2000.

[9] N. Ikeda and S. Watanabe. *Stochastic Differential Equations and Diffusion Processes*. North-Holland Publishing Company, New York, 2nd edition, 1989.

[10] C. Jia, M. Qian, Y. Kang, and D. Jiang. Modeling stochastic phenotype switching and bet-hedging in bacteria: stochastic nonlinear dynamics and critical state identification. *Quant. Biol.*, 2(3):110–125, 2014.

[11] I. Karatzas and S. Shreve. *Brownian Motion and Stochastic Calculus*. Springer, New York, 2nd edition, 1998.

[12] R. Liu, M. Li, Z.-P. Liu, J. Wu, L. Chen, and K. Aihara. Identifying critical transitions and their leading biomolecular networks in complex diseases. *Sci. Rep.*, 2:813, 2012.

[13] R. S. Maier and D. L. Stein. Oscillatory behavior of the rate of escape through an unstable limit cycle. *Phys. Rev. Lett.*, 77(24):4860, 1996.

[14] E. Olivieri and M. E. Vares. *Large Deviations and Metastability*. Cambridge University Press, Cambridge, 2004.

[15] G. Peskir. Bounding the maximal height of a diffusion by the time elapsed. *J. Theor. Probab.*, 14(3):845–855, 2001.

[16] H. Qian. Nonlinear stochastic dynamics of mesoscopic homogeneous biochemical reaction system–an analytical theory. *Nonlinearity*, 24(6):R19, 2011.

[17] D. Revuz and M. Yor. *Continuous Martingales and Brownian Motion*. Springer, Berlin, 3rd edition, 1999.

[18] M. E. Vares. A note on small random perturbations of dynamical systems. *Stoch. Proc. Appl.*, 35(2):225–230, 1990.

[19] D. V. Widder. *The Laplace Transform*. Princeton University Press, Princeton, 1946.

[20] T. Wilhelm. The smallest chemical reaction system with bistability. *BMC Syst. Biol.*, 3(1):90, 2009.

[21] L. Yan, L. Lu, and Z. Xu. L^p estimates on a time-inhomogeneous diffusion process. *J. Math. Phys.*, 46(8):3513, 2005.

[22] L. Yan and B. Zhu. L^p-estimates on diffusion processes. *J. Math. Anal. Appl.*, 303(2):418–435, 2005.