Rolling with Random Slipping and Twisting: 
A Large Deviation Point of View

Qiao Huang∗, Wei Wei†, Jinqiao Duan‡

Abstract

We revisit the rolling model from the perspective of probability. More precisely, we consider a Riemannian manifold rolling against the Euclidean space, where the rolling is combined with random slipping and twisting. The system is modelled as a stochastic differential equation of Stratonovich-type driven by semimartingales, on the orthonormal frame bundle. We prove the large deviation principles for the projection curves on the base manifold and their horizontal lifts respectively, provided the large deviations hold for the random Euclidean curves as semimartingales. A general large deviation result for the compact manifold and two special results for the noncompact manifold are established.

AMS 2010 Mathematics Subject Classification: 60H10, 60F10, 58J65, 60G44.

Keywords and Phrases: Large deviations, rolling with slipping and twisting, semimartingales, controlled stochastic differential equations.

1 Introduction

Differential geometry has been inextricably related to classical mechanics, since its very conception in the 18th century. A well-known early example of the systems in mechanics is the sphere rolling on the plane without slipping or twisting. The first time the problem of a sphere rolling on the plane was considered as worthy of study was in the seminal papers of S. A. Chaplygin [6, 7], one of the fathers of non-holonomic mechanics. Nowadays the general systems of rolling a manifold are often studied in connection to intrinsic Riemannian geometry, sub-Riemannian geometry [20, 24] and geometric control theory [1]. They are also employed by many application areas, such as robotics and geometric interpolation theory.

A geometric operation, which is known as development, plays a central role in the rolling model. It was originally given by É. Cartan in [5], where he developed a manifold onto a tangent space, in order to define holonomy in terms of Euclidean displacements. The relation between Cartan’s development and rolling model became clear in Nomizu’s breakthrough introduction of the dynamics of rolling in higher dimensions through embedded submanifolds of Euclidean space.

Perhaps something that might have been not expected by Cartan is that this concept of development would play a fundamental role in the definition of Brownian motion on a manifold, and the subsequent explosion of interest that stochastic analysis in Riemannian manifolds has had in later decades, see [13]. For a long time, mathematicians have had the intuition that by rolling a d-dimensional Riemannian manifold M along a given curve γ = {γt}0≤t≤T in Rd with the Euclidean

∗Center for Mathematical Sciences, Huazhong University of Science and Technology, Wuhan, Hubei 430074, P.R. China. Email: hq932309@alumni.hust.edu.cn
†Center for Mathematical Sciences, Huazhong University of Science and Technology, Wuhan, Hubei 430074, P.R. China. Email: weiw16@hust.edu.cn
‡Department of Applied Mathematics, Illinois Institute of Technology, Chicago, IL 60616, USA. Email: duan@iit.edu
structure, one would obtain a curve in $M$ which resembles the original curve $\gamma$, see [19]. The main outstanding idea (as far as we know due to Malliavin) was to use Cartan’s development through the orthonormal frame bundle and Wiener’s measure, see [25].

The idea of how to define Brownian trajectories on manifolds is similar to the procedure of rolling sphere. Intuitively, one can draw a Brownian path $B_t$ in $\mathbb{R}^d$, and then one can consider the system of $M$ rolling against $\mathbb{R}^d$ following the path $B_t$ (see Fig. 1 for a visualization). The precise definition uses a less regular version of Cartan’s development and parallel transport. This naive notion allows one to recover the Laplace-Beltrami operator $\Delta_M$ of the manifold. It is often interpreted as if Brownian paths are the “integral curves” for $\Delta_M$. Of course this assertion lacks of mathematical precision, but it introduces the idea that second order differential operators induce “diffusions” on the manifold. This point of view has been exploited significantly in the study of stochastic differential equations (SDEs) on manifolds, see [4, 14].

![Fig. 1. Brownian motion on the sphere $S^2$ by rolling along a Brownian motion on $\mathbb{R}^2$.](image)

Our aim in the present paper is to model the systems of rolling with random slipping and twisting in higher dimensions and study the stability. If we denote by $H_\xi$ the standard horizontal vector field on the orthonormal frame bundle $OM$ corresponding to $\xi \in \mathbb{R}^d$, then given a smooth curve $\gamma = \{\gamma_t\}_{0 \leq t \leq T}$ in the Euclidean space $\mathbb{R}^d$, the system of rolling along $\gamma$ without slipping or twisting is modelled by Cartan’s development as the following ordinary differential equation (ODE) on $OM$,

$$\dot{u}_t = H_{\dot{\gamma}_t}(u_t).$$

(1.1)

Projecting the solution curve $u = \{u_t\}_{0 \leq t \leq T}$ onto the base manifold $M$, we get a curve $x = \{x_t\}_{0 \leq t \leq T}$, which is exactly the trace left on the manifold $M$ by rolling $M$ along the pathway $\gamma$ on $\mathbb{R}^d$. If $\gamma$ is a straight line on the Euclidean space, then the projection curve $x$ is a geodesic on $M$ starting from $\pi(u_0)$ with initial velocity $u_0(\dot{\gamma}_0)$. Let $D = d(d-1)/2$ and let $\{A_1, \cdots, A_D\}$ be an orthonormal basis of $\mathfrak{so}(d)$, the space of skew symmetric matrices in dimension $d$. For $A \in \mathfrak{so}(d)$, we denote by $A^*$ the fundamental vertical vector field on $OM$ induced by $A$. Let $\{e_1, \cdots, e_d\}$ be the canonical basis of $\mathbb{R}^d$. Denote for shorthand $H_i := H_{e_i}$. If we involve the slipping and twisting ingredients (in a random fashion) in the rolling model, then the differential equation on the bundle $OM$ becomes

$$du_t = H_\gamma(u_t) \circ d\tilde{\gamma}_t^i + A^*_\alpha(u_t) \circ dW_\alpha^i,$$

(1.2)

where $\tilde{\gamma}$ is an equivalent curve that involves the slipping constituent, it is a randomization of the original Euclidean curve $\gamma$. $W = \{W_t\}_{t \geq 0}$ is a $D$-dimensional standard Brownian motion, the symbol $\circ$ means that the stochastic differential is in Stratonovich sense. Here and after, we use
Einstein’s convention that the repeated indices in a product will be summed automatically. The random twisting is indicated in the second term, i.e., the vertical component, of right hand side of the last equation. Some examples of rolling models with specific pattern of slipping or twisting are given in Section 5.

Recently, there are several works concerned with some similar models as (1.1) or (1.2). The paper [18] studied the homogenization a kind of random perturbation to the geodesic equation. The author took $\tilde{\gamma}_t = e_t t$ in (1.2) and add one more deterministic vertical perturbation on the right hand side, the horizontal and deterministic vertical components are in normal scale while the random vertical components are in ‘fast’ scale. In the paper [2], the authors introduced a diffusion process with a finite speed of propagation which they call kinetic Brownian motion. This kind of diffusion process is exactly modelled by (1.2) with $\tilde{\gamma}_t = e_t t$ as well but with the family of Lie elements $\{A_\alpha\}_{\alpha=1}^D$ replaced by $\{\tilde{A}_i\}_{i=2}^d$, $\tilde{A}_i = e_i \otimes e_i^* - e_1 \otimes e_1^*$. That is, the kinetic Brownian motion is the trace curve of rolling along the straight line pointing to the first Euclidean direction with unit speed, coupling with some special twisting. Both papers give some interpolation between geodesic and Brownian motions. In [3], the authors studied the small mass limit of Langevin equations solving an SDE on $\mathbb{R}^d$.

We are concerned with a small perturbation version of (1.2), that is, the system of rolling perturbed by small random slipping and twisting, which is modelled by the following stochastic differential equation (SDE) on $OM$,

$$
du^\epsilon_t = H_i(u^\epsilon_t) \circ d\gamma^\epsilon_t^i + A_\alpha^*(u^\epsilon_t) \circ dW_t^\epsilon,\alpha,$$

where $\epsilon > 0$ is a small parameter, $W^\epsilon := \sqrt{\epsilon}W$, the curve $\gamma^\epsilon$ is a random perturbation of $\gamma$ indicating the slipping, which is independent of $W^\epsilon$. When $\epsilon \to 0$, the process $W^\epsilon$ converges in distribution to 0. If in addition, $\gamma^\epsilon$ converges in distribution to the original curve $\gamma$ as $\epsilon \to 0$, then the classical stability theory of stochastic differential equations tells that the bundle-valued curve $u^\epsilon$ converges in distribution to the deterministic curve $u$ in (1.1) (see [15, Section IX.6]), and then the continuity of the projection of $OM$ to $M$ yields that the trace curve $x^\epsilon$ converges in distribution as well, to the curve $x$ which is the projection of the curve $u$. This means that the stability in distribution sense holds for the rolling systems, in the terminology of dynamics.

A further question problem is how to qualify the path-wise stability or instability? The theory of large deviations comes in handy. Roughly speaking, the theory of large deviations is concerned with the asymptotic estimation of probabilities of rare events. If the large deviation principle (LDP) holds for the trace family $\{x\}_{\epsilon>0}$, then one can say that the probability of rare events that the sample path of trace process $x^\epsilon$ is not close to the limit curve $x$ is exponentially small in $\epsilon$. In this sense, we can say that the rolling system ‘exponentially’ stable. So our question is: if the family of Euclidean curves $\{\gamma^\epsilon\}_{\epsilon>0}$ satisfies the large deviation principle, is it true that the family of bundle-valued curves $\{u^\epsilon\}_{\epsilon>0}$ or manifold-valued curves $\{x^\epsilon\}_{\epsilon>0}$ still satisfies the large deviation principle?

We refer to [9, 10] for more details on the large deviation theory. We would like to point out that the classical large deviation theorems for Brownian motions and random walks was generalized to the setting of complete Riemannian manifolds in [17].

In general, the perturbation curve $\gamma^\epsilon$ is a continuous semimartingale (not necessarily be Markovian) for each $\epsilon > 0$. Therefore, our problem translates to prove that the solution family of the following semimartingale-driven stochastic differential equation satisfies the large deviation principle:

$$
dX^\epsilon_t = F(X^\epsilon_t) dY^\epsilon_t, \quad X^\epsilon_0 = x_0, \quad (1.3)$$
provided that the family of noise \( \{Y^\epsilon\}_{\epsilon > 0} \) satisfies the large deviation principle. A similar problem of large deviations for SDE (1.3) was investigated in [12, 11]. In both papers, the \textit{exponentially tight} assumption on the family \( \{(X^\epsilon, Y^\epsilon)\} \) was proposed to prove the LDP for the solution family \( \{X^\epsilon\} \), provided that the LDP holds for \( \{Y^\epsilon\} \). By the classical large deviation theory, in the presence of exponential tightness of the family \( \{(X^\epsilon, Y^\epsilon)\} \), to prove that this family satisfies the large deviation principle with some rate function, it is enough to assume it holds and then identify the rate function to make sure that this rate function does not depend on the choice of subsequences. The results in those two papers do not completely solve the preceding problem, because the assumption of exponential tightness of \( \{(X^\epsilon, Y^\epsilon)\} \) is rather strong and not easy to verify, and it involves the additional condition on the solutions \( X^\epsilon \). We will reconsider the large deviations of semimartingale-driven SDEs in the presence of controls, with all assumptions made only in terms of the driven noise and controls (see Section 2). The result of LDP for semimartingale-driven SDEs with no assumptions on solutions \( X^\epsilon \) (Proposition 2.6) are new to our knowledge, and are of independent interest.

The main results of this paper are Theorem 3.3, Theorem 4.1 and Theorem 4.2. The first theorem aims to build the large deviations for the rolling problems on compact manifolds, while the other two deal with the case of noncompact manifolds. When the rolled manifold is compact, we can accomplish the task for each Euclidean curve \( \gamma^\epsilon \) being general semimartingale. But in noncompact case, it is only possible for us to treat for some special semimartingales. Theorem 4.1 requires that each Euclidean curve \( \gamma^\epsilon \) is locally of finite variation, and Theorem 4.2 is devoted for each \( \gamma^\epsilon \) satisfying an SDE driven by Brownian noise.

The sequel of this paper is organized as follows. In the next section, we prove the LDP for a class of controlled stochastic differential equations driven by semimartingales. The large deviations for stochastic differential equations of Stratonovich-type are also given there. In Section 3 and 4, we establish various large deviation principles for the rolling systems. Section 3 deals with the rolling on compact manifolds, the random Euclidean curves are general continuous semimartingales. Section 3 is devoted to the noncompact case, where Euclidean curves are assumed to be either locally of finite variation or driven by stochastic differential equations. Finally, Section 5 is reserved for a few typical examples of rolling systems.

## 2 Large deviations in stochastic control

The aim of this section is to prove the large deviation principle for the controlled stochastic differential equations. The reason why the `controls` appear is that the horizontal lifts of the projections of solution processes satisfy a sort of stochastic differential equations with controls, as we will see in Lemma 3.1.

First of all, we recall some definitions in the large deviation theory. Let \( \mathcal{X} \) be a topological space equipped with Borel \( \sigma \)-algebra \( \mathcal{B}(\mathcal{X}) \). A \textit{good rate function} \( I \) is a lower semicontinuous mapping \( I : \mathcal{X} \to [0, \infty) \) such that for all \( \alpha \in [0, \infty) \), the level set \( \Phi_I(\alpha) := \{x \in \mathcal{X} : I(x) \leq \alpha\} \) is a closed, compact subset of \( \mathcal{X} \). A family of probability measure \( \{\mu_\epsilon\}_{\epsilon > 0} \) on \( (\mathcal{X}, \mathcal{B}(\mathcal{X})) \) is said to satisfies the large deviation principle (LDP) with a good rate function \( I \) if, for all \( \Gamma \in \mathcal{B}(\mathcal{X}) \),

\[
- \inf_{x \in \Gamma^o} I(x) = \liminf_{\epsilon \to 0} \epsilon \log \mu_\epsilon(\Gamma) \leq \limsup_{\epsilon \to 0} \epsilon \log \mu_\epsilon(\Gamma) \leq -\inf_{x \in \Gamma} I(x).
\]

A family of \( \mathcal{X} \)-valued random elements \( \{X^\epsilon\}_{\epsilon > 0} \) is said to satisfies the large deviation principle if the family of probability measures induced by each \( X^\epsilon \) on \( \mathcal{X} \) satisfies the large deviation principle. We refer to [9] for more details of the large deviation theory.
2.1 Exponential tightness

A family of probability measure \( \{ \mu_\epsilon \}_{\epsilon > 0} \) on \((X, B(X))\) is said to be exponentially tight if for every \( \alpha < \infty \), there exists a compact set \( K \subset X \) such that

\[
\limsup_{\epsilon \to 0} \epsilon \log \mu_\epsilon(K^c) < -\alpha.
\]

Or equivalently, for every \( 0 < \delta < 1 \), there exists a compact set \( K \subset X \) and \( \epsilon_0 > 0 \), such that for all \( 0 < \epsilon < \epsilon_0 \),

\[
[\mu_\epsilon(K^c)]^\epsilon < \delta.
\]

Similarly, a family of \( X \)-valued random elements \( \{ X_\epsilon \}_{\epsilon > 0} \) is said to exponentially tight if the family of induced probability measures on \( X \) is exponentially tight. If \( X \) has a countable base, from [9, Lemma 4.1.23] we know that an exponentially tight sequence of measures has a subsequence that satisfies the large deviation principle with some good rate function (see [21, Theorem (P)]). In case of \( X = \mathbb{R}^d \), the exponential tightness is equivalent to the exponential stochastic boundness found in [12].

Now denote by \( C := C(\mathbb{R}_+; \mathbb{R}^d) \) the space of all \( \mathbb{R}^d \)-valued continuous functions on \( \mathbb{R}_+ \), equipped with the local uniform topology which is the topology of uniform convergence on compact intervals. It is well known that \( C \) is a Polish space. We associate with \( C \) the Borel \( \sigma \)-algebra \( B(C) \). The exponential tightness of a family of probability measures on \((C, B(C))\) is implied by the LDP with good rate function. For \( \rho > 0 \), \( T > 0 \) and \( x = \{x(t)\}_{t \geq 0} \in C \), define the uniform norm and modulus of continuity

\[
\|x\|_T := \sup_{0 \leq t \leq T} |x(t)|, \quad w_T(x, \rho) := \sup_{0 \leq t, s \leq T} |x(t) - x(s)|.
\]

The following proposition is a criterion for the exponential tightness of probability measures in \( C \), referring to [21, Theorem 4.2].

**Lemma 2.1.** A family of probability measures \( \{ \mu_\epsilon \}_{\epsilon > 0} \) on \( C \) is exponentially tight if and only if

(i) for any \( T > 0 \),

\[
\lim_{\rho \to 0} \limsup_{\epsilon \to 0} \epsilon \log \mu_\epsilon(x \in C : \|x\|_T \geq \rho) = -\infty,
\]

(ii) for any \( T > 0, \eta > 0 \),

\[
\lim_{\rho \to 0} \limsup_{\epsilon \to 0} \epsilon \log \mu_\epsilon(x \in C : w_T(x, \rho) \geq \eta) = -\infty.
\]

2.2 Large deviations

For each \( \epsilon > 0 \), we have a filtered probability space \((\Omega^\epsilon, \mathcal{F}^\epsilon, \{\mathcal{F}_t^\epsilon\}_{t \geq 0}, \mathbf{P}^\epsilon)\) endowed with an \( n \)-dimensional continuous semimartingale \( Y^\epsilon \) and an \( m \)-dimensional continuous adapted process \( U^\epsilon \). We also have a global Lipschitz function \( F : \mathbb{R}^d \times \mathbb{R}^m \to \mathbb{R}^{d \times n} \), so that each controlled stochastic differential equation

\[
dX_t^\epsilon = F(X_t^\epsilon, U_t^\epsilon) dY_t^\epsilon, \quad X_0^\epsilon = x_0,
\]

has a unique global solution \( X^\epsilon \) which is a \( d \)-dimensional process. Suppose the Lipschitz constant of the function \( F \) is \( C > 0 \), that is, \( |F(x_1, u_1) - F(x_2, u_2)| \leq C(|x_1 - x_2| + |u_1 - u_2|) \) for all \( x_1, x_2, u_1, u_2 \). It is obvious that \( F \) has linear growth, that is, \( |F(x, u)| \leq C(1 + |x| + |u|) \) for all \( x, u \).
We assume for each $\epsilon > 0$, the canonical decomposition of the continuous semimartingale $Y^\epsilon$ is

\[ Y^\epsilon = M^\epsilon + A^\epsilon, \]

where $M^\epsilon$ is a continuous local martingale, and $A^\epsilon$ is a continuous predictable process with locally finite variation. We associate with $Y^\epsilon$ an increasing process $G(Y^\epsilon)$ via

\[ G(Y^\epsilon)_t := |V(A^\epsilon)|_t + \frac{1}{\epsilon}(|M^\epsilon, M^\epsilon|)_t. \quad (2.1) \]

The following lemma is a small modification of [12, Theorem 1.2, Lemma 2.5], which will be useful in proving the LDP results.

**Lemma 2.2.** Suppose the family $\{G(Y^\epsilon)_t\}_{\epsilon > 0}$ is exponentially tight for each $t > 0$. Then

(i). The family $\{Y^\epsilon\}_{\epsilon > 0}$ is uniformly exponentially tight.

(ii). Let $\{Z^\epsilon\}_{\epsilon > 0}$ be a family of continuous $\{H^\epsilon_t\}_{t \geq 0}$-adapted processes. If the family $\{(Y^\epsilon, U^\epsilon, Z^\epsilon)\}_{\epsilon > 0}$ satisfies the LDP with good rate function $I^\epsilon$, then the family $\{(Z^\epsilon \cdot Y^\epsilon, U^\epsilon, Z^\epsilon)\}_{\epsilon > 0}$ also satisfies the LDP with the following good rate function:

\[ I(w, y, u, z) = \begin{cases} I^\epsilon(y, u, z), & w = z \cdot y, \text{and } y \text{ is locally of finite variation,} \\ \infty, & \text{otherwise.} \end{cases} \]

**Proof.** The first assertion is taken from [12, Lemma 2.5]. For the second assertion, we observe that $Z^\epsilon \cdot Y^\epsilon = (U^\epsilon, Z^\epsilon) \cdot (0, Y^\epsilon)^T$. Then the result is a corollary of [12, Theorem 1.2].

**Lemma 2.3.** Suppose the family of increasing processes $\{G(Y^\epsilon)\}_{\epsilon > 0}$ is uniformly bounded by a constant $K > 0$. Let $F^\epsilon$ be an $\{F^\epsilon_t\}$-adapted processes, and let

\[ dZ^\epsilon_t = F^\epsilon_t dY^\epsilon_t, \quad Z^\epsilon_0 = z_0. \]

Let $T^\epsilon$ be an $\{F^\epsilon_t\}$-stopping time. Suppose there exist positive constants $C$ and $\rho$, such that for any $t \in [0, T^\epsilon]$,

\[ |F^\epsilon_t| \leq C(\rho^2 + |Z^\epsilon_t|^2)^{1/2}. \quad (2.2) \]

Then for any $a > 0$ and $0 < \epsilon \leq 1$,

\[ \epsilon \log P^\epsilon \left( \sup_{t \in [0, T^\epsilon]} |Z^\epsilon_t| \geq a \right) \leq K' + \log \left( \frac{\rho^2 + |z^2_0|}{\rho^2 + a^2} \right), \]

where $K' = (2C + (2 + n)C^2)K$.

**Proof.** Define $\phi'(z) = (\rho^2 + |z|^2)^{1/\epsilon}$ for each $\epsilon > 0$. Then

\[ \partial_\epsilon \phi'(z) = \frac{2\phi'(z)}{\epsilon(\rho^2 + |z|^2)} z_i, \quad \partial_i \partial_\epsilon \phi'(z) = \frac{2\phi'(z)}{\epsilon(\rho^2 + |z|^2)} \left( \delta_{ij} + 2 \left( \frac{1}{\epsilon} - 1 \right) \frac{z_i z_j}{\rho^2 + |z|^2} \right). \quad (2.3) \]

Let $\Phi^\epsilon := \phi'(Z^\epsilon)$. By Itô’s formula (see, e.g., [15, Theorem I.4.57]),

\[ \Phi^\epsilon = \phi'(z_0) + \partial_\epsilon \phi'(Z^\epsilon) F^\epsilon_{ij} \cdot B^\epsilon_{ij} + \partial_i \phi'(Z^\epsilon) F^\epsilon_{ij} \cdot M^\epsilon_{ij} \left( \frac{1}{2} \partial_i \partial_j \phi'(Z^\epsilon) F^\epsilon_{ik} F^\epsilon_{lj} \cdot (M^\epsilon_{ik}, M^\epsilon_{lj}) \right) \]

Define a stopping time $T^{\epsilon, a} := \inf\{t \geq 0 : |Z^\epsilon_t| \geq a \} \wedge T^\epsilon$. Using the bound (2.2), for each $\epsilon > 0$,

\[ E^\epsilon \left( \int_0^{T^{\epsilon, a}} |D\phi'(Z^\epsilon)|^2 d\langle M^\epsilon, M^\epsilon \rangle_t \right) < \infty. \]
Then the stochastic integrals $\partial_t \phi^\varepsilon(Z^\varepsilon)F_j^\varepsilon \cdot M$ is martingale up to $T$, which yields

$$
E' (\Phi_{t,T}^\varepsilon, a) = \phi^\varepsilon(z_0) + E' \left( \int_0^{t \wedge T} \partial_t \phi(Z^\varepsilon) F_j^\varepsilon dB_j^\varepsilon \right) + \frac{1}{2} E' \left( \int_0^{t \wedge T} \partial_t \partial_j \phi^\varepsilon(Z^\varepsilon) F_k^\varepsilon F_j^\varepsilon d(M^\varepsilon, M^\varepsilon, l) \right) =: \phi^\varepsilon(z_0) + I_1 + I_2.
$$

By (2.2) and (2.3), it is easy to get

$$
|I_1| \leq \frac{2C}{\varepsilon} E' \int_0^{t \wedge T} \Phi^\varepsilon d\langle A^\varepsilon \rangle, \quad |I_2| \leq \frac{(2 + n)C^2}{\varepsilon^2} E' \int_0^{t \wedge T} \Phi^\varepsilon d\langle M^\varepsilon, M^\varepsilon \rangle.
$$

Then

$$
E' (\Phi_{t,T}^\varepsilon, a) \leq \phi^\varepsilon(z_0) + (2C + (2 + n)C^2) \frac{1}{\varepsilon} E' \int_0^{t \wedge T} \Phi^\varepsilon dG_s^\varepsilon.
$$

By virtue of the Gronwall-type inequality in [15, Lemma IX.6.3], we have

$$
E' (\Phi_{T}^\varepsilon, a) \leq \phi^\varepsilon(z_0) e^{K'/\varepsilon}.
$$

Then by Chebycheff’s inequality,

$$
\varepsilon \log P \left( \sup_{0 \leq t \leq T} |Z^\varepsilon_t| \geq a \right) = \varepsilon \log P (|Z^\varepsilon_{T} | \geq a) = \varepsilon \log P (\Phi_{T}^\varepsilon, a \geq \phi(a)) \leq K' + \varepsilon \log \phi^\varepsilon(z_0) - \varepsilon \log \phi^\varepsilon(a) = K' + \log \left( \frac{\rho^2 + |z_0|^2}{\rho^2 + a^2} \right).
$$

The proof is completed.

For any $t > 0$ and $m \in \mathbb{N}_+$, we denote $[t]_m := [t] + \frac{[m(t - [t])]}{m}$, then obviously $|t - [t]_m| < \frac{1}{m}$. For each $m \in \mathbb{N}_+$, define $X^\varepsilon, m$ to be the solution of the following SDE:

$$
dX^\varepsilon, m_t = F(X^\varepsilon, m, [t]_m, U^\varepsilon, m) dY^\varepsilon_t, \quad X^\varepsilon, m_0 = x_0. \quad (2.4)
$$

The following lemma shows that $X^\varepsilon, m$ are exponentially good approximations of $X^\varepsilon$.

**Lemma 2.4.** Suppose the family of increasing processes $\{G(Y^\varepsilon)\}_{\varepsilon > 0}$ is uniformly bounded by a constant $K > 0$. Assume the function $F$ to be bounded and global Lipschitz. Then for any $T > 0$ and $a > 0$,

$$
\lim_{m \to \infty} \limsup_{\varepsilon \to 0} \varepsilon \log P (\|X^\varepsilon, m - X^\varepsilon\|_T > a) = -\infty.
$$

**Proof.** Fix $m \in \mathbb{N}_+$. Let $Z^\varepsilon := X^\varepsilon, m - X^\varepsilon$. Then

$$
dZ^\varepsilon_t = \left( F(X^\varepsilon, m, [t]_m, U^\varepsilon, m) - F(X^\varepsilon, U^\varepsilon) \right) dY^\varepsilon_t, \quad Z^\varepsilon_0 = 0.
$$

For $\rho > 0$, we define an $\{F^\varepsilon_t\}$-stopping time

$$
T^\varepsilon, m, \rho := \inf \{ t \geq 0 : |X^\varepsilon, m_t - X^\varepsilon, [t]_m| + |U^\varepsilon_t - U^\varepsilon, [t]_m| \geq \rho \}.
$$
Then for \( t \in [0, T^\varepsilon,m,\rho] \), we have
\[
|F(X_{t|m}^{\varepsilon,m}, U_{t|m}^{\varepsilon}) - F(X_t^\varepsilon, U_t^\varepsilon)| \leq |F(X_{t|m}^{\varepsilon,m}, U_{t|m}^{\varepsilon})| + |F(X_t^\varepsilon, U_t^\varepsilon) - F(X_t^\varepsilon, U_t^\varepsilon)| \\
\leq C(\rho + |X_t^{\varepsilon,m} - X_t^\varepsilon|).
\]
Together with Lemma 2.3, it yields that for any \( a > 0 \) and \( 0 < \varepsilon \leq 1 \),
\[
\epsilon \log \mathbf{P}^\varepsilon \left( \sup_{t \in [0, T^\varepsilon,m,\rho]} |X_t^{\varepsilon,m} - X_t^\varepsilon| \geq a \right) \leq K' + \log \left( \frac{\rho^2}{\rho^2 + a^2} \right).
\]
Hence,
\[
\lim_{\rho \to 0} \sup_{m \in \mathbb{N}_+} \epsilon \log \mathbf{P}^\varepsilon \left( \sup_{t \in [0, T^\varepsilon,m,\rho]} |X_t^{\varepsilon,m} - X_t^\varepsilon| \geq a \right) = 0. \tag{2.5}
\]
On the other hand, since
\[
|X_t^{\varepsilon,m} - X_t^\varepsilon| = |F(X_{t|m}^{\varepsilon,m}, U_{t|m}^{\varepsilon})(Y_t^\varepsilon - Y_{t|m}^\varepsilon)| \leq Cw_T(Y^\varepsilon, 1/m).
\]
we have
\[
\mathbf{P}^\varepsilon \{ T^\varepsilon,m,\rho < T \} = \mathbf{P}^\varepsilon \left( \sup_{0 \leq t \leq T} \left( |X_t^{\varepsilon,m} - X_t^\varepsilon| + |U_t^\varepsilon - U_{t|m}^\varepsilon| \right) \geq \rho \right) \\
\leq \mathbf{P}^\varepsilon \left( \sup_{0 \leq t \leq T} |X_t^{\varepsilon,m} - X_t^\varepsilon| \geq \frac{\rho}{2} \right) + \mathbf{P}^\varepsilon \left( \sup_{0 \leq t \leq T} |U_t^\varepsilon - U_{t|m}^\varepsilon| \geq \frac{\rho}{2} \right) \\
\leq \mathbf{P}^\varepsilon \left( w_T(Y^\varepsilon, 1/m) \geq \rho/2 \right) + \mathbf{P}^\varepsilon \left( w_T(U^\varepsilon, 1/m) \geq \rho/2 \right).
\]
By virtue of the exponential tightness of \((Y^\varepsilon, U^\varepsilon)\) and Lemma 2.1, for any \( T > 0 \) and \( \rho > 0 \)
\[
\lim_{m \to \infty} \limsup_{\varepsilon \to 0} \varepsilon \log \mathbf{P}^\varepsilon \{ T^\varepsilon,m,\rho < T \} = -\infty. \tag{2.6}
\]
Note that for any \( T > 0 \),
\[
\{ \|X^{\varepsilon,m} - X^{\varepsilon}\|_T > a \} \subset \left\{ \sup_{t \in [0,T^\varepsilon,m,\rho]} |X_t^{\varepsilon,m} - X_t^\varepsilon| \geq a \right\} \cup \{ T^\varepsilon,m,\rho < T \}.
\]
Combining this with (2.5) and (2.6), we obtain the desired result. \( \square \)

\textbf{Corollary 2.5.} \textit{Under the assumptions in Lemma 2.4, the family \( \{ X^\varepsilon \}_{\varepsilon > 0} \) is exponentially tight.}

\textit{Proof.} We first show that for each \( m \in \mathbb{N}_+ \), the family \( \{ X^{\varepsilon,m} \}_{\varepsilon > 0} \) defined in (2.4) obeys the LDP with a good rate function. To this end, we define for each \( m \) a map \( \Psi^m : C \to C \) via \( f = \Phi^m(g, h) \), where \( f(0) = 0 \) and for \( t \in (N + \frac{k}{m}, N + \frac{k+1}{m}] \), \( N \in \mathbb{N}, k = 0, 1, \ldots, m-1 \),
\[
f(t) = f(N + \frac{k}{m}) + F(f(N + \frac{k}{m}), h(N + \frac{k}{m})) (g(t) - g(N + \frac{k}{m})).
\]
Then it is easy to see \( X^{\varepsilon,m} = \Psi^m(Y^\varepsilon, U^\varepsilon) \). Let \( f_i = \Psi^m(g_i, h_i) \) for \( g_i, h_i \in C, i = 1, 2 \). Let \( e = f_1 - f_2 \). Then for fixed \( N \in \mathbb{N} \),
\[
\sup_{t \in (N + \frac{k}{m}, N + \frac{k+1}{m}]} e(t) \leq 2C(\|g_1\|_{N+1} + 1) (e(N + \frac{k}{m}) + \|g_1 - g_2\|_{N+1} + \|h_1 - h_2\|_{N+1}).
\]
Since \( e(0) = 0 \), by iterating this bound over \( k = 0, 1, \cdots, m - 1 \) and \( N \), we have
\[
\|e(t)\|_{N+1} \leq C(\|g_1\|_{N+1} N, \|g_1 - g_2\|_{N+1} + \|h_1 - h_2\|_{N+1}).
\]
Hence \( \Phi^m \) is continuous from \( C([0, N+1]; \mathbb{R}^d) \) to \( C([0, N+1]; \mathbb{R}^d) \) for each \( N \in \mathbb{N} \), and consequently continuous from \( C \) to \( C \) (cf. [8, Proposition 5.6, 5.7]). The contraction principle (see, e.g., [9, Theorem 4.2.1]) yields that the family \( \{X^m\} \) obeys the LDP with a good rate function, and thus is exponentially tight for each \( m \).

Fix \( T > 0 \). Observe that for all \( a > 0 \), \( \eta > 0 \) and \( \rho > 0 \)
\[
\{\|X^\epsilon\|_T \geq a \} \subset \{\|X^\epsilon - X^{e,m}\|_T \geq \frac{a}{2} \} \cup \{\|X^{e,m}\|_T \geq \frac{a}{2} \},
\]
\[
\{w_T(X^\epsilon, \rho) \geq \eta \} \subset \{\|X^\epsilon - X^{e,m}\|_T \geq \frac{\eta}{2} \} \cup \{w_T(X^\epsilon, \rho) \geq \frac{\eta}{2} \}.
\]
Then the exponential tightness of the family \( \{X^\epsilon\} \) follows from that of \( \{X^{e,m}\} \), Lemma 2.4 and Lemma 2.1.

**Proposition 2.6.** Assume the family \( \{G(Y^\epsilon)_t\}_{\epsilon>0} \) defined in (2.1) is exponentially tight for each \( t > 0 \), and the family \( \{(Y^\epsilon, U^\epsilon)\}_{\epsilon>0} \) satisfies the LDP with good rate function \( I' \). Suppose the function \( F \) to be bounded and global Lipschitz. Then
(i). the family \( \{(X^\epsilon, Y^\epsilon, U^\epsilon)\}_{\epsilon>0} \) satisfies the LDP with a good rate function \( I \),
(ii). the rate function \( I \) is given by
\[
I(x,y,u) = \begin{cases} I'(y,u), & x = F(x,u) \cdot y, y \text{ is locally of finite variation}, \\
\infty, & \text{otherwise}. \end{cases} \tag{2.7}
\]
In particular, the family \( \{X^\epsilon\}_{\epsilon>0} \) satisfies the LDP with following good rate function:
\[
\hat{I}(x) = \inf \{I'(y,u) : x = F(x,u) \cdot y, y \text{ is locally of finite variation} \}.
\]

**Proof.** Step 1 (Identification of the rate function). Suppose the family \( \{(X^\epsilon, Y^\epsilon, U^\epsilon)\}_{\epsilon>0} \) is exponentially tight. We show that for any subsequence \( \{(X^{\epsilon_k}, Y^{\epsilon_k}, U^{\epsilon_k})\}_{k=1}^\infty \), with \( \epsilon_k \to 0 \) as \( k \to \infty \), which obeys the LDP, the rate function \( I \) is given by (2.7). For notational simplicity, we still denote the subsequence \( \epsilon_k \) by \( \epsilon \).

We follow the lines of [11, Theorem 6.1]. By the contraction principle, the family \( (Y^\epsilon, U^\epsilon, F(X^\epsilon, U^\epsilon)) \) obeys the LDP with good rate function \( \hat{I}(y,u,f) = \inf \{I(x,y,u) : f = F(x,u) \} \). Since \( X^\epsilon = F(X^\epsilon, U^\epsilon) \cdot Y^\epsilon \), Lemma 2.2 yields that the family \( \{X^\epsilon, Y^\epsilon, U^\epsilon, F(X^\epsilon, U^\epsilon)\} \) obeys the LDP with good rate function
\[
J(x,y,u,f) = \begin{cases} \hat{I}(y,u,f), & x = f \cdot y \text{ and } y \text{ is locally of finite variation}, \\
\infty, & \text{otherwise}. \end{cases}
\]
\[
= \begin{cases} \inf \{I(x',y,u) : f = F(x',u)\}, & x = f \cdot y \text{ and } y \text{ is locally of finite variation}, \\
\infty, & \text{otherwise}. \end{cases} \tag{2.8}
\]
But the contraction principle yields \( I(x,y,u) = \inf_f \{J(x,y,u,f)\} \). Hence, if \( y \) is of infinite variation, then by (2.8), \( J(x,y,u,f) = \infty \) and \( I(x,y,u) = \infty \).

On the other hand, using contraction principle once again, the rate function \( J \) is
\[
J(x,y,u,f) = \begin{cases} I(x,y,u), & f = F(x,u), \\
\infty, & \text{otherwise}. \end{cases} \tag{2.9}
\]
Suppose $y$ is locally of finite variation but $x \neq F(x, u) \cdot y$, we will prove that $J(x, y, u, f) = \infty$ and then $I(x, y, u) = \infty$. If $x \neq f \cdot y$, then (2.8) yields the desired results. If $x = f \cdot y$, then $f \neq F(x, u)$, and the results follow from (2.9).

Suppose now $I(x, y, u) < \infty$. Then the previous arguments yield that $y$ is locally of finite variation and $x = F(x, u) \cdot y$. Again by the contraction principle, $I'(y, u) = \inf_{x'} \{I(x', y, u)\}$, and obviously $I'(y, u) \leq I(x, y, u)$. If $I'(y, u) < I(x, y, u)$, then there exists $x'$ such that $I(x', y, u) < I(x, y, u) < \infty$. Hence, $x' = F(x', u) \cdot y$, which yields $x = x'$ by the uniqueness. Therefore, we have $I'(y, u) = I(x, y, u)$ in this case. The representation (2.7) follows.

In the rest steps, we only need to show the exponential tightness of $\{(X^\epsilon, Y^\epsilon, U^\epsilon)\}_{\epsilon > 0}$, provided the exponential tightness of $\{(Y^\epsilon, U^\epsilon)\}_{\epsilon > 0}$ and $\{G(Y^\epsilon)\}_{\epsilon > 0}$ for each $t$.

Step 2 (Localization). Suppose $\{(X^\epsilon, Y^\epsilon, U^\epsilon)\}$ is exponentially tight if in addition, the following conditions are satisfied: the family of increasing processes $\{G(Y^\epsilon)\}$ are uniformly bounded. We deduce it holds in general.

For each $\epsilon, p > 0$, define a stopping time
\[
T^{\epsilon, p} := \inf\{t \geq 1 : G(Y^\epsilon)_t \geq p\}.
\]
Then $T^{\epsilon, p}$ is nondecreasing in $p$. Let $U^{\epsilon, p} := U^{\epsilon}_{\Lambda_{T^{\epsilon, p}}}$, $Y^{\epsilon, p} := Y^{\epsilon}_{\Lambda_{T^{\epsilon, p}}}$, and each $X^{\epsilon, p}$ be the solution of the SDE
\[
dX^{\epsilon, p}_t = F(X^{\epsilon, p}_t, U^{\epsilon, p}_t) dY^{\epsilon, p}_t, \quad X^{\epsilon, p}_0 = x_0, \tag{2.10}
\]
Then $X^{\epsilon, p} = X^{\epsilon}_{\Lambda_{T^{\epsilon, p}}}$. It is obvious that (cf. [15, Eq. (VI.1.9)])
\[
\sup_{t \leq T} (|U^{\epsilon, p}_t| + |Y^{\epsilon, p}_t|) = \sup_{t \leq T} (|U^{\epsilon}_t| + |Y^{\epsilon}_t|) \leq \sup_{t \leq T} (|U^{\epsilon}_t| + |Y^{\epsilon}_t|),
\]
\[
w_T(U^{\epsilon, p}, Y^{\epsilon, p}, X^{\epsilon, p}, \rho) = w_T(U^{\epsilon}, Y^{\epsilon}, X^\epsilon, \rho) \leq w_T(U^{\epsilon}, Y^{\epsilon}, X^\epsilon, \rho).
\]

Lemma 2.1 yields that $\{(X^{\epsilon, p}, U^{\epsilon, p})\}$ is also exponentially tight, for each $p > 0$. Note that the increasing process $G(Y^{\epsilon, p})$ associated to each $Y^{\epsilon, p}$ is exactly $G(Y^{\epsilon})_{\Lambda_{T^{\epsilon, p}}}$. Hence, $G(Y^{\epsilon, p})$ is uniformly bounded by $p$, for each $p > 0$, and $G(Y^{\epsilon, p}) \leq G(Y^{\epsilon})$, which yields that the family $\{G(Y^{\epsilon, p})\}_{\epsilon > 0}$ is exponentially tight for each $t > 0$ and $p > 0$. Then our assumption yields each family $\{(X^{\epsilon, p}, Y^{\epsilon, p}, U^{\epsilon, p})\}_{\epsilon > 0}$ is exponentially tight for every $p > 0$. Using Lemma 2.1 again, for each $p > 0$ and all $T > 0$, $\eta > 0$, and any $\delta > 0$, there exist $a > 0$, $\rho > 0$ and $\epsilon_0 > 0$, such that for all $0 < \epsilon \leq \epsilon_0$,
\[
\mathbb{P}^{\epsilon} \left( \sup_{t \leq T} (|U^{\epsilon, p}_t| + |Y^{\epsilon, p}_t| + |X^{\epsilon, p}_t|) \geq a \right) < \delta^{1/\epsilon}, \tag{2.11}
\]
\[
\mathbb{P}^{\epsilon} \left( w_T((U^{\epsilon, p}, Y^{\epsilon, p}, X^{\epsilon, p}), \rho) \geq \eta \right) < \delta^{1/\epsilon}, \tag{2.12}
\]

Note that for each $\epsilon, p, T > 0$,
\[
\{T^{\epsilon, p} > T\} \subset \{(U^{\epsilon, p}_t, Y^{\epsilon, p}_t, X^{\epsilon, p}_t) = (U^{\epsilon}_t, Y^{\epsilon}_t, X^\epsilon_t), \forall t \in [0, T]\}. \tag{2.13}
\]

Fix $T > 0$. Using the exponential tightness of $\{G(Y^{\epsilon})_T\}$, there exist $p_0 > 0$ and $\epsilon_0 > 0$, such that for all $0 < \epsilon \leq \epsilon_0$,
\[
\mathbb{P}^{\epsilon}(T^{\epsilon, p_0} \leq T) = \mathbb{P}^{\epsilon}(G(Y^{\epsilon})_T \geq p_0) < \delta^{1/\epsilon}. \tag{2.14}
\]
Hence, combining (2.11), (2.12), (2.13) and (2.14), for all $0 < \epsilon \leq \epsilon_0 \wedge \epsilon_0' \wedge 1$,

$$
P^\epsilon \left( \sup_{t \leq T} (|U_t^\epsilon| + |Y_t^\epsilon| + |X_t^\epsilon|) \geq a \right)
= P^\epsilon \left( \sup_{t \leq T} (|U_t^\epsilon| + |Y_t^\epsilon| + |X_t^\epsilon|) \geq a, T^{\epsilon, p_0} > T \right)
+ P^\epsilon \left( \sup_{t \leq T} (|U_t^\epsilon| + |Y_t^\epsilon| + |X_t^\epsilon|) \geq a, T^{\epsilon, p_0} \leq T \right)
\leq P^\epsilon \left( \sup_{t \leq T} (|U_t^{\epsilon, p_0}| + |Y_t^{\epsilon, p_0}| + |X_t^{\epsilon, p_0}|) \geq a \right)
+ P^\epsilon (T^{\epsilon, p_0} \leq T)
< 2\delta^{1/\epsilon},
$$

and

$$
P^\epsilon (w_T((U^\epsilon, Y^\epsilon, X^\epsilon), \rho) \geq \eta) \leq P^\epsilon (w_T(U^{\epsilon, p_0}, Y^{\epsilon, p_0}, X^{\epsilon, p_0}, \rho) \geq \eta) + P^\epsilon (T^{\epsilon, p_0} \leq T) < 2\delta^{1/\epsilon}.
$$

Once again, Lemma 2.1 tells that the family $\{(X^\epsilon, Y^\epsilon, U^\epsilon)\}$ is exponentially tight.

**Step 3 (Exponential tightness of $\{(X^\epsilon, Y^\epsilon, U^\epsilon)\}$).** In this step, we will assume the family of processes $\{G(Y^\epsilon)\}$ to be uniformly bounded by a constant $K > 0$. By Corollary 2.5, the family $\{X^\epsilon\}$ is exponentially tight. Observe that for all $a > 0$, $\eta > 0$ and $\rho > 0$,

$$
\left\{ \|X^\epsilon, Y^\epsilon, U^\epsilon\|_T \geq a \right\} \subset \left\{ \|X^\epsilon\|_T \geq \frac{a}{2} \right\} \cup \left\{ \|Y^\epsilon, U^\epsilon\|_T \geq \frac{a}{2} \right\},
\left\{ w_T(X^\epsilon, Y^\epsilon, U^\epsilon), \rho \geq \eta \right\} \subset \left\{ w_T(X^\epsilon, \rho) \geq \frac{\eta}{2} \right\} \cup \left\{ w_T(Y^\epsilon, U^\epsilon), \rho \geq \frac{\eta}{2} \right\}.
$$

The exponential tightness of $\{(X^\epsilon, Y^\epsilon, U^\epsilon)\}$ follows from that of $\{X^\epsilon\}$ and $\{(Y^\epsilon, U^\epsilon)\}$, together with Lemma 2.1.

**Corollary 2.7.** Let the assumptions in Proposition 2.6 hold, except that $F$ is only locally Lipschitz. Assume in addition, for every $t > 0$, the family $\{\sup_{0 \leq s \leq t} |X_s^\epsilon|\}_{\epsilon > 0}$ is exponentially tight. Then the same results as in Proposition 2.6 hold.

**Proof.** Since we have already identified the rate function in Step 1 of the proof of Proposition 2.6, it is only needed to show the family $\{(X^\epsilon, Y^\epsilon, U^\epsilon)\}$ is exponentially tight. Suppose first that the families $\{X^\epsilon\}$ and $\{U^\epsilon\}$ are uniformly bounded by a constant $K > 0$. Let $f : \mathbb{R}^d \times \mathbb{R}^m \rightarrow \mathbb{R}^d \times \mathbb{R}^m$ be a $C^1$ bounded function with $f(x, u) = (x, u)$ when $|x| \leq K$ and $|u| \leq K$. Define $\tilde{F}(x, u) := F(f(x, u))$. Then obviously $\tilde{F}$ is global Lipschitz and each $X^\epsilon$ is the solution of the following SDE:

$$
dX_t^\epsilon = \tilde{F}(X_t^\epsilon, U_t^\epsilon)dY_t^\epsilon, \quad X_0^\epsilon = x_0.
$$

Hence, Proposition 2.6 yields that $\{(X^\epsilon, Y^\epsilon, U^\epsilon)\}$ is exponential tight. Now by virtue of the assumption and the exponential tightness of $\{U^\epsilon\}$, the family $\{\sup_{0 \leq s \leq t} (|X_s^\epsilon| + |U_s^\epsilon|)\}$ is exponential tight for every $t > 0$. The general case follows from a similar localization argument in Step 2 of the proof of Proposition 2.6, with $\sup_{0 \leq s \leq t} (|X_s^\epsilon| + |U_s^\epsilon|)$ in place of $G(Y^\epsilon)_t$.

Now we consider the following SDE of Stratonovich-type:

$$
dX_t^\epsilon = F(X_t^\epsilon, U_t^\epsilon) \circ dY_t^\epsilon, \quad X_0^\epsilon = x_0.
$$

**Corollary 2.8.** Assume the family $\{G(Y^\epsilon)_t\}_{\epsilon > 0}$ is exponentially tight for each $t > 0$, and the family $\{(Y^\epsilon, (M^\epsilon, M^\epsilon'), U^\epsilon)\}_{\epsilon > 0}$ satisfies the LDP with good rate function $I''$. Suppose that the function $F$ is bounded, global Lipschitz and $\nabla F$ is global Lipschitz. Then
(i). the family \( \{ (X^\epsilon, Y^\epsilon, U^\epsilon) \}_{\epsilon > 0} \) satisfies the LDP with a good rate function \( I \),
(ii). the rate function \( I \) is given by
\[
I(x, y, u) = \begin{cases}
\inf \{ I''(y, q, u) : x = F(x, u) \cdot y + \frac{1}{2}(F^j \partial_{x^j} F)(x, u) \cdot q, q \text{ is locally of finite variation} \}, & y \text{ is locally of finite variation}, \\
\infty, & \text{otherwise}.
\end{cases}
\]

In particular, the family \( \{ X^\epsilon \}_{\epsilon > 0} \) satisfies the LDP with following good rate function:
\[
\tilde{I}(x, y, q, u) = \inf \{ I''(y, q, u) : x = F(x, u) \cdot y + \frac{1}{2}(F^j \partial_{x^j} F)(x, u) \cdot q, y \text{ and } q \text{ are locally of finite variation} \}.
\]

Proof. We can rewrite the Stratonovich-type SDE as
\[
d\left( \begin{array}{c}
X^\epsilon_t \\
U^\epsilon_t
\end{array} \right) = \left( \begin{array}{cc}
F(X^\epsilon_t, U^\epsilon_t) & 0 \\
0 & I_m
\end{array} \right) \circ d\left( \begin{array}{c}
Y^\epsilon_t \\
U^\epsilon_t
\end{array} \right).
\]

Then using the decomposition \( Y^\epsilon = M^\epsilon + A^\epsilon \), we have
\[
dx^\epsilon_t = F^j(X^\epsilon_t, U^\epsilon_t)dM^\epsilon_t + F^j(X^\epsilon_t, U^\epsilon_t)dB^\epsilon_t + \frac{1}{2}(F^j \partial_{x^j} F)(X^\epsilon_t, U^\epsilon_t)d\langle M^\epsilon, k \rangle_t,
\]
that is
\[
dx^\epsilon_t = \left( \begin{array}{cc}
F(X^\epsilon_t, U^\epsilon_t) & 0 \\
0 & \frac{1}{2}(F^j \partial_{x^j} F)(X^\epsilon_t, U^\epsilon_t)
\end{array} \right) d\left( \begin{array}{c}
Y^\epsilon_t \\
U^\epsilon_t
\end{array} \right),
\]
where the coefficients is global Lipschitz by assumptions. Using the exponential tightness of \( \{(M^\epsilon, M^\epsilon)\} \) and Lemma 2.1.(i), the family of increasing processes \( \{(M^\epsilon, M^\epsilon)\}_{\epsilon > 0} \) is exponentially tight for each \( t > 0 \). Thus, we can see that Proposition 2.6 is applicable here, with \( (Y^\epsilon, (M^\epsilon, M^\epsilon)) \) in place of \( Y^\epsilon \) and \( |V(A^\epsilon)| + \frac{1}{2}|(M^\epsilon, M^\epsilon)| \) in place of \( G(Y^\epsilon) \). The family \( \{(X^\epsilon, Y^\epsilon, (M^\epsilon, M^\epsilon), U^\epsilon)\}_{\epsilon > 0} \) satisfies the LDP with following good rate function
\[
\tilde{I}(x, y, q, u) = \begin{cases}
I''(y, q, u), & x = F(x, u) \cdot y + \frac{1}{2}(F^j \partial_{x^j} F)(x, u) \cdot q, y \text{ and } q \text{ are locally of finite variation}, \\
\infty, & \text{otherwise}.
\end{cases}
\]

The results follow from the contraction principle.

Remark 2.9. The exponential tightness of \( \{(M^\epsilon, M^\epsilon)\}_{\epsilon > 0} \) also follows from that of \( \frac{1}{\epsilon^2} (M^\epsilon, M^\epsilon)_{\epsilon > 0} \).

3 Rolling problems on compact manifolds

In this and next sections, we will show the large deviations for various rolling systems. The manifold rolled against the Euclidean is assumed to be compact in this section. The random Euclidean curves that the manifold rolls along are general continuous semimartingales.

3.1 Preliminaries

Let \((M, g)\) be a \(d\)-dimensional Riemannian manifold equipped with the Levi-Civita connection, \(\pi : OM \to M\) be its orthonormal frame bundle with structure group \(O(d)\). We denote by \(\omega\) the connection form associated with the Levi-Civita connection. Denote by \(H_\xi\) the standard horizontal vector field on \(P\) corresponding to \(\xi \in \mathbb{R}^d\), uniquely characterized by the property that \(\pi_*(H_\xi(u)) = u(\xi)\) for all \(u \in OM\). Let \(\{e_1, ..., e_d\}\) be the canonical basis of \(\mathbb{R}^d\), with dual basis \(\{e_1^*, ..., e_d^*\}\). We
denote in shorthand $H_i := H_{e_i}$. Then $H_\xi = \xi^i H_i$. Every $A \in \frs\frd$, Lie algebra of the rotation group $SO(d)$, induces a vector field $A^*$ on $OM$, called the fundamental vector field corresponding to $A$. It is well-known that the Lie algebra $\frs\frd$ consists of all skew-symmetric $d \times d$ matrices, and its dimension is $D = d(d - 1)/2$. If we denote by $a_i^j = e_i \otimes e_j^\ast - e_j \otimes e_i^\ast$ the $d \times d$ matrix such that the entry at the $i$-th column and the $j$-th row is $1$, the entry at the $j$-th column and the $i$-th row is $-1$ and other entries are all zero. Then $\{a_i^j \mid 1 \leq i < j \leq d\}$ form a basis of $\frs\frd$. We rearrange this basis into $\{A_1, \cdots , A_D\}$. We refer the reader to [KN] and [Elton] for more details on the issues of differential geometry or stochastic analysis on manifolds.

Recall that the equation of rolling perturbed by small random slipping and twisting is set up on $OM$ as

$$du_t^\varepsilon = H_i(u_t^\varepsilon) \circ d\gamma_t^\varepsilon + A_\alpha^i(u_t^\varepsilon) \circ dW_t^{\varepsilon,\alpha}, \quad u_0^\varepsilon = u_0 \in OM,$$

(3.1)

Let $x^\varepsilon = \{x_t^\varepsilon\}_{t \geq 0}$ be the projection curve $x_t = \pi(u_t^\varepsilon)$. We denote by $\tilde{x}^\varepsilon = \{\tilde{x}_t^\varepsilon\}_{t \geq 0}$ the horizontal lift of the curve $\{x_t^\varepsilon\}$ starting at $u_0$. Then we have

Lemma 3.1. Let $M$ be geodesically complete. Then

$$d\tilde{x}_t^\varepsilon = H_i(\tilde{x}_t^\varepsilon)(g_j^\varepsilon) \circ d\gamma_t^\varepsilon, \quad \tilde{x}_0^\varepsilon = u_0,$$

(3.2)

$$dg_t^\varepsilon = g_t^\varepsilon A_\alpha \circ dW_t^{\varepsilon,\alpha}, \quad g_0 = I_d,$$

(3.3)

where $I_d$ denotes the $d \times d$ identity matrix.

Proof. We follow the lines of [18, Lemma 3.1]. Set

$$u_t^\varepsilon = \tilde{x}_t^\varepsilon g_t^\varepsilon.$$

(3.4)

Then by (3.1) and $\tilde{x}^\varepsilon = \{\tilde{x}_t^\varepsilon\}_{t \geq 0}$,

$$dx_t^\varepsilon = \pi^\varepsilon(du_t^\varepsilon) = \pi^\varepsilon((du_t^\varepsilon) h) = \pi^\varepsilon(H_i(u_t^\varepsilon) \circ d\gamma_t^\varepsilon) = u_t^\varepsilon \circ d\gamma_t^\varepsilon = \tilde{x}_t^\varepsilon g_t^\varepsilon \circ d\gamma_t^\varepsilon,$$

which implies (3.2). On the other hand, (3.4) yields that

$$du_t^\varepsilon = d(R_{g_t^\varepsilon})(\circ d\tilde{x}_t^\varepsilon) + \tilde{x}_t^\varepsilon \circ dg_t^\varepsilon.$$

For a $g \in O(d)$, we denote by $R_g$ and $L_g$ the right and left action of $g$ on $OM$, and denote by $d(R_g)$ and $d(L_g)$ their differentials. Then

$$\omega(\circ du_t^\varepsilon) = Ad((g_t^\varepsilon)^{-1}) \omega(\circ d\tilde{x}_t^\varepsilon) + d(L(g_t^\varepsilon)^{-1})(\circ dg_t^\varepsilon) = d(L(g_t^\varepsilon)^{-1})(\circ dg_t^\varepsilon).$$

We also have by (3.1) that

$$\omega(\circ du_t^\varepsilon) = A_\alpha \circ dW_t^{\varepsilon,\alpha}.$$

Hence

$$dg_t^\varepsilon = d(L(g_t^\varepsilon))(A_\alpha \circ dW_t^{\varepsilon,\alpha}) = g_t^\varepsilon A_\alpha \circ dW_t^{\varepsilon,\alpha}.$$

This gives (3.3). \qed

We will also need some probabilistic preparation. Given two continuous local martingales $X$ and $Y$ on a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq 0}, \mathbb{P})$, the quadratic covariation of $X$ and $Y$ is defined by

$$\langle X, Y \rangle_t := \lim_{\|\Delta\| \to 0} \sum_i (X_{t_{i+1}} - X_{t_i})(Y_{t_{i+1}} - Y_{t_i})$$

where $\Delta$ ranges over all partitions $\{t_i\}_{i=0}^k$ of the interval $[0, t]$ with $0 = t_0 \leq t_1 \leq \cdots \leq t_k = t$, and $\|\Delta\| = \max_{1 \leq i < k} |t_{i+1} - t_i|$ is the mesh of $\Delta$. This limit, if it exists, is defined using convergence in probability. The following lemma will be used.
Lemma 3.2. Let $X$ be a one-dimensional continuous local martingale, $W$ be a one-dimensional standard Brownian motion independent of $X$. Then $\langle X, W \rangle = 0$.

Proof. Fix $t > 0$. By the definition of quadratic covariation, it is sufficient to prove that
\[ \sum_i (X_{t_i+1} - X_{t_i})(W_{t_i+1} - W_{t_i}) \to 0, \text{ in } L^2, \text{ as } \|\Delta\| \to 0. \]
We denote by $S^\Delta$ the summation in the previous sentence. Then
\[ |S^\Delta|^2 = \sum_i (X_{t_i+1} - X_{t_i})^2(W_{t_i+1} - W_{t_i})^2 + \sum_{i \neq j} (X_{t_i+1} - X_{t_i})(X_{t_j+1} - X_{t_j})(W_{t_i+1} - W_{t_i})(W_{t_j+1} - W_{t_j}) =: I_1 + I_2. \]
For $I_1$, using the independence of $X$ and $W$, we have
\[ \mathbb{E}(I_1) = \sum_i (t_{i+1} - t_i)\mathbb{E}(X_{t_i+1} - X_{t_i})^2 \leq \|\Delta\| \sum_i \mathbb{E}(X_{t_i+1} - X_{t_i})^2, \]
which goes to zero as $\|\Delta\| \to 0$, since $\sum_i \mathbb{E}(X_{t_i+1} - X_{t_i})^2 \to \mathbb{E}(\langle X, X \rangle)$. For $I_2$, since $W$ has independent increments,
\[ \mathbb{E}(I_2) = \sum_{i \neq j} \mathbb{E}[(X_{t_i+1} - X_{t_i})(X_{t_j+1} - X_{t_j})]\mathbb{E}(W_{t_i+1} - W_{t_i})\mathbb{E}(W_{t_j+1} - W_{t_j}) = 0. \]
Combining these, $\mathbb{E}|S^\Delta|^2 \to 0$ as $\|\Delta\| \to 0$, the result follows. \qed

3.2 Large deviations for rolling on compact manifolds

Now we denote the canonical decomposition of the continuous semimartingale $\gamma^\epsilon$ by

\[
\gamma^\epsilon = m^\epsilon + a^\epsilon,
\]
where $m^\epsilon$ is the continuous local martingale part, and $a^\epsilon$ is the continuous finite variation part.

We define the space $H^1$ by
\[
H^1 := H^1(\mathbb{R}_+, \mathbb{R}^d) = \{ f : \mathbb{R}_+ \to \mathbb{R}^d \mid f \text{ is absolutely continuous and } \int_0^\infty |\dot{f}(t)|^2 dt < \infty \},
\]
equipped with norm $\|f\|_{H^1} = \int_0^\infty |\dot{f}(t)|^2 dt$.

Theorem 3.3. Let $M$ be a compact Riemannian manifold. Then for each $\epsilon > 0$, the SDE (3.1) is conservative. If the family $\{(\gamma^\epsilon, \langle m^\epsilon, m^\epsilon \rangle)\}_{\epsilon > 0}$ satisfies the LDP with good rate function $I$, and the family $\{G(\gamma^\epsilon)\}_{\epsilon > 0}$ is exponentially tight for each $t > 0$, then $\{u^\epsilon\}_{\epsilon > 0}$, $\{\tilde{\epsilon}^\epsilon\}_{\epsilon > 0}$ and $\{x^\epsilon\}_{\epsilon > 0}$ all satisfy the LDP with the following good rate functions respectively:
\[
I_{OM}(u) = \inf \{ I(y, q) + \frac{1}{2} \int_0^\infty |\dot{f}(t)|^2 dt : u = u_0 + H(u) \cdot y + A^\star(u) \cdot f + \frac{1}{2} (\nabla_H H)(u) \cdot q, f \in H^1, y \text{ and } q \text{ are locally of finite variation} \},
\]
\[
\bar{I}_{OM}(v) = \inf \{ I(y, q) + \frac{1}{2} \int_0^\infty |\dot{f}(t)|^2 dt : v = v_0 + H(v) g \cdot y + \frac{1}{2} (\nabla_H H)(v) g \cdot q, g = 1 + g A \cdot f^\alpha, f \in H^1, y \text{ and } q \text{ are locally of finite variation} \},
\]
\[
I_M(x) = \inf \{ I_{OM}(u) : \pi(u) = x \} = \inf \{ \bar{I}_{OM}(v) : \pi(v) = x \}. \]
Proof. The well-known Whitney’s embedding theorem tells that $OM$ admits a proper smooth embedding into $\mathbb{R}^p$ with some $p \in \mathbb{N}$, which we denote as $\iota : OM \to \mathbb{R}^p$. Using this embedding, the rolling equation becomes

$$d(\iota \circ u^t)_t = \iota_*(H_1)((\iota \circ u^t)_t) \circ d\epsilon_t^\iota + \iota_*(A_\alpha^*)((\iota \circ u^t)_t) \circ dW_t^\epsilon,\alpha,$$

Since $M$ is compact, its orthonormal frame bundle $OM$ with structure group $O(d)$ is also compact. Thus, the smooth vector fields $\iota_*(H_i)$’s and $\iota_*(A_\alpha^*)$’s on $\iota(OM)$ are all bounded and Lipschitz, with Lipschitz derivatives. The properness of the embedding $\iota$ yields that $\iota(OM)$ is a closed submanifold of $\mathbb{R}^p$. Thus, one can extend each $\iota_*(H_i)$ and $\iota_*(A_\alpha^*)$ to bounded and Lipschitz vector fields on $\mathbb{R}^p$ with Lipschitz derivatives, which we denote as $\tilde{\iota}_*(H_i)$ and $\tilde{\iota}_*(A_\alpha^*)$. Since these extended vector fields are tangent to $\iota(OM)$ when restricted on it, $\iota \circ u^t$ is also the solution to the following equation

$$d(\iota \circ u^t)_t = \iota_*(\tilde{H}_1)((\iota \circ u^t)_t) \circ d\epsilon_t^\iota + \iota_*(\tilde{A}_\alpha^*)((\iota \circ u^t)_t) \circ dW_t^\epsilon,\alpha,$$

Since $\langle W^\epsilon, W^\epsilon \rangle_t = \epsilon t$, the family $\langle W^\epsilon, W^\epsilon \rangle$ converges to the constant path 0 in $C$ as $\epsilon \to 0$. It follows from the classical results that the family $\{(W^\epsilon, (\langle W^\epsilon, W^\epsilon \rangle))\}$ satisfies the LDP with good rate function

$$I'(f, g) = \begin{cases} \frac{1}{2} \int_0^\infty |\dot{f}(t)|^2 dt, & \text{if } f \in H^1, g \equiv 0, \\ \infty, & \text{otherwise}. \end{cases}$$

The independence of $\gamma^\epsilon$ and $W^\epsilon$ yields that the LDP holds for $\{(\gamma^\epsilon, \langle m^\epsilon, m^\epsilon \rangle, W^\epsilon, (\langle W^\epsilon, W^\epsilon \rangle))\}$ with good rate function $I''(y, q, f, g) = I(y, q) + I'(f, g)$. It follows from Lemma 3.2 that

$$\langle (m^\epsilon, W^\epsilon), (m^\epsilon, W^\epsilon) \rangle = \begin{pmatrix} \langle m^\epsilon, m^\epsilon \rangle & 0 \\ 0 & \langle W^\epsilon, W^\epsilon \rangle \end{pmatrix}.$$  

On the other hand, $G((\gamma^\epsilon, W^\epsilon)) = G(\gamma^\epsilon) + \frac{1}{\epsilon} I((\gamma^\epsilon, W^\epsilon)) = G(\gamma^\epsilon) + Nt$. For each $t$, the exponential tightness of $\{G(\gamma^\epsilon)_t\}$ yields that $\{G(\gamma^\epsilon, W^\epsilon)_t\}$ is also exponentially tight.

Now we can apply Corollary 2.8, the family $\{(\iota \circ u^t)\}$ satisfies the LDP with good rate function

$$I_{\iota(OM)}(x) = \inf \{I''(y, q, f, g) : x = \iota(u_0) + \iota_*(H)(x) \cdot y + \iota_*(A^*)(x) \cdot f + \frac{1}{2} (\iota_*(A^*) \partial_x \iota_*(A^*))(x) \cdot g, y, q, f \text{ and } g \text{ are locally of finite variation} \}$$

$$= \inf \{I(y, q) + \frac{1}{2} \int_0^\infty |\dot{f}(t)|^2 dt : x = \iota(u_0) + \iota_*(H)(x) \cdot y + \iota_*(A^*)(x) \cdot f \\ + \frac{1}{2} (\iota_*(A^*) \partial_x \iota_*(A^*))(x) \cdot g, y, q, f \text{ and } g \text{ are locally of finite variation} \}$$

$$= \inf \{I(y, q) + \frac{1}{2} \int_0^\infty |\dot{f}(t)|^2 dt : x = \iota(u_0) + \iota_*(H)(x) \cdot y + \iota_*(A^*)(x) \cdot f \\ + \frac{1}{2} (\iota_*(H) \partial_x \iota_*(H))(x) \cdot q, f \in H^1, y \text{ and } q \text{ are locally of finite variation} \}$$

where the last equality follows from the fact that the path $x$ always stays on $\iota(OM)$ since all the vector fields $\iota_*(H_i)$ and $\iota_*(A_\alpha^*)$ are tangent to $\iota(OM)$ when restricted on it and the initial value $\iota(u_0) \in \iota(OM)$. Since $\iota$ is a smooth embedding, the inverse contraction principle (see, e.g., [9, Theorem 4.2.4]) implies that the family $\{u^t\}$ satisfies the LDP with good rate

$$I_{OM}(u) = I_{\iota(OM)}(\iota \circ u),$$

which yields (3.5).
Now we turn to \( \{ \tilde{x}^\epsilon \} \). Since \( M \) is compact, it is geodesically complete. We will use Lemma 3.1. It is easy to see that the coefficient function of SDE (3.3) is global Lipschitz and bounded. The classical Freidlin-Wentzell theory for Stratonovich-type SDE (see [17, Theorem 2.5]), or using our general result Corollary 2.8) yields that the family \( \{ g^\epsilon \} \) satisfies the LDP with good rate function
\[
\tilde{I}(g) = \inf \{ \frac{1}{2} \int_0^\infty |\dot{f}(t)|^2 dt : g = I + gA_\alpha \cdot f^\alpha, f \in H^1 \}.
\]
Since \( \gamma^\epsilon \) is independent of \( W^\epsilon \), it is also independent of \( g^\epsilon \). Thus the family \( \{ (\gamma^\epsilon, \langle m^\epsilon, m^\epsilon \rangle, g^\epsilon) \} \) satisfies the LDP with good rate function \( \tilde{I}^\epsilon(y, q, g) = I(y, q) + \tilde{I}(g) \). Similar as \( u^\epsilon \), Corollary 2.8 yields that the family \( \{ \tilde{x}^\epsilon \} \) satisfies the LDP with good rate
\[
\tilde{I}_{OM}(v) = \inf \{ I(y, g, q) : v = u_0 + H(v)g \cdot y + \frac{1}{2}(\nabla H)(v)g \cdot q, y \text{ and } q \text{ are locally of finite variation} \},
\]
the representation (3.6) follows. The large deviation principle of \( \{ x^\epsilon \} \) and its rate function (3.7) follows the continuity of the projection \( \pi \) and the contraction principle. \( \square \)

## 4 Rolling problems on noncompact manifolds

In this section, we deal with the rolling systems where the rolled manifold \( M \) is a complete Riemannian manifold. For the Euclidean curves, we only consider two special case: the random curves with locally finite variation and the random curve driven by stochastic differential equations.

### 4.1 Along pathwise rectifiable random curves

We consider the case that \( m^\epsilon \equiv 0 \), that is, each \( \gamma^\epsilon \) is continuous process with locally finite variation. In this case, both Stratonovich circles in (3.1) and (3.2) can be removed, because the integrals therein are in the Riemann-Stieltjes sense. Note also that in this case, \( G(\gamma^\epsilon) = V(\gamma^\epsilon) \).

**Theorem 4.1.** Let \( M \) be a geodesically complete Riemannian manifold. Let \( \{ \gamma^\epsilon \}_{\epsilon > 0} \) be a family of adapted continuous processes with locally finite variation. Then for each \( \epsilon > 0 \), the SDE (3.1) is conservative. If the family \( \{ \gamma^\epsilon \}_{\epsilon > 0} \) satisfies the LDP with good rate function \( \tilde{I}^\epsilon \) and the family \( \{ |V(\gamma^\epsilon)|_t \}_{\epsilon > 0} \) is exponentially tight for each \( t > 0 \), then \( \{ u^\epsilon \}_{\epsilon > 0} \) and \( \{ \tilde{x}^\epsilon \}_{\epsilon > 0} \) all satisfy the LDP with the following good rate functions respectively:
\[
I_{OM}(u) = \inf \{ \tilde{I}^\epsilon(y) + \frac{1}{2} \int_0^\infty |\dot{f}(t)|^2 dt : u = u_0 + H(u) \cdot y + A^\epsilon(u) \cdot f, \\
\quad f \in H^1, y \text{ and } q \text{ are locally of finite variation} \}, \tag{4.1}
\]
\[
\tilde{I}_{OM}(v) = \inf \{ \tilde{I}^\epsilon(y) + \frac{1}{2} \int_0^\infty |\dot{f}(t)|^2 dt : v = u_0 + H(v)g \cdot y, q = I + gA_\alpha \cdot f^\alpha, \\
\quad f \in H^1, y \text{ and } q \text{ are locally of finite variation} \}, \tag{4.2}
\]
the family \( \{ x^\epsilon \}_{\epsilon > 0} \) satisfies the LDP with the good rate function (3.7).

**Proof.** Fix \( \epsilon > 0 \). Denote by \( T^\epsilon \) the lifetime of \( u^\epsilon \). Assume, by contradiction, \( P^\epsilon(T^\epsilon < \infty) > 0 \), that is, \( u^\epsilon \) would explode at finite time \( T^\epsilon \) with positive probability. Since \( \gamma^\epsilon \) is locally of finite variation, each sample path of it is rectifiable, and for each \( t > 0 \), the length of \( \gamma^\epsilon|_{[0,t]} \) is \( V(\gamma^\epsilon)_t \). Then by (3.2), for every \( t > 0 \),
\[
L(x^\epsilon|_{[0,t]}) = L \left( \left\{ \int_0^r g^\epsilon_s d\gamma^\epsilon_s : r \in [0,t] \right\} \right) = \int_0^t |g^\epsilon_s| dV(\gamma^\epsilon)_s = V(\gamma^\epsilon)_t.
\]
That is, \( x^\epsilon \) is also rectifiable. Thus, the completeness of \( M \) yields that the continuous path \( x^\epsilon|_{[0,T^\epsilon]} \) converges as \( t \to T^\epsilon \), on the event \( \{T^\epsilon < \infty\} \). As a consequence, the horizontal lift \( \tilde{x}^\epsilon|_{[0,T^\epsilon]} \) also converge as \( t \to T^\epsilon \). Since the equation (3.3) can be globally solved, by (3.4), the process \( u^\epsilon|_{[0,T^\epsilon]} \) would converge as well, as \( t \to T^\epsilon \). This leads to a contradiction with the necessary explosion of \( u^\epsilon \). Therefore, the SDE (3.1) is conservative.

For the second statement, we note that we have identified the rate functions in Theorem 3.3. Since each \( \gamma^\epsilon \) is locally of finite variation, \( m^\epsilon \equiv 0 \) and the family \( \{(\gamma^\epsilon, \langle m^\epsilon, m^\epsilon \rangle)\} \) satisfies the LDP with good rate function

\[
I(y, q) = \begin{cases} 
  \bar{I}^\epsilon(y), & \text{if } q \equiv 0, \\
  \infty, & \text{otherwise.}
\end{cases}
\]

Thus, the representation (4.1) and (4.2) follow from (3.5) and (3.6) respectively. Now it is only needed to prove that the family \( \{u^\epsilon\} \) and \( \{\tilde{x}^\epsilon\} \) are exponentially tight. We will use a similar argument as Corollary 2.7.

Suppose first that there exists a constant \( R > 0 \) such that \( d(x_0, x^\epsilon_t) \leq R \) for all \( t > 0 \). Let \( K_R := B\bar{M}(x_0, R) \). The completeness of \( M \) implies \( K_R = \exp(B_{T_{x_0}M}(0, R)) \). Thus, \( K_R \) is a compact submanifold of \( M \). Since all the paths of \( x^\epsilon \) are contained in the submanifold \( K_R \), all paths of \( u^\epsilon \) and \( \tilde{x}^\epsilon \) are contained in the bundle \( OM|_{K_R} = O(K_R) \). Consequently, the SDEs (3.1) and (3.2) can be view as equations on \( O(K_R) \). The exponential tightness of \( \tilde{x}^\epsilon \) follows from Theorem 3.3.

Now we prove the general case. For each \( \epsilon > 0 \) and \( p > 0 \), define a stopping time

\[
T^{\epsilon,p} = \inf\{t \geq 0 : d(x_0, x^\epsilon_t) \geq p\}.
\]

Observe that \( d(x_0, x^\epsilon_t) \leq L(x^\epsilon|_{[0,t]}) = V(\gamma^\epsilon)_t \).

Then

\[
P^\epsilon(T^{\epsilon,p} \leq T) = P^\epsilon(d(x_0, x^\epsilon_t) \geq p) \leq P^\epsilon(V(\gamma^\epsilon)_t \geq p).
\]

Then a similar localization argument in Step 2 of the proof of Proposition 2.6 yields the exponential tightness.

\[
\square
\]

### 4.2 Along random curves driven by stochastic differential equations

For each \( \epsilon > 0 \), let \( \gamma^\epsilon \) be the solution of the following SDE:

\[
d\gamma^\epsilon_t = b(t, \gamma^\epsilon_t)dt + \sqrt{\epsilon}dB_t, \quad \gamma^\epsilon_0 = \gamma_0.
\]

(4.3)

where \( B \) is a \( d \)-dimensional Brownian motion independent with \( W, b : [0, \infty) \times \mathbb{R}^d \to \mathbb{R}^d \) is a function such that the SDE has a unique strong solution.

**Theorem 4.2.** Let \( M \) be a geodesically and stochastically complete Riemannian manifold and assume there exists a constant \( L > 1 \), such that the Ricci curvature is bounded below by \(-L\). For each \( \epsilon > 0 \), let \( \gamma^\epsilon \) be the unique solution of (4.3). If the function \( b \) is bounded and satisfies

\[
|b(t, x_1) - b(t, x_2)| \leq C|x_1 - x_2|
\]

for all \( t \geq 0 \) and \( x_1, x_2 \in \mathbb{R}^d \) with some constant \( C > 0 \), then for each \( \epsilon > 0 \), the SDE (3.1) is conservative, and the families \( \{u^\epsilon\}_{\epsilon > 0} \) and \( \{\tilde{x}^\epsilon\}_{\epsilon > 0} \) all satisfy the LDP with the following good rate functions respectively:

\[
I_{OM}(u) = \inf\{\frac{1}{2} \int_0^\infty (|\dot{y}(t) - b(t, y(t))|^2 + |\dot{f}(t)|^2)dt : u = u_0 + H(u) \cdot y + A^*(u) \cdot f, y, f \in H^1\},
\]

(4.4)
\[ I_{OM}(v) = \inf \left\{ \frac{1}{2} \int_0^\infty \left( |\dot{y}(t) - b(t, y(t))|^2 + |f(t)|^2 \right) dt : v = u_0 + H(t)g \cdot y, g = I + gA_\alpha \cdot f_\alpha, y, f \in H^1 \right\}, \]

the family \( \{ x^\epsilon \}_{\epsilon > 0} \) satisfies the LDP with the good rate function (3.7).

**Proof.** Firstly, we show the non-explosion of (3.1). For each \( \epsilon > 0 \), define \( \zeta^\epsilon \) to be the following integral process

\[ \zeta^\epsilon_t := \int_0^t g^\epsilon_s \circ d\gamma_s = \int_0^t g^\epsilon_s b(s, \gamma^\epsilon_s)ds + \sqrt{\epsilon} \int_0^t g^\epsilon_s \circ dB_s = \int_0^t g^\epsilon_s b(s, \gamma^\epsilon_s)ds + \sqrt{\epsilon} \int_0^t g^\epsilon_s dB_s. \]

Here the last equality follows from \( \langle g^\epsilon, B \rangle = 0 \) by (3.3) and the independence. Since the process \( g^\epsilon \) take values in \( O(d) \),

\[ \left( \int_0^t g^\epsilon_s dB_s, \int_0^t g^\epsilon_s dB_s \right)_t = \int_0^t g^\epsilon_s (g^\epsilon_s)^* ds = tI. \]

By Lévy’s characterization of Brownian motion (see, e.g., [16, Theorem 3.3.16]), the stochastic integral \( \int_0 g^\epsilon_s dB_s \) is a \( d \)-dimensional Brownian motion under \( P^\epsilon \), which we denote as \( \hat{A}^\epsilon \). Then

\[ \zeta^\epsilon_t = \int_0^t g^\epsilon_s b(s, \gamma^\epsilon_s)ds + \sqrt{\epsilon} \hat{A}^\epsilon_t. \]

By (3.2) and the definition of \( \zeta^\epsilon \), we have

\[ d\bar{x}^\epsilon_t = H_i(\bar{x}^\epsilon_t) \circ d\zeta^\epsilon_t. \]

Let

\[ Z^\epsilon_t = \exp \left\{ -\frac{1}{\sqrt{\epsilon}} \int_0^t g^\epsilon_s b(s, \gamma^\epsilon_s) d\hat{A}^\epsilon_s - \frac{1}{2\epsilon} \int_0^t |g^\epsilon_s b(s, \gamma^\epsilon_s)|^2 ds \right\}. \]

Then the boundedness of \( b \) and \( g^\epsilon \) yields that \( Z^\epsilon \) is a continuous martingale. Define, for each \( t > 0 \), a probability measure on \( F^\epsilon_t \) by \( \bar{P}^\epsilon_t(A) := E^\epsilon_t(Z^\epsilon_t; A). \) By Girsanov theorem (see [16, Theorem 3.5.1]), each process \( \{ \zeta^\epsilon_t \}_{0 \leq s \leq t} \) in (4.6) is a Brownian motion with covariance matrix \( \epsilon I \) under \( \bar{P}^\epsilon_t \).

Denote by \( T^\epsilon \) the lifetime of each \( x^\epsilon \). Then for each \( \epsilon > 0 \) and \( t > 0 \), the lifetime of the process \( \{ x^\epsilon_{s/\epsilon} \}_{0 \leq s \leq \epsilon t} \) is \( (\epsilon T^\epsilon) \wedge t \). By (4.7), we know that the process \( \{ x^\epsilon_{s/\epsilon} \}_{0 \leq s \leq \epsilon t} \) is a standard Riemannian Brownian motion under \( \bar{P}^\epsilon_t \). Since \( M \) is stochastically complete, we have

\[ \bar{P}^\epsilon_t(T^\epsilon \geq t/\epsilon) = \bar{P}^\epsilon_t((\epsilon T^\epsilon) \wedge t = t) = 1. \]

Meanwhile, by (4.8) and (4.6), it is easy to deduce

\[ (Z^\epsilon_t)^{-1} = \exp \left\{ \frac{1}{\epsilon} \int_0^t g^\epsilon_s b(s, \gamma^\epsilon_s) d\zeta^\epsilon_s - \frac{1}{2\epsilon} \int_0^t |g^\epsilon_s b(s, \gamma^\epsilon_s)|^2 ds \right\}. \]

Then \( \{(Z^\epsilon_t)^{-1} \}_{0 \leq s \leq \epsilon t} \) is a continuous martingale under \( \bar{P}^\epsilon_t \) with \( \bar{E}^\epsilon_t((Z^\epsilon_t)^{-1}) = \bar{E}^\epsilon_t((Z^\epsilon_0)^{-1}) = 1, \) since \( \{ \zeta^\epsilon_t \}_{0 \leq s \leq t} \) is Brownian motion with covariance matrix \( \epsilon I \) under this probability measure. Hence, for every \( t > 0 \),

\[ P^\epsilon(T^\epsilon \geq t/\epsilon) = \bar{E}^\epsilon_t((Z^\epsilon_t)^{-1}; T^\epsilon \geq t/\epsilon) = \bar{E}^\epsilon_t((Z^\epsilon_t)^{-1}) = 1. \]

This lead to \( P^\epsilon(T^\epsilon = \infty) = 1 \), which means for each \( \epsilon > 0 \), with probability 1 the process \( x^\epsilon \) does not explode. The horizontal lifts \( \bar{x}^\epsilon \) will also not explode. Since \( u^\epsilon = \bar{x}^\epsilon g^\epsilon \) with \( SO(d) \)-valued process \( g^\epsilon \) globally defined by Lemma 3.1, the process \( u^\epsilon \) does not explode as well, for every \( \epsilon > 0 \).
We now identify the rate functions. The canonical decomposition of each semimartingale $\gamma^\epsilon$ is $\gamma^\epsilon = a^\epsilon + m^\epsilon$, with $a^\epsilon_t = \int_0^t b(s, \gamma^\epsilon_s)ds$ and $m^\epsilon_t = \sqrt{\epsilon} B_t$. Then, $\langle m^\epsilon, m^\epsilon \rangle = et$ and

$$G(\gamma^\epsilon)_t = V(a^\epsilon)_t + t = \int_0^t |b(s, \gamma^\epsilon_s)|ds + t \leq (C + 1)t.$$ 

It follows that $\{G(\gamma^\epsilon)_t\}$ is exponentially tight for each $t > 0$. By the classical Freidlin-Wentzell theory and the fact the deterministic path $\langle m^\epsilon, m^\epsilon \rangle$ converges to zero in $C$ as $\epsilon \to 0$, the family $\{(\gamma^\epsilon, \langle m^\epsilon, m^\epsilon \rangle)\}$ satisfies the LDP with good rate function

$$I(y, q) = \left\{ \frac{1}{2} \int_0^\infty |\dot{y}(t) - b(t, y(t))|^2 dt, \text{ if } y \in H^1, q = 0, \right.$$ 

otherwise.

The representation (4.4) and (4.5) follow from (3.5) and (3.6) respectively.

To prove the exponential tightness of $\{a^\epsilon\}$ and $\{\tilde{x}^\epsilon\}$, as in the proof of Theorem 4.1, it is enough to show that for each $t > 0$, the family $\{\sup_{0 \leq s \leq t} d(x_0, x^\epsilon_s)\}_{\epsilon > 0}$ is exponentially tight. Suppose first $b \equiv 0$. Then $d\tilde{x}^\epsilon_t = H_t(\tilde{x}^\epsilon_t) \circ d(\sqrt{\epsilon} \tilde{B}^\epsilon_t) = H_t(\tilde{x}^\epsilon_t) \circ d(B^\epsilon_t)$, that is, the rescaled process $x^\epsilon/\epsilon$ is a standard Riemannian Brownian motion. By [17, Proposition 3.7], we have

$$\mathbb{P}^\epsilon \left( \sup_{0 \leq s \leq t} d(x_0, x^\epsilon_s) \geq a \right) = \mathbb{P}^\epsilon \left( \sup_{0 \leq s \leq t} d(x_0, x^\epsilon_{s/\epsilon}) \geq a \right) \leq 2 \exp \left\{ -\frac{(kL\epsilon - \frac{1}{2}a^2)^2}{2a^2\epsilon t} \right\}. $$

The exponential tightness follows. For the general case, we apply Girsanov’s transform. Since $\{x^\epsilon_{s/\epsilon}\}_{0 \leq s \leq t}$ is a standard Riemannian Brownian motion under $\mathbb{P}_t^\epsilon$, we have

$$\tilde{\mathbb{P}}^\epsilon_t \left( \sup_{0 \leq s \leq t} d(x_0, x^\epsilon_s) \geq a \right) \leq 2 \exp \left\{ -\frac{(kL\epsilon - \frac{1}{2}a^2)^2}{2a^2\epsilon t} \right\}. \quad (4.9)$$

By Chebyshev’s inequality, for any $A > 0$,

$$\tilde{\mathbb{P}}^\epsilon_t((Z^\epsilon_t)^{-1} \geq e^{A/\epsilon}) \leq e^{-A/\epsilon}. \quad (4.10)$$

Hence, by the definition of $\tilde{\mathbb{P}}^\epsilon_t$ and (4.9), (4.10),

$$\mathbb{P}^\epsilon \left( \sup_{0 \leq s \leq t} d(x_0, x^\epsilon_s) \geq a \right) = \tilde{\mathbb{E}}^\epsilon_t \left( (Z^\epsilon_t)^{-1}; \sup_{0 \leq s \leq t} d(x_0, x^\epsilon_s) \geq a \right)$$

$$= \tilde{\mathbb{E}}^\epsilon_t \left( (Z^\epsilon_t)^{-1}; \sup_{0 \leq s \leq t} d(x_0, x^\epsilon_s) \geq a \right) = e^{A/\epsilon} \tilde{\mathbb{E}}^\epsilon_t \left( \sup_{0 \leq s \leq t} d(x_0, x^\epsilon_s) \geq a \right) + e^{-A/\epsilon}$$

$$\leq 2 \exp \left\{ \frac{A}{\epsilon} - \frac{(kL\epsilon - \frac{1}{2}a^2)^2}{2a^2\epsilon t} \right\} + e^{-A/\epsilon}. $$

Then

$$\limsup_{\epsilon \to 0} \epsilon \log \mathbb{P}^\epsilon \left( \sup_{0 \leq s \leq t} d(x_0, x^\epsilon_s) \geq a \right) \leq \left( A - \frac{a^2}{8t} \right) \vee (-A).$$

The exponential tightness of $\{\sup_{0 \leq s \leq t} d(x_0, x^\epsilon_s)\}_{\epsilon > 0}$ follows by letting first $a \to \infty$ and then $A \to \infty$. \qed
5 Examples

Throughout this section, we fix a differentiable curve $\gamma : [0, \infty) \to \mathbb{R}^d$ satisfying $|\gamma_t| \leq C$ for all $t \geq 0$ with some constant $C > 0$. Then obviously $\gamma$ is rectifiable.

**Example 5.1 (Random perturbation).** A simple case is the Euclidean curve which the manifold rolls along is a random perturbation of a given curve. To be precise, consider the following family of random perturbation of $\gamma$:

$$\gamma_t^\epsilon = \gamma_t + \sqrt{\epsilon} B_t.$$  

Then $\gamma_t^\epsilon$ solves the SDE (4.3) with $b \equiv \dot{\gamma}$. And Theorem 4.2 is applicable here. \hfill \Box

In the next examples, we will study the rolling mode along pathwise rectifiable random perturbation. We will make sure Theorem 4.1 is applicable, by checking that the family $\{\gamma_t^\epsilon\}_{\epsilon > 0}$ satisfies the LDP with some good rate function, and the family $\{|V(\gamma_t^\epsilon)|_t\}_{\epsilon > 0}$ is exponentially tight for each $t > 0$. The following lemma is a criterion for large deviation principle for the processes with locally finite variation, which is adapted from [22, Theorem 2.1].

**Lemma 5.2.** Let $\{A^\epsilon\}_{\epsilon > 0}$ be a family of continuous adapted processes with locally finite variation. Let $g : [0, \infty) \to \mathbb{R}^d$ be a Borel function such that $\int_0^\infty |g(t)| dt < \infty$. Let $a_t = \int_0^t g(s) ds$. If for every $T > 0$ and any $\eta > 0$,

$$\limsup_{\epsilon \to 0} \epsilon \log P^\epsilon \left( \sup_{0 \leq t \leq T} |A_t^\epsilon - a_t| \geq \eta \right) = -\infty.$$

Then the family $\{A^\epsilon\}_{\epsilon > 0}$ satisfies the LDP with good rate function

$$I(f) = \begin{cases} \int_0^\infty \sup_{\lambda \in \mathbb{R}^d} \langle \lambda, \dot{f}(t) - g(t) \rangle dt, & f \text{ is absolutely continuous}, \\ \infty, & \text{otherwise}. \end{cases}$$

In the next two examples, we will consider the rolling procedure mixed by random slipping. To indicate the slipping moments and slipping duration, we need a Poisson point process. For this purpose, we introduce a subordinator $S^\epsilon = \{S_t^\epsilon\}_{t \geq 0}$, which has zero drift and jump measure $\nu^\epsilon$ with $\nu^\epsilon((0, \infty)) < \infty$, for each $\epsilon > 0$. Then, almost surely, the jumping times of each $S_t^\epsilon$ are infinitely many and countable in increasing order (see [23, Theorem 21.3]). We define the jumping time of each $S_t^\epsilon$ recursively by

$$\tau_1^\epsilon = \inf\{t \geq 0 : \Delta S_t^\epsilon \neq 0\},$$

$$\tau_{k+1}^\epsilon = \inf\{t > \tau_k^\epsilon : \Delta S_t^\epsilon \neq 0\}, \quad k \in \mathbb{N}_+.$$  

These stopping times indicate the moments when every random slipping will occur. Let $e^\epsilon = \{e^\epsilon(t)\}_{t \geq 0}$ be the associated Poisson point process valued on $[0, \infty)$, namely, $e^\epsilon = \Delta S^\epsilon$. The value of $e^\epsilon$ on each jumping time $\tau_k^\epsilon$ indicates the duration of each random slipping.

**Example 5.3 (Translational slipping).** Consider the case that each slipping is translational along the given curve, that is, the contact point on the manifold will be rest while the contact point on the Euclidean space will move along $\gamma$ from the slipping starting point to the slipping end point. Then the equivalent Euclidean curve that one can slip the manifold along is given by

$$\gamma_t^\epsilon = \begin{cases} \gamma_t, & 0 \leq t < \tau_1^\epsilon, \\ \gamma_{\tau_1^\epsilon}, & \tau_1^\epsilon \leq t < \tau_1^\epsilon + e^\epsilon(\tau_1^\epsilon), \\ \gamma_t - \gamma_{\tau_1^\epsilon} + e^\epsilon(\tau_1^\epsilon) + \gamma_{\tau_1^\epsilon}, & \tau_1^\epsilon + e^\epsilon(\tau_1^\epsilon) \leq t < \tau_1^\epsilon + e^\epsilon(\tau_1^\epsilon) + \tau_2^\epsilon, \\ \vdots, & \end{cases}$$
for each $\epsilon > 0$. Then each $\gamma^\epsilon$ is locally of finite variation, and we have
\[
\sup_{0 \leq t \leq T} |\gamma^\epsilon_t - \gamma_t| \leq C \sum_{0 \leq t \leq T} e^\epsilon(t) = CS^T.
\]
If the family $\{S^\epsilon\}_{\epsilon>0}$ satisfies that for each $T > 0$ and $\eta > 0$,
\[
\limsup_{\epsilon \to 0} \epsilon \log P^\epsilon (S^\epsilon_T \geq \eta) = -\infty,
\]
then the LDP for $\{\gamma^\epsilon\}_{\epsilon>0}$ holds, by Lemma 5.2, with good rate function
\[
P^\epsilon(y) = \left\{ \begin{array}{ll}
\int_0^\infty \sup_{\lambda \in \mathbb{R}^d} \langle \lambda, \dot{y}(t) - \dot{\gamma}(t) \rangle dt, & y \text{ is absolutely continuous}, \\
\infty, & \text{otherwise}.
\end{array} \right.
\]
Moreover, it is easy to see that $|V(\gamma^\epsilon)| \leq |V(\gamma)|$, and the exponential tightness of $\{|V(\gamma^\epsilon)|_t\}$, for each $t > 0$, follows. Therefore, Theorem 4.1 is applicable here.

Example 5.4 (Slipping in place). There is another slipping mode: the slipping only happens in place. In this mode, the contact point on the Euclidean space will rest, but the contact point on the manifold will move along the geodesic that starts from the slipping starting point with initial speed the tangent vector of the curve $\gamma$ at the slipping starting point. For each $\epsilon > 0$, the equivalent curve is given by
\[
\gamma^\epsilon_t = \begin{cases} 
\gamma_t, & 0 \leq t < \tau^\epsilon_1, \\
\gamma_{\tau^\epsilon_1} + (t - \tau^\epsilon_1)\dot{\gamma}_{\tau^\epsilon_1}, & \tau^\epsilon_1 \leq t < \tau^\epsilon_1 + e^\epsilon(\tau^\epsilon_1), \\
\gamma_t - e^\epsilon(\tau^\epsilon_1) + e^\epsilon(\tau^\epsilon_1)\dot{\gamma}_{\tau^\epsilon_1}, & \tau^\epsilon_1 + e^\epsilon(\tau^\epsilon_1) \leq t < \tau_1 + e^\epsilon(\tau^\epsilon_1) + \tau^\epsilon_2, \\
\cdots.
\end{cases}
\]
Then we have
\[
\sup_{0 \leq t \leq T} |\gamma^\epsilon_t - \gamma_t| \leq 2C \sum_{0 \leq t \leq T} e^\epsilon(t) = 2CS^T.
\]
Same as before, the condition (5.2) yields the LDP for $\{\gamma^\epsilon\}_{\epsilon>0}$, with good rate function (5.3). On the other hand, it is easy to derive
\[
|V(\gamma^\epsilon)|_t \leq |V(\gamma)|_t + C \sum_{0 \leq s \leq t} e^\epsilon(s) \leq C(t + S^T).
\]
Hence, condition (5.2) also implies the exponential tightness of $\{|V(\gamma^\epsilon)|_t\}$, for each $t > 0$.

Example 5.5 (Piecewise linear approximation). In this example, we will consider the rolling along the piecewise linear approximation of $\gamma$. Assume the speed function $\dot{\gamma}$ is Lipschitz. For each $\epsilon > 0$, let $\{\tau^\epsilon_k\}_{k=1}^\infty$ be the sequence of stopping times defined in (5.1). Set $\tau^\epsilon_0 \equiv 0$ for simplicity. We define the speed function of the approximation curve by
\[
\xi^\epsilon_t = \dot{\gamma}_{\tau^\epsilon_k}, \quad \text{when } \tau^\epsilon_k \leq t < \tau^\epsilon_{k+1}, k \geq 0.
\]
Then $|\xi^\epsilon_t| < C$ for $t \geq 0$ and $\epsilon > 0$. Define the approximation curves by
\[
\gamma^\epsilon_t = \int_0^t \xi^\epsilon_s ds, \quad \epsilon > 0.
\]
Then each $\gamma^t$ is piecewise linear and locally of finite variation with variation satisfying $|V(\gamma^t)| \leq Ct$. Hence, for each $t > 0$, the family $\{V(\gamma^t)\}_{t>0}$ is exponentially tight.

Since $\dot{\gamma}$ is Lipschitz, we have

$$\sup_{0 \leq t \leq T} |\gamma^t_t - \gamma_t| \leq \int_0^T |\xi^t_s - \dot{\gamma}_s| ds \leq \begin{cases} CT^2, & 0 < T \leq \tau_1^t, \\ CT \sup_{0 \leq k \leq m - 1} (\tau_{k+1}^t - \tau_k^t) \vee (T - \tau_m^t), & \tau_m^t < T \leq \tau_{m+1}^t, \end{cases}$$

Then for every $T > 0$ and any $\eta > 0$,

$$P^\epsilon \left( \sup_{0 \leq t \leq T} |\gamma^t_t - \gamma_t| \geq \eta \right) \leq P^\epsilon \left( CT^2 \geq \eta, \tau_1^t \geq T \right) + \sum_{m=1}^\infty P^\epsilon \left( CT \sup_{0 \leq k \leq m - 1} (\tau_{k+1}^t - \tau_k^t) \vee (T - \tau_m^t) \geq \eta, \tau_m^t < T \leq \tau_{m+1}^t \right)$$

$$=: I^\epsilon + \sum_{m=1}^\infty J_m^\epsilon. \quad (5.4)$$

Denote $\lambda(\epsilon) := \nu^\epsilon((0, \infty))$. Since the sequence $\{\tau_{k+1}^t - \tau_k^t\}_{k=0}^\infty$ constitutes independent identically distributed random variables, each exponentially distributed with mean $1/\lambda(\epsilon)$ (see the proof of [23, Theorem 21.3]), we have

$$I^\epsilon = 1_{\{CT^2 \geq \eta\}} P^\epsilon (\tau_1^t \geq T) = 1_{\{CT^2 \geq \eta\}} e^{-\lambda(\epsilon)T}, \quad (5.5)$$

and

$$J_m^\epsilon = P^\epsilon \left( CT \sup_{0 \leq k \leq m - 1} (\tau_{k+1}^t - \tau_k^t) \geq \eta \right) P^\epsilon (T - \tau_m^t \geq \eta, \tau_m^t < T \leq \tau_{m+1}^t) \quad (5.6)$$

Using the fact that $\tau_m^t$ obeys the Gamma distribution $\text{Gamma}(m, \lambda(\epsilon))$, we deduce

$$P^\epsilon (\tau_m^t < T \leq \tau_{m+1}^t) = E^\epsilon \left[ P^\epsilon (\tau_m^t < T \leq \tau_{m+1}^t | \tau_m^t < T \leq \tau_{m+1}^t - \tau_m^t) \right]$$

$$= \int_0^T P^\epsilon (\tau_m^t < T \leq \tau_m^t + t) \lambda(\epsilon) e^{-\lambda(\epsilon)t} dt + \int_T^\infty P^\epsilon (\tau_m^t < T) \lambda(\epsilon) e^{-\lambda(\epsilon)t} dt$$

$$= \left( \int_0^T e^{-\lambda(\epsilon)x} \frac{\lambda(\epsilon)^m x^{m-1} e^{-\lambda(\epsilon)x}}{(m-1)!} dx \right) \int_T^\infty \frac{\lambda(\epsilon)^m e^{-\lambda(\epsilon)x}}{m!} dx$$

$$= e^{-\lambda(\epsilon)T} \frac{\lambda(\epsilon)^m T^m}{m!}, \quad (5.7)$$

and similarly,

$$P^\epsilon (CT(T - \tau_m^t) \geq \eta, \tau_m^t < T \leq \tau_{m+1}^t)$$

$$= \begin{cases} 1 & \{CT^2 \geq \eta\} \quad (5.8) \\ P^\epsilon (\tau_m^t + \eta/(CT) < T \leq \tau_{m+1}^t) \\ = 1_{\{CT^2 \geq \eta\}} e^{-\lambda(\epsilon)T} \frac{\lambda(\epsilon)^m (T - \frac{\eta}{CT})^m}{m!}. \end{cases}$$
Using again the fact that the sequence \( \{\tau_{k+1} - \tau_k\}_{k=0}^{\infty} \) is i.i.d. with exponential distribution,

\[
P^\varepsilon \left( C T \sup_{0 \leq k \leq m-1} (\tau_{k+1} - \tau_k) < \eta \right) = \prod_{k=0}^{m-1} P^\varepsilon (C T (\tau_{k+1} - \tau_k) < \eta) = \left( 1 - e^{-\lambda(\varepsilon) \frac{\eta}{C T}} \right)^m.
\] 

(5.9)

Combining (5.4)–(5.9), we get

\[
P^\varepsilon \left( \sup_{0 \leq t \leq T} |\gamma^e_t - \gamma_t| \geq \eta \right) \leq 1 \{CT^2 \geq \eta\} e^{-\lambda(\varepsilon)T} \sum_{m=0}^{\infty} \frac{\lambda(\varepsilon)^m (T - \frac{\eta}{C T})^m}{m!} \left( 1 - e^{-\lambda(\varepsilon) \frac{\eta}{C T}} \right)^m
\]

\[+ e^{-\lambda(\varepsilon)T} \sum_{m=1}^{\infty} \frac{\lambda(\varepsilon)^m T^m}{m!} \left[ 1 - \left( 1 - e^{-\lambda(\varepsilon) \frac{\eta}{C T}} \right)^m \right] \]

\[
\leq 1 \{CT^2 \geq \eta\} \exp \left\{ -\lambda(\varepsilon) \left( \frac{\eta}{C T} + Te^{-\lambda(\varepsilon) \frac{\eta}{C T}} - \frac{\eta}{C T} e^{-\lambda(\varepsilon) \frac{\eta}{C T}} \right) \right\}
\]

\[+ \left( 1 - \exp \left\{ -\lambda(\varepsilon) Te^{-\lambda(\varepsilon) \frac{\eta}{C T}} \right\} \right).
\] 

(5.10)

We assume that \( \lim_{\varepsilon \to 0} \varepsilon \lambda(\varepsilon) = \infty \), that is,

\[
\lim_{\varepsilon \to 0} \varepsilon \nu(0, \infty) = \infty.
\]

In particular, \( \lim_{\varepsilon \to 0} \lambda(\varepsilon) = \infty \). Now we take \( \varepsilon \log \) on both sides of (5.10). It is obvious that the \( \varepsilon \log \) of the first term behind the last inequality of (5.10) goes to \( -\infty \) as \( \varepsilon \to 0 \). For the second term, we use Taylor’s theorem to obtain that there exists a \((0,1)\)-valued function \( \theta(\varepsilon) \) such that

\[
\varepsilon \log \left( 1 - \exp \left\{ -\lambda(\varepsilon) Te^{-\lambda(\varepsilon) \frac{\eta}{C T}} \right\} \right) = \varepsilon \log \left( \lambda(\varepsilon) Te^{-\lambda(\varepsilon) \frac{\eta}{C T}} \exp \left\{ -\theta(\varepsilon) \lambda(\varepsilon) Te^{-\lambda(\varepsilon) \frac{\eta}{C T}} \right\} \right)
\]

\[= \varepsilon \log \lambda(\varepsilon) + \varepsilon \log T - \varepsilon \lambda(\varepsilon) \frac{\eta}{C T} - \varepsilon \theta(\varepsilon) \lambda(\varepsilon) Te^{-\lambda(\varepsilon) \frac{\eta}{C T}},
\]

which goes to \( -\infty \) as \( \varepsilon \to 0 \). Therefore,

\[
\limsup_{\varepsilon \to 0} \varepsilon \log P^\varepsilon \left( \sup_{0 \leq t \leq T} |\gamma^e_t - \gamma_t| \geq \eta \right) = -\infty.
\]

By Lemma 5.2, the family \( \{\gamma^e\}_{\varepsilon > 0} \) satisfies the LDP with good rate function (5.3).

\[\square\]

**Acknowledgements.** The research of J. Duan was partly supported by the NSF grant 1620449. The research of Q. Huang was partly supported by NSFC grants 11531006 and 11771449.

**References**

[1] A.A. Agrachev and Y. Sachkov. *Control theory from the geometric viewpoint*, vol. 87. Springer-Verlag Berlin Heidelberg, 2004.

[2] J. Angst, I. Bailleul, and C. Tardif. Kinetic Brownian motion on Riemannian manifolds. *Electron. J. Probab.*, 20, 2015.

[3] J. Birrell, S. Hottovy, G. Volpe, and J. Wehr. Small mass limit of a Langevin equation on a manifold. *Ann. Henri Poincaré*, 18:707–755, 2017.
[4] J.-M. Bismut. *Large deviations and the Malliavin calculus*, vol. 45. Birkhauser Boston Inc., 1984.

[5] É Cartan. *La géométrie des espaces de Riemann*, mémorial des sciences mathématiques. *Gauthier-Villars*, 9:1–61, 1925.

[6] S.A. Chaplygin. On some feasible generalization of the theorem of area, with an application to the problem of rolling spheres (in Russian). *Mat. Sbornik*, XX:1–32, 1897.

[7] S.A. Chaplygin. On the rolling of a sphere on a horizontal plane (in Russian). *Mat. Sbornik*, XXIV:139–168, 1903.

[8] J.B. Conway. *A course in functional analysis*, vol. 96. Springer-Verlag New York, Inc, 1990.

[9] A. Dembo and O. Zeitouni. *Large deviations techniques and applications*, vol. 38. Springer-Verlag New York, Inc, 1998.

[10] P. Dupuis and R.S. Ellis. *A weak convergence approach to the theory of large deviations*, vol. 902. John Wiley & Sons, Inc, 2011.

[11] A. Ganguly. Large deviation principle for stochastic integrals and stochastic differential equations driven by infinite-dimensional semimartingales. *Stoch. Process. Their Appl.*, 128(7): 2179–2227, 2018.

[12] J. Garcia. A large deviation principle for stochastic integrals. *J. Theor. Probab.*, 21(2):476–501, 2008.

[13] E.P. Hsu. *Stochastic analysis on manifolds*, vol. 38. American Mathematical Society, 2002.

[14] N. Ikeda and S. Watanabe. *Stochastic differential equations and diffusion processes, second edition*, vol. 24. North-Holland Publishing Company, 1989.

[15] J. Jacod and A. Shiryaev. *Limit theorems for stochastic processes, second edition*, vol. 288. Springer-Verlag Berlin Heidelberg, 2013.

[16] I. Karatzas and S. Shreve. *Brownian motion and stochastic calculus*, vol. 113. Springer-Verlag New York, Inc., 1991.

[17] R.C. Kraaij, F. Redig, and R. Versendaal. Classical large deviation theorems on complete Riemannian manifolds. *Stoch. Process. Their Appl.*, 2018.

[18] X.-M. Li. Random perturbation to the geodesic equation. *Ann. Probab.*, 44(1):544–566, 2016.

[19] M.G. Molina and E. Grong. Geometric conditions for the existence of a rolling without twisting or slipping. *Commun. Pure Appl. Anal.*, 13:435–452, 2014.

[20] R. Montgomery. *A tour of subriemannian geometries, their geodesics and applications*. Number 91. American Mathematical Society, 2002.

[21] A. Puhalskii. On functional principle of large deviations. In *Proceedings of the Bakuriani Colloquium in Honour of Yu. V. Prohorov: Bakuriani, Georgia, USSR, 24 February–4 March, 1990*, vol. 1, pp. 198. VSP, 1991.

[22] A Puhalskii. Large deviations of semimartingales via convergence of the predictable characteristics. *Stochastics and Stochastics Reports*, 49:27–85, 1994.
[23] K.-I. Sato. *Lévy processes and infinitely divisible distributions*. Cambridge university press, 1999.

[24] R.W. Sharpe. *Differential geometry: Cartan’s generalization of Klein’s Erlangen program*, vol. 166. Springer-Verlag New York, Inc., 1997.

[25] D.W. Stroock. *An introduction to the analysis of paths on a Riemannian manifold*. Number 74. American Mathematical Society, 2000.