Publ. RIMS Kyoto Univ. **60** (2024), 145–184
DOI 10.4171/PRIMS/60-1-4

Large Deviations for Small Noise Hypoelliptic Diffusion Bridges on Sub-Riemannian Manifolds

by

Yuzuru Inahama

Abstract

In this paper we study a large deviation principle of Freidlin–Wentzell type for pinned hypoelliptic diffusion measures associated with a natural sub-Laplacian on a compact sub-Riemannian manifold. To prove this large deviation principle, we use rough path theory and manifold-valued Malliavin calculus.

Mathematics Subject Classification 2020: 60F10 (primary); 53C17, 58J65, 60H07, 60L90 (secondary).
Keywords: large deviation principle, sub-Riemannian geometry, pinned diffusion process, Malliavin calculus, rough path theory.

§1. Introduction

In the theory of stochastic differential equations (SDEs), small noise problems for SDEs are considered very important and have been studied intensively and extensively. A large deviation principle (LDP) associated with them is called Freidlin–Wentzell's LDP. One of its typical formulations is as follows. Let $\mathcal{M}$ be a Euclidean space or a manifold and let V_i, $0 \le i \le d$, be sufficiently nice vector fields on $\mathcal{M}$. For a standard d-dimensional Brownian motion $(w_t)_{0\le t\le 1}$, consider the following Stratonovich-type SDE:

$$dX_t^\varepsilon = \varepsilon \sum_{i=1}^{d} V_i(X_t^\varepsilon) \circ dw_t^i + \varepsilon^2 V_0(X_t^\varepsilon)\, dt, \quad X_0^\varepsilon = x.$$

Here, $x \in \mathcal{M}$ is a given initial point and $0 < \varepsilon \ll 1$ is a small parameter. Note that $X^\varepsilon = (X_t^\varepsilon)_{0\le t\le 1}$ is the diffusion process associated with the generator

Communicated by K. Nakanishi. Received November 8, 2021. Revised April 6, 2022.

Y. Inahama: Faculty of Mathematics, Kyushu University, 744 Motooka, Nishi-ku, Fukuoka 819-0395, Japan;
e-mail: inahama@math.kyushu-u.ac.jp

$\varepsilon^2(\frac{1}{2}\sum_{i=1}^d V_i^2 + V_0)$ and the starting point x. Then, as is well known, the law of X^ε satisfies an LDP as $\varepsilon \searrow 0$.

Let us consider the case that X_t^ε has a (sufficiently nice) strictly positive density with respect to a reference measure on $\mathcal{M}$ (e.g. the Lebesgue measure when $\mathcal{M}$ is a Euclidean space). Then the pinned diffusion process from x to a associated with the above generator exists, where a is a given end point. It seems quite natural to ask whether an LDP of Freidlin–Wentzell type holds for these scaled pinned diffusion measures as $\varepsilon \searrow 0$. In this work we take up this problem.

Although there are a large number of papers on the standard version of Freidlin–Wentzell's LDP, only a few papers have been published on this type of LDP. To the author's knowledge, the first one is Hsu [12]. He proved a pinned version of Freidlin–Wentzell's LDP for pinned Brownian motion on a compact Riemannian manifold (i.e. the pinned diffusion process associated with one half of the Laplace–Beltrami operator). Then, by studying SDEs under a suitable bracket-generating condition on the coefficient vector fields, the author [15, 16] proved this type of LDP in the Euclidean setting for rather general pinned diffusion processes. His method is a combination of rough path theory and quasi-sure analysis, which is a potential-theoretic part of Malliavin calculus. Also, Bailleul [3] proved this type of LDP on a compact manifold for pinned diffusions associated with the sum-of-square-type generator as above with a suitable bracket-generating condition on these vector fields. He combined a probabilistic method (rough path theory) and an analytic method (Sanchez-Calle's estimate for the semigroup generated by a sum-of-square-type operator).

The purpose of this paper is to prove an analogous LDP for the pinned diffusion process on a sub-Riemannian manifold $\mathcal{M}$ associated with the generator $\varepsilon^2(\Delta_{\text{sub}}/2+V)$, where Δ_{sub} is a natural "div-grad-type" sub-Laplacian on $\mathcal{M}$ and V is an arbitrary smooth vector field on $\mathcal{M}$. Our proof is basically similar to those in the preceding works [15, 16]. However, there are two new ingredients in this work. First, in order to realize the $\varepsilon^2(\Delta_{\text{sub}}/2 + V)$-diffusion process via an SDE, we use Eells–Elworthy's construction on a frame bundle over $\mathcal{M}$. Second, since we work on the manifold $\mathcal{M}$ and its frame bundle, we need manifold-valued Malliavin calculus developed in Taniguchi [28].

The organization of this paper is as follows. In Section 2 we formulate our LDP precisely and then state our main result (Theorem 2.1). Section 3 is devoted to recalling the stochastic parallel transport over a sub-Riemannian manifold. Stochastic tools such as Malliavin calculus, quasi-sure analysis, and rough path theory are collected in Section 4. In Section 5 we provide an LDP for the rough path lifts of certain positive Watanabe distributions on the geometric rough path space (Theorem 5.1). Our main result is almost immediate from this, thanks to Lyons' continuity theorem. The lower estimate of the LDP in Theorem 5.1 is proved in Section 6, while the upper estimate is proved in Section 7. In the appendix, we show the strict positivity of the heat kernel associated with $\varepsilon^2(\Delta_{\mathrm{sub}}/2+V)$. This fact ensures the well-definedness of the pinned diffusion measure.

Throughout this paper we will use the following notation:

- $\mathbb{N}_+ = \{1,2,3,\dots\}$ and $\mathbb{N} = \{0\}\cup\mathbb{N}_+$. The set of all real numbers is denoted by $\mathbb{R}$.
- The time interval of (rough) paths is basically $[0,1]$ unless otherwise specified.
- Let U be an open set of a manifold $\mathcal{N}$ and $\mathcal{V}$ be a vector (or a fiber) bundle over U. We denote by $\Gamma(U;\mathcal{V})$ the set of all smooth sections of $\mathcal{V}$ on U. When $U=\mathcal{N}$, we will often simply write $\Gamma(\mathcal{V})$. For instance, we will write $\Gamma(T\mathcal{N})$ or $\Gamma(\mathcal{N};T\mathcal{N})$ for the set of all smooth vector fields on $\mathcal{N}$.
- For a manifold $\mathcal{N}$, $y\in\mathcal{N}$, and a subbundle $\mathcal{V}$ of $T\mathcal{N}$ with a metric, we set
$$\begin{aligned}\mathcal{H}_y(\mathcal{N},\mathcal{V}) = \Bigl\{\gamma\colon [0,1]\to\mathcal{N} \Bigm|\ & \text{absolutely continuous}, \gamma_0 = y,\\ & \gamma'_t\in\mathcal{V}_{\gamma_t} \text{ for almost all } t,\\ & \mathcal{E}(\gamma) := \textstyle\int_0^1 |\gamma'_t|^2_{\mathcal{V}}\,dt<\infty\Bigr\}.\end{aligned}$$
We call $\mathcal{E}(\gamma)$ the energy of γ.
- The Cameron–Martin space $\mathcal{H}^d$ over $\mathbb{R}^d$ is a real Hilbert space defined by
$$\begin{aligned}\mathcal{H}^d = \Bigl\{h\colon [0,1]\to\mathbb{R}^d \Bigm|\ & \text{absolutely continuous}, h_0 = 0,\\ & \text{with } \|h\|^2_{\mathcal{H}^d} := \textstyle\int_0^1 |h'_t|^2_{\mathbb{R}^d}\,dt<\infty\Bigr\}.\end{aligned}$$
In other words, $\mathcal{H}^d = \mathcal{H}_0(\mathbb{R}^d, T\mathbb{R}^d)$ and $\mathcal{E}(h) = \|h\|^2_{\mathcal{H}^d}$. When d is obvious from the context, we simply write $\mathcal{H}$.
- The classical d-dimensional Wiener space $(\mathcal{W},\mathcal{H},\mu) = (\mathcal{W}^d,\mathcal{H}^d,\mu^d)$ is defined as follows: (1) $\mathcal{H}$ is a Cameron–Martin Hilbert space as above. (2) $\mathcal{W}$ is the Banach space of all continuous paths from $[0,1]$ to $\mathbb{R}^d$ which start at the origin. The topology of $\mathcal{W}$ is that of uniform convergence as usual. (3) μ is the d-dimensional Wiener measure. A generic element of $\mathcal{W}$ is denoted by w. The coordinate process $(w_t)_{0\le t\le 1}$ defined on $(\mathcal{W},\mu)$ is called the canonical realization of d-dimensional Brownian motion.
- Let $Z_1,\dots,Z_m$ be smooth vector fields on an open subset U of a manifold $\mathcal{N}$. We set $\Sigma^1 = \{Z_1,\dots,Z_m\}$ and $\Sigma^k = \{[Z_i,Y] \mid Y\in\Sigma^{k-1}\}$ for $k\ge 2$,

recursively. We next set, for $k \geq 1$,

$$\mathrm{Lie}^{(k)}(Z_1, \ldots, Z_m) = \mathrm{span}\left[\bigcup_{j=1}^{k} \Sigma^j\right] \quad \text{and} \quad \mathrm{Lie}(Z_1, \ldots, Z_m) = \mathrm{span}\left[\bigcup_{j=1}^{\infty} \Sigma^j\right].$$

Here, span A means the linear span of A. For $x \in U$, we set

$$\mathrm{Lie}^{(k)}(Z_1, \ldots, Z_m)(x) = \{Y(x) \mid Y \in \mathrm{Lie}^{(k)}(Z_1, \ldots, Z_m)\} \subset T_x\mathcal{N}$$

and also set $\mathrm{Lie}(Z_1, \ldots, Z_m)(x) \subset T_x\mathcal{N}$ in the same way.

As for the last item, one should note the following simple fact. Suppose that there is a smooth function $F \colon U \to \mathrm{GL}(m, \mathbb{R})$ such that

$$[Z_1, \ldots, Z_m] = [\widehat{Z}_1, \ldots, \widehat{Z}_m]F \quad \text{on } U.$$

Then, for every $x \in U$, we have $\mathrm{Lie}^{(k)}(Z_1, \ldots, Z_m)(x) = \mathrm{Lie}^{(k)}(\widehat{Z}_1, \ldots, \widehat{Z}_m)(x)$ for every $k \geq 1$ and therefore $\mathrm{Lie}(Z_1, \ldots, Z_m)(x) = \mathrm{Lie}(\widehat{Z}_1, \ldots, \widehat{Z}_m)(x)$.

§2. Setting and main result

First we recall the essentials of sub-Riemannian geometry following [24]. We say that $(\mathcal{M}, \mathcal{D}, g)$ is a sub-Riemannian manifold if (i) $\mathcal{M}$ is a connected, smooth manifold of dimension n, (ii) $\mathcal{D} \subset T\mathcal{M}$, $T\mathcal{M}$ being the tangent bundle of $\mathcal{M}$, is a smooth distribution of constant rank d ($1 \leq d \leq n$) which satisfies the Hörmander condition at every $x \in \mathcal{M}$, and (iii) $g = (g_x)_{x \in \mathcal{M}}$, where each g_x is an inner product on the fiber $\mathcal{D}_x$, and $x \mapsto g_x$ is smooth. (When there is no risk of confusion, we simply say that $\mathcal{M}$ is a sub-Riemannian manifold.) When $n = d$, this definition coincides with that of a (connected) Riemannian manifold. Throughout this paper $\mathcal{M}$ is assumed to be compact.

The precise statement of the Hörmander condition on $\mathcal{D}$ at $x \in \mathcal{M}$ is as follows: If $\{Z_1, \ldots, Z_d\}$ is a local frame of $\mathcal{D}$ over a coordinate neighborhood U around x, then $\mathrm{Lie}(Z_1, \ldots, Z_d)(x) = T_x\mathcal{M}$. As is well known, this condition does not depend on the choice of U and $\{Z_1, \ldots, Z_d\}$.

Now we recall a "div-grad-type" sub-Laplacian on a sub-Riemannian manifold $\mathcal{M}$. Let **vol** be a smooth volume on $\mathcal{M}$, that is, **vol** is a measure on $\mathcal{M}$ whose restriction to every local coordinate chart is written as a strictly positive smooth density function times the Lebesgue measure on the chart. We consider the second-order differential operator of the form $\Delta_{\mathrm{sub}} = \mathrm{div} \circ \mathrm{grad}_{\mathcal{D}}$, where $\mathrm{grad}_{\mathcal{D}}$ is the horizontal gradient in the direction of $\mathcal{D}$ and div is the divergence with respect to **vol** (i.e. $\mathrm{div} = -(\mathrm{grad}_{\mathcal{D}})^*$ at least formally, where the adjoint is taken with respect to **vol**).

A continuous path $\gamma\colon [0,1] \to \mathcal{M}$ is said to be an admissible path[1] if γ is absolutely continuous, $\gamma'_t \in \mathcal{D}_{\gamma_t}$ for almost all $t \in [0,1]$, and

$$\mathcal{E}(\gamma) := \int_0^1 |\gamma'_t|^2_{g_{\gamma_t}}\, dt < \infty. \tag{2.1}$$

We call $\mathcal{E}(\gamma)$ the energy of γ. (If γ is not admissible, we set $\mathcal{E}(\gamma) = +\infty$ by convention.) By Chow–Rashevsky's theorem [24, Thm. 1.14], for every $x, y \in \mathcal{M}$ there exists an admissible path γ such that $\gamma_0 = x$ and $\gamma_1 = y$.

We define $d_{\mathrm{SR}}\colon \mathcal{M} \times \mathcal{M} \to [0,\infty)$ by

$$d_{\mathrm{SR}}(x,y) = \inf\bigl\{\textstyle\int_0^1 |\gamma'_t|_{g_{\gamma_t}}\, dt \bigm| \gamma\colon [0,1] \to \mathcal{M}, \text{ admissible with } \gamma_0 = x, \gamma_1 = y\bigr\}.$$

Then $d_{\mathrm{SR}}(x,y) < \infty$ for every $x, y \in \mathcal{M}$. It is well known that d_{SR} becomes a distance on $\mathcal{M}$, which generates the same topology as the original manifold topology of $\mathcal{M}$. This is called the sub-Riemannian distance of $\mathcal{M}$. According to [24, Prop. 2.1], it holds that

$$d_{\mathrm{SR}}(x,y)^2 = \inf\{\mathcal{E}(\gamma) \mid \gamma\colon [0,1] \to \mathcal{M}, \text{ admissible with } \gamma_0 = x, \gamma_1 = y\}. \tag{2.2}$$

Let V be any smooth vector field on $\mathcal{M}$ (i.e. $V \in \Gamma(T\mathcal{M})$) and $\varepsilon \in (0,1]$. Then the diffusion process on $\mathcal{M}$ associated with $\varepsilon^2(\Delta_{\mathrm{sub}}/2 + V)$ starting at $x \in \mathcal{M}$ has a density $p^\varepsilon_t(x,a)$ with respect to $\mathbf{vol}(da)$ at every $t > 0$. In fact, $p^\varepsilon_t(x,a) > 0$ for all $x, a \in \mathcal{M}$ and $t, \varepsilon \in (0,1]$ and $a \mapsto p^\varepsilon_t(x,a)$ is smooth for all $x \in \mathcal{M}$ and $t, \varepsilon \in (0,1]$.

The pinned diffusion measure $\mathbb{Q}^\varepsilon_{x,a}$ associated with $\varepsilon^2(\Delta_{\mathrm{sub}}/2 + V)$ from x to a is a unique probability measure on $C([0,1],\mathcal{M})$, the continuous path space over $\mathcal{M}$, such that the following holds: for every $k \geq 1$, $0 = t_0 < t_1 < \cdots < t_k < t_{k+1} = 1$, and $G \in C^\infty(\mathcal{M}^k)$,

$$\begin{aligned}\int_{C([0,1],\mathcal{M})} & G(\xi_{t_1},\ldots,\xi_{t_k})\mathbb{Q}^\varepsilon_{x,a}(d\xi) \\ &= p^\varepsilon_1(x,a)^{-1} \int_{\mathcal{M}^k} G(x_1,\ldots,x_k) \prod_{i=0}^k p^\varepsilon_{t_{i+1}-t_i}(x_i,x_{i+1}) \prod_{i=1}^k \mathrm{vol}(dx_i),\end{aligned}$$

with the convention that $x_0 = x$ and $x_{k+1} = a$. (We will see that $\mathbb{Q}^\varepsilon_{x,a}$ does exist.) Here, $C([0,1],\mathcal{M})$ is the set of all continuous paths on $\mathcal{M}$ equipped with the compact-open topology. The closed subset of all continuous paths which start at x and end at a is denoted by $C_{x,a}([0,1],\mathcal{M})$, in which $\mathbb{Q}^\varepsilon_{x,a}$ is supported.

[1] It is called a *horizontal* path in [24]. We avoid this term, however, because the term "horizontal" is also used in the theory of connections on a principal bundle.

Now we state our main theorem. This can be viewed as a version of Freidlin–Wentzell-type LDP for pinned hypoelliptic diffusion processes on a sub-Riemannian manifold. As one can easily expect, the rate function equals one-half of the energy functional on the path space (up to an additive constant). This theorem includes the main theorem of [12] as a special case. We will prove it in Section 5 as a simple application of Theorem 5.1. The goodness of the rate function J defined by (2.3) below is part of our claim. Note that J actually attains its minimum 0 because of its goodness and (2.2).

Theorem 2.1. *Let $(\mathcal{M}, \mathcal{D}, g)$ be a compact sub-Riemannian manifold with a smooth volume* **vol**. *For $x, a \in \mathcal{M}$, $V \in \Gamma(T\mathcal{M})$, and $\varepsilon \in (0,1]$, let $\mathbb{Q}^{\varepsilon}_{x,a}$ be the pinned diffusion measure as above. Then $\{\mathbb{Q}^{\varepsilon}_{x,a}\}_{0<\varepsilon\le 1}$ satisfies an LDP on $C_{x,a}([0,1],\mathcal{M})$ as $\varepsilon \searrow 0$ with speed ε^{-2} and good rate function $J\colon C_{x,a}([0,1],\mathcal{M}) \to [0,\infty]$, where*

$$J(\gamma) = \frac{1}{2}\{\mathcal{E}(\gamma) - d_{\mathrm{SR}}(x,a)^2\}. \tag{2.3}$$

Here, $\mathcal{E}(\gamma)$ is the energy of γ defined by (2.1).

Remark 2.2. Take any distance d on $\mathcal{M}$ which generates the topology of $\mathcal{M}$. Then

$$\mathrm{dist}(\gamma, \widehat{\gamma}) := \sup_{0\le t\le 1} d(\gamma_t, \widehat{\gamma}_t)$$

defines a distance on $C([0,1],\mathcal{M})$. It is known that dist generates the compact-open topology of $C([0,1],\mathcal{M})$ regardless of the choice of d. Typical examples of d include (1) the sub-Riemannian distance d_{SR} on $\mathcal{M}$, (2) the Riemannian distance on $\mathcal{M}$ with respect to any Riemannian metric on $\mathcal{M}$, (3) the Euclidean distance of $\mathbb{R}^m$ restricted to $\mathcal{M}$ for any embedding $\mathcal{M} \hookrightarrow \mathbb{R}^m$.

Remark 2.3. In Theorem 2.1 above and Proposition 2.4 below, the topology of the path spaces can be slightly strengthened as we now explain. In this remark, $\alpha \in (1/3, 1/2)$.

Let $\mathcal{N}$ be a compact smooth manifold and let $\iota\colon \mathcal{N} \hookrightarrow \mathbb{R}^m$ be an embedding for some $m \in \mathbb{N}_+$. We denote by $C^{\alpha-H}([0,1],\mathbb{R}^m)$ the set of all α-Hölder continuous paths taking values in $\mathbb{R}^m$ and define

$$C^{\alpha-H}([0,1],\mathcal{N}) := C^{\alpha-H}([0,1],\mathbb{R}^m) \cap C([0,1],\mathcal{N}),$$

whose distance is the restriction of the natural one of $C^{\alpha-H}([0,1],\mathbb{R}^m)$.

Let $\mathcal{N}'$ be another compact smooth manifold and let $\iota'\colon \mathcal{N}' \hookrightarrow \mathbb{R}^{m'}$ be an embedding for some $m' \in \mathbb{N}_+$ and suppose that $\phi\colon \mathcal{N} \to \mathcal{N}'$ is a smooth map. Noting that ϕ extends to a smooth map from $\mathbb{R}^m$ to $\mathbb{R}^{m'}$ with compact support,

ϕ naturally induces a continuous map from $C^{\alpha-H}([0,1],\mathcal{N})$ to $C^{\alpha-H}([0,1],\mathcal{N}')$ in an obvious way. In particular, this topology of $C^{\alpha-H}([0,1],\mathcal{N})$ is independent of the choice of ι.

The reason why the LDP in Theorem 2.1 (and in Proposition 2.4 below) also holds on $C^{\alpha-H}([0,1],\mathcal{M})$ is as follows. When we derive Theorem 2.1 from Theorem 5.1, we consider a rough differential equation (RDE) on $\mathcal{P}$, a principal bundle over $\mathcal{M}$, and embed $\mathcal{P}$ into a Euclidean space and then use Lyons' continuity theorem with respect to the α-Hölder rough path topology, together with the contraction principle for LDPs. So, one can easily see that our LDP holds on $C^{\alpha-H}([0,1],\mathcal{M})$ too. (However, we do not write this fact in our main theorem since this set does not look very beautiful from the geometric viewpoint.)

Before closing this section, we claim that our method can re-prove a very similar LDP in [3] on a compact manifold when the generator of the diffusion process is of sum-of-squares type.

Let $\mathcal{M}'$ be a compact smooth manifold and $V_0, V_1, \dots, V_k$ ($k \in \mathbb{N}_+$) be smooth vector fields on $\mathcal{M}'$. Consider the second-order differential operator $\varepsilon^2(V_0 + \frac{1}{2}\sum_{i=1}^k V_i^2)$. Assume the bracket generating condition that $\mathrm{Lie}(V_1,\dots,V_k)(x) = T_x\mathcal{M}'$ for every $x \in \mathcal{M}'$. Then the pinned diffusion measure from x to a associated with this operator exists uniquely, and is denoted by $\widetilde{Q}^{\varepsilon}_{x,a}$.

The following proposition, together with Remark 2.3, is the main result of [3, Thm. 1]. See [3, eq. (4)] for an explicit expression of J', which may not always have a deep geometric meaning.

Proposition 2.4. *Let the notation and assumptions be as above. Then, for every $x, a \in \mathcal{M}'$, $\{\widetilde{Q}^{\varepsilon}_{x,a}\}_{0<\varepsilon\le 1}$ satisfies an LDP on $C_{x,a}([0,1],\mathcal{M}')$ as $\varepsilon \searrow 0$ with speed ε^{-2} and good rate function J'.*

Proof. We can prove this proposition using the same method for Theorem 2.1. The actual proof of this proposition is much simpler than that of Theorem 2.1 since we need not use a principal bundle over $\mathcal{M}'$. So we omit the proof. □

Remark 2.5. Suppose that Δ_{sub} on $\mathcal{M}$ admits a sum-of-squares form in the following sense: there exist $k \ge d$ and $V_i \in \Gamma(T\mathcal{M})$ ($1 \le i \le k$) such that the following condition holds:

$$\Delta_{\mathrm{sub}} - \sum_{i=1}^k V_i^2 \in \Gamma(T\mathcal{M}), \quad \mathrm{Lie}(V_1,\dots,V_k)(x) = T_x\mathcal{M} \text{ for every } x \in \mathcal{M}. \tag{2.4}$$

When Δ_{sub} admits the above expression, it is in a sense true that the LDP in Proposition 2.4 immediately implies the LDP in Theorem 2.1. However, it does not seem easy to obtain the explicit expression (2.3) of the rate function J from J' in this way. The reason is as follows. In this case, k equals the dimension of Brownian motion that plays a key role in the proof and can be very large. Hence, a one-to-one correspondence of Cameron–Martin paths and admissible paths on $\mathcal{M}$ as in Proposition 3.2 breaks down in general. Therefore, even when (2.4) holds, we believe Theorem 2.1 is worth proving. (Unfortunately, the author does not know for which sub-Riemannian manifolds condition (2.4) holds. For instance, for the Laplace–Beltrami operator on a Riemannian manifold, (2.4) is known to be satisfied for sufficiently large k.)

§3. Stochastic parallel transport over a sub-Riemannian manifold

In this section we recall the stochastic parallel transport over a compact sub-Riemannian manifold $(\mathcal{M}, \mathcal{D}, g)$. The history of Eells–Elworthy's construction of Itô's stochastic parallel transport is quite long. The Riemannian case is classical (see [11, Chap. 2] or [14, Sect. V-4] for example), but non-Riemannian cases have also been studied. For diffusion processes associated with semielliptic second-order differential operators, see [7, 8]. For those associated with sub-Laplacians on various kinds of sub-Riemannian manifolds, see [4, 5, 10, 18, 29] among others. In this section we will mainly follow [10].

Now we construct a principal bundle over $\mathcal{M}$ and a connection which are associated with the sub-Riemannian structure of $(\mathcal{M}, \mathcal{D}, g)$. We write $n = \dim \mathcal{M}$ and the rank of $d = \operatorname{rank} \mathcal{D}$ with $1 \leq d \leq n$. Since the Riemannian case (i.e. the case $d = n$) is classical and simpler, we only elaborate the case $1 \leq d < n$ in this section. But the results in this section still hold for the case $d = n$.

First we take a subbundle $\mathcal{D}^\perp = (\mathcal{D}_x^\perp)_{x \in \mathcal{M}}$ of the tangent bundle $T\mathcal{M}$ such that

$$T_x\mathcal{M} = \mathcal{D}_x \oplus \mathcal{D}_x^\perp, \quad x \in \mathcal{M} \tag{3.1}$$

holds. To ensure the existence of such a $\mathcal{D}^\perp$, one just needs to take any Riemannian metric of $\mathcal{M}$ and set $\mathcal{D}_x^\perp$ to be the orthogonal complement of $\mathcal{D}_x$ in $T_x\mathcal{M}$ with respect to the metric. The projections associated with (3.1) are denoted by $\mathbf{pr}_x \colon T_x\mathcal{M} \to \mathcal{D}_x$ and $\mathbf{pr}_x^\perp \colon T_x\mathcal{M} \to \mathcal{D}_x^\perp$, respectively. Then $\mathbf{pr} = (\mathbf{pr}_x)_{x \in \mathcal{M}}$ and $\mathbf{pr}^\perp = (\mathbf{pr}_x^\perp)_{x \in \mathcal{M}}$ belong to $\Gamma(\mathbf{End}(T\mathcal{M}, \mathcal{D}))$ and $\Gamma(\mathbf{End}(T\mathcal{M}, \mathcal{D}^\perp))$, respectively. For any metric h on $\mathcal{D}^\perp$, we set

$$\hat{g}_x\langle v, \hat{v}\rangle = g_x\langle \mathbf{pr}_x v, \mathbf{pr}_x \hat{v}\rangle + h_x\langle \mathbf{pr}_x^\perp v, \mathbf{pr}_x^\perp \hat{v}\rangle, \quad v, \hat{v} \in T_x\mathcal{M}.$$

Then $\hat{g}$ is a Riemannian metric on $\mathcal{M}$ which tames g, that is, the restriction of $\hat{g}_x$ to $\mathcal{D}_x \times \mathcal{D}_x$ equals g_x. Moreover, the decomposition (3.1) is an orthogonal decomposition with respect to $\hat{g}_x$.

Next choose metric Koszul connections $\widetilde{\nabla}$ and $\widetilde{\nabla}^{\perp}$ on $\Gamma(\mathcal{D})$ and $\Gamma(\mathcal{D}^{\perp})$ with respect to the metrics g and h, respectively. Note that they always exist. Define

$$\nabla_X V = \widetilde{\nabla}_X(\mathbf{pr}V) + \widetilde{\nabla}^{\perp}_X(\mathbf{pr}^{\perp}V), \quad X, V \in \Gamma(T\mathcal{M}). \tag{3.2}$$

Since the projections are pointwise operations, ∇ is a $\hat{g}$-metric Koszul connection on $\Gamma(T\mathcal{M})$. It is immediate from the definition that, for all $X \in \Gamma(T\mathcal{M})$, $Y \in \Gamma(\mathcal{D})$ and $Z \in \Gamma(\mathcal{D}^{\perp})$, we have $\nabla_X Y \in \Gamma(\mathcal{D})$, and $\nabla_X Z \in \Gamma(\mathcal{D}^{\perp})$.

Now we introduce a principal bundle over $\mathcal{M}$. We choose any $(\mathcal{D}^{\perp}, h)$ and ∇ (or $\widetilde{\nabla}$ and $\widetilde{\nabla}^{\perp}$) as above and will fix them in what follows. The product $O(d) \times O(n-d)$ of two orthogonal groups naturally acts on it from the right:

$$\begin{aligned} O(\mathcal{M}; \mathcal{D} \oplus \mathcal{D}^{\perp})_x = \{ & u = \xi \oplus \eta \colon \mathbb{R}^d \oplus \mathbb{R}^{n-d} = \mathbb{R}^n \to T_x\mathcal{M} = \mathcal{D}_x \oplus \mathcal{D}^{\perp}_x \\ & \mid \xi \colon \mathbb{R}^d \to \mathcal{D}_x \text{ and } \eta \colon \mathbb{R}^{n-d} \to \mathcal{D}^{\perp}_x \text{ are linear isometries}\}, \\ O(\mathcal{M}; \mathcal{D} \oplus \mathcal{D}^{\perp}) = & \bigsqcup_{x \in \mathcal{M}} O(\mathcal{M}; \mathcal{D} \oplus \mathcal{D}^{\perp})_x. \end{aligned}$$

This is a subbundle of the orthonormal frame bundle over the Riemannian manifold $(\mathcal{M}, \hat{g})$. For notational simplicity we will write $\mathcal{P} := O(\mathcal{M}; \mathcal{D} \oplus \mathcal{D}^{\perp})$ and $G := O(d) \times O(n-d)$. The Lie algebra of G is $\mathfrak{o}(d) \times \mathfrak{o}(n-d)$, which will be denoted by $\mathfrak{g}$. Here, $\mathfrak{o}(d)$ stands for the set of real $d \times d$ anti-symmetric matrices. In the usual way, we view G as a subgroup of $O(n)$ via the injection

$$G \ni (U, V) \mapsto \left(\begin{array}{c|c} U & O \\ \hline O & V \end{array}\right) \in O(n).$$

We denote the natural projection by $\pi \colon \mathcal{P} \to \mathcal{M}$. The right action on $\mathcal{P}$ by $a \in G$ is denoted by R_a. The vertical vector (field) associated with $X \in \mathfrak{g}$ is denoted by X^*, which is defined by $X^*(u) = (d/dt)|_{t=0} R_{\exp(tX)}(u)$ at $u \in \mathcal{P}$.

Let $U \subset \mathcal{M}$ be a local coordinate neighborhood with a local chart $(x^1, \ldots, x^n)$ and let $\{Z_1, \ldots, Z_d\}$ and $\{Z_{d+1}, \ldots, Z_n\}$ be a local orthonormal frame over U of $\mathcal{D}$ and $\mathcal{D}^{\perp}$, respectively. The canonical orthonormal basis of $\mathbb{R}^n$ is denoted by $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$. For $x \in U$, set $\sigma_U(x) \oplus \tilde{\sigma}_U(x) \in \mathcal{P}_x$ by $\sigma_U(x)(\mathbf{e}_i) = Z_i(x)$ for all $1 \le i \le d$ and $\tilde{\sigma}_U(x)(\mathbf{e}_i) = Z_i(x)$ for all $d+1 \le i \le n$. Then, for every $\xi \oplus \eta \in \mathcal{P}_x$, there is a unique $e = \{e_{\alpha\beta}\}_{1 \le \alpha, \beta \le n} \in G$ such that

$$\begin{aligned} \xi \oplus \eta &= (\sigma_U(x) \oplus \tilde{\sigma}_U(x)) \circ e \\ &= (\sigma_U(x) \circ \{e_{\alpha\beta}\}_{1 \le \alpha, \beta \le d}) \oplus (\tilde{\sigma}_U(x) \circ \{e_{\alpha\beta}\}_{d+1 \le \alpha, \beta \le n}). \end{aligned}$$

Of course, $e_{\alpha\beta} = 0$ if (α, β) belongs to

$$(\{1, \ldots, d\} \times \{d+1, \ldots, n\}) \cup (\{d+1, \ldots, n\} \times \{1, \ldots, d\}). \tag{3.3}$$

We can identify $\pi^{-1}(U) \cong U \times G$ with a local chart $(x^1, \ldots, x^n; \{e_{\alpha\beta}\}_{1\le\alpha,\beta\le n})$ in this way.

Next, set $\widehat{\omega} = \{\widehat{\omega}^\alpha_\beta\}_{1\le\beta,\gamma\le n}$ by

$$\nabla Z_\beta = \sum_{\alpha=1}^n \widehat{\omega}^\alpha_\beta Z_\alpha \in \Gamma(U; T^*\mathcal{M} \otimes T\mathcal{M}), \quad 1 \le \beta \le n. \tag{3.4}$$

Since $\widetilde{\nabla}$ and $\widetilde{\nabla}^\perp$ are metric and ∇ is defined by (3.2), we can easily see that $\widehat{\omega}^\beta_\alpha = -\widehat{\omega}^\alpha_\beta$ for all (α, β) and $\widehat{\omega}^\alpha_\beta = 0$ if (α, β) belongs to (3.3). Hence, $\widehat{\omega}$ is an $\mathfrak{g}$-valued one-form on U, i.e. $\widehat{\omega} \in \Gamma(U; T^*\mathcal{M} \otimes \mathfrak{g})$. Unfortunately, $\widehat{\omega}$ does not define a global one-form on $\mathcal{M}$. We need to lift it to obtain a globally defined one-form on the principal bundle $\mathcal{P}$.

Define $\omega \in \Gamma(U \times G; T^*\mathcal{P} \otimes \mathfrak{g})$ as

$$\omega = e^{-1}\widehat{\omega} e + e^{-1}\, de. \tag{3.5}$$

This is a $\mathfrak{g}$-valued one-form on $\pi^{-1}(U)$. Note that $e^{-1}\, de$ stands for Maurer–Cartan form (i.e. the left translation of tangent vectors to the unit element) on G. By a standard argument, we can see that ω actually defines a global one-form on $\mathcal{P}$, namely, we have $\omega \in \Gamma(\mathcal{P}; T^*\mathcal{P} \otimes \mathfrak{g})$. It is easy to see that $R_a^*\omega = a^{-1}\omega a$ for all $a \in G$ and that $\omega(X^*) = X$ for all $X \in \mathfrak{g}$. Thus, ω is an Ehresmann connection on $\mathcal{P}$.

On the above coordinate chart $\pi^{-1}(U)$, we set $\Gamma^\alpha_{\beta\gamma}$ by $\Gamma^\alpha_{\beta\gamma} = \widehat{\omega}^\alpha_\beta\langle Z_\gamma\rangle$, or equivalently,

$$\nabla_{Z_\gamma} Z_\beta = \sum_{\alpha=1}^n \Gamma^\alpha_{\beta\gamma} Z_\alpha, \quad 1 \le \alpha, \beta, \gamma \le n.$$

We can easily see that $\Gamma^\alpha_{\beta\gamma} = 0$ if (α, β) belongs to (3.3) and that $\Gamma^\alpha_{\beta\gamma} = -\Gamma^\beta_{\alpha\gamma}$ for all α, β, γ. For $u = \xi \oplus \eta \in \pi^{-1}(x)$, we define a linear injection $\ell_u \colon T_x\mathcal{M} \to T_u\mathcal{P}$ as

$$\begin{aligned}&\ell_u\Big\langle \sum_{\gamma=1}^n c_\gamma Z_\gamma(x)\Big\rangle \\ &\quad = \sum_{\gamma=1}^n c_\gamma \Big(Z_\gamma(x) - \sum_{\alpha,\beta,\delta=1}^n \Gamma^\alpha_{\beta\gamma}(x) e_{\beta\delta} \frac{\partial}{\partial e_{\alpha\delta}} \Big), \quad (c_1, \ldots, c_n) \in \mathbb{R}^n\end{aligned} \tag{3.6}$$

and set $\mathcal{K}_u = \ell_u\langle \mathcal{D}_x\rangle$ and $\mathcal{K}_u^\perp = \ell_u\langle \mathcal{D}_x^\perp\rangle$. It is known that

$$\ker(\omega_u \colon T_u\mathcal{P} \to \mathfrak{g}) = \mathcal{K}_u \oplus \mathcal{K}_u^\perp.$$

This is called the horizontal subspace and ℓ_u is called the horizontal lift. The subspace $\mathcal{V}_u = \{X^*(u) \mid X \in \mathfrak{g}\}$ is called the vertical subspace and it holds that $T_u\mathcal{P} = \ker \omega_u \oplus \mathcal{V}_u$. (From these explicit expressions too, we can see that the horizontal lift and the horizontal subspaces are compatible with the right action of G.)

We define the canonical horizontal vector fields $A_i = \ell_u\langle u\langle \mathbf{e}_i\rangle\rangle$ $(1 \le i \le n)$ on $\mathcal{P}$. Since $u\langle \mathbf{e}_i\rangle = \sum_{\gamma=1}^n e_{\gamma i} Z_\gamma(x)$, A_i reads

$$A_i(u) = \sum_{\gamma=1}^n e_{\gamma i}\left(Z_\gamma(x) - \sum_{\alpha,\beta,\delta=1}^n \Gamma^\alpha_{\beta\gamma}(x) e_{\beta\delta} \frac{\partial}{\partial e_{\alpha\delta}}\right) \tag{3.7}$$

in the local chart. At every u, $\{A_i(u)\}_{i=1}^d$ and $\{A_i(u)\}_{i=d+1}^n$ are linear bases of $\mathcal{K}_u$ and $\mathcal{K}_u^\perp$, respectively. We equip $\mathcal{K}_u \oplus \mathcal{K}_u^\perp$ with a unique inner product so that $\{A_i(u)\}_{i=1}^n$ becomes an orthonormal basis. Set a linear isometry $\theta_u \colon \mathcal{K}_u \oplus \mathcal{K}_u^\perp \to \mathbb{R}^n$ by $\theta\langle A_i(u)\rangle = \mathbf{e}_i$ for all $1 \le i \le n$. We can naturally view θ as an $\mathbb{R}^n$-valued one-form on $\mathcal{P}$. To summarize, the following four bijective linear maps are all isometric (in either pair of maps, those in the opposite directions are inverse to each other):

$$\mathbb{R}^n = \mathbb{R}^d \oplus \mathbb{R}^{n-d} \underset{\theta_u}{\overset{\theta_u^{-1}}{\rightleftarrows}} \ker \omega_u = \mathcal{K}_u \oplus \mathcal{K}_u^\perp \underset{\ell_u}{\overset{(\pi_*)_u}{\rightleftarrows}} T_x\mathcal{M} = \mathcal{D}_x \oplus \mathcal{D}_x^\perp.$$

By restricting this to the first components, we obtain the isometric correspondence

$$\mathbb{R}^d \underset{\theta_u}{\overset{\theta_u^{-1}}{\rightleftarrows}} \mathcal{K}_u \underset{\ell_u}{\overset{(\pi_*)_u}{\rightleftarrows}} \mathcal{D}_x.$$

Lemma 3.1. *Canonical horizontal vector fields $\{A_1, \ldots, A_d\}$ on $\mathcal{P}$ satisfy the partial Hörmander condition at every $u \in \mathcal{P}$, that is,*

$$\pi_* \operatorname{Lie}(A_1, \ldots, A_d)(u) = T_{\pi(u)}\mathcal{M}, \quad u \in \mathcal{P}.$$

Proof. Let $u \in \pi^{-1}(x)$ and use the local chart as above. Note that

$$[A_1, \ldots, A_d] = [\ell\langle Z_1\rangle, \ldots, \ell\langle Z_d\rangle]E, \quad \text{where } E = (e_{\alpha\beta})_{1\le\alpha,\beta\le d}.$$

Since $E = E(u)$ is a smooth $O(d)$-valued function in u, we have

$$\operatorname{Lie}(A_1, \ldots, A_d)(u) = \operatorname{Lie}(\ell\langle Z_1\rangle, \ldots, \ell\langle Z_d\rangle)(u).$$

Since we have $\ell\langle Z_i\rangle(u) = Z_i(x) + (\text{a vertical vector field})$ from (3.6), we see that

$$[\ell\langle Z_i\rangle, \ell\langle Z_j\rangle](u) = [Z_i, Z_j](x) + (\text{a vertical vector field}).$$

Repeating this, we have

$$\pi_* \operatorname{Lie}(\ell\langle Z_1\rangle, \ldots, \ell\langle Z_d\rangle)(u) = \operatorname{Lie}(Z_1, \ldots, Z_d)(x).$$

Since $\mathcal{M}$ is sub-Riemannian, the right-hand side equals $T_x\mathcal{M}$. □

Now we turn to the (anti-)development of finite energy paths. Let $\mathcal{H}^n$ be the Cameron–Martin space over $\mathbb{R}^n$. For $h \in \mathcal{H}^n$ and $u \in \mathcal{P}$, we consider the controlled ODE (skeleton ODE)

$$d\phi_t = \sum_{i=1}^n A_i(\phi_t)\, dh_t^i, \quad \phi_0 = u, \tag{3.8}$$

and set $\psi_t = \pi(\phi_t)$. To emphasize the dependency on h, we sometimes write $\phi(h)$ and $\psi(h)$. It is clear that $\phi(h) \in \mathcal{H}_u(\mathcal{P}, \mathcal{K} \oplus \mathcal{K}^\perp)$ and $\psi(h) \in \mathcal{H}_{\pi(u)}(\mathcal{M}, T\mathcal{M})$. Moreover, the energy is preserved, that is, $\|h\|_{\mathcal{H}^n}^2 = \mathcal{E}(\phi(h)) = \mathcal{E}(\psi(h))$. It should also be noted that $\psi(h)'_t = \phi(h)_t \langle h'_t \rangle$.

The map $h \mapsto \psi(h)$ is bijective. The inverse of

$$h \in \mathcal{H}^n \mapsto \phi(h) \in \mathcal{H}_u(\mathcal{P}, \mathcal{K} \oplus \mathcal{K}^\perp)$$

is given by the line integral of θ:

$$(u_t)_{t \in [0,1]} \in \mathcal{H}_u(\mathcal{P}, \mathcal{K} \oplus \mathcal{K}^\perp) \mapsto \int_0^\cdot \theta_{u_t} \langle u'_t \rangle \in \mathcal{H}^n.$$

The inverse of the projection $\pi\colon \mathcal{H}_u(\mathcal{P}, \mathcal{K} \oplus \mathcal{K}^\perp) \to \mathcal{H}_{\pi(u)}(\mathcal{M}, T\mathcal{M})$ is the horizontal lift of paths. Recall that the horizontal lift of $(x_t)_{t \in [0,1]} \in \mathcal{H}_{\pi(u)}(\mathcal{M}, T\mathcal{M})$ is a unique $(u_t)_{t \in [0,1]} \in \mathcal{H}_u(\mathcal{P}, \mathcal{K} \oplus \mathcal{K}^\perp)$ such that $u'_t = \ell_{u_t} \langle x'_t \rangle$ for almost all t. Locally, (u_t) satisfies a simple controlled ODE as follows. Suppose that $0 < \tau \le 1$ and that $(x_t)_{t \in [0,\tau]}$ stays in a local chart. Then there exists a unique Cameron–Martin path $(k_t)_{t \in [0,\tau]}$ over $\mathbb{R}^n$ such that

$$dx_t = \sum_{i=1}^n Z_i(x_t) k'_t\, dt = \sum_{i=1}^n Z_i(x_t)\, dk_t \quad \text{on } [0,\tau] \text{ with } x_0 = x = \pi(u).$$

Therefore, the horizontal lift $(u_t)_{t \in [0,\tau]}$ solves the controlled ODE

$$du_t = \sum_{i=1}^n \ell \langle Z_i \rangle (u_t)\, dk_t \quad \text{on } [0,\tau] \text{ with } u_0 = u.$$

The local expression of $\ell\langle Z_i \rangle$ was given in (3.6). Hence, we can write down this ODE for $u_t = (x_t^1, \dots, x_t^n; \{e_{\alpha\beta,t}\}_{1 \le \alpha,\beta \le n})$ concretely using these coordinates. (However, we do not elaborate it because it is well known and cumbersome. The point here is to explain that the lift map is explicitly computable and dependency on the data such as x, k can be tracked.) Thus, we have seen that the development map $h \mapsto \phi(h)$ is bijective and preserves energy.

Restricting this correspondence to the first component, we have the following proposition. Note that we can naturally view $\mathcal{H}^n = \mathcal{H}^d \oplus \mathcal{H}^{n-d}$ as a direct sum of Hilbert spaces.

Proposition 3.2. *For $h \in \mathcal{H}^d$ and $u \in \mathcal{P}$, consider the ODE*

$$d\phi(h)_t = \sum_{i=1}^{d} A_i(\phi(h)_t)\, dh_t^i, \quad \phi(h)_0 = u, \tag{3.9}$$

and set $\psi(h)_t = \pi(\phi(h)_t)$. Then the development map $h \mapsto \psi(h)$ is an energy-preserving bijection from $\mathcal{H}^d$ to $\mathcal{H}_{\pi(u)}(\mathcal{M}, \mathcal{D})$.

We now provide a generalization of Chow–Rashevsky's theorem for future purposes. We will use this in the appendix.

Proposition 3.3. *Let $V \in \Gamma(T\mathcal{M})$. Then, for every $x, y \in \mathcal{M}$ and $\tau \in (0, 1]$, there exists $k \in \mathcal{H}_x(\mathcal{M}, T\mathcal{M})$ such that $k_\tau = y$ and $k_t' - V(k_t) \in \mathcal{D}_{k_t}$ for almost all $t \in [0, \tau]$.*

Proof. We may simply work on $[0, \tau]$ since we can just set k to be constant on $[\tau, 1]$. Note that if we take $l \in \mathbb{N}_+$ large enough, then we can find $B_1, \dots, B_l \in \Gamma(\mathcal{D})$ such that

$$\mathrm{Lie}(B_1, \dots, B_l)(x) = T_x\mathcal{M} \quad \text{for every } x \in \mathcal{M}.$$

Let us consider the following ODE on $\mathcal{M}$ controlled by an l-dimensional Cameron–Martin path $h\colon [0, \tau] \to \mathbb{R}^l$:

$$dk_t = \sum_{i=1}^{l} B_i(k_t)\, dh_t^i + V(k_t)\, dt, \quad k_0 = x.$$

Thanks to the above condition on Lie brackets of the B_i, this controlled ODE is strongly completely controllable. It implies that for every x, y, τ, we can find h such that the solution $k = k(h)$ satisfies $k_\tau = y$ (see [19, Sect. 5] for example). Hence, this solution k is a desired path. □

Define a second-order differential operator $\tilde{\Delta}$ on $\mathcal{M}$ by

$$\tilde{\Delta} f = \mathrm{Trace}_{\mathcal{D}}(\nabla \operatorname{grad}_{\mathcal{D}} f), \quad f \in C^2(\mathcal{M}).$$

The precise meaning is as follows. First, $\nabla \operatorname{grad}_{\mathcal{D}} f \in \Gamma(T^*\mathcal{M} \otimes \mathcal{D})$ and therefore $v \mapsto \nabla_v \operatorname{grad}_{\mathcal{D}} f$ can be viewed as a linear map from $\mathcal{D}_x$ to itself at every $x \in \mathcal{D}$. The right-hand side at x is defined to be the trace of this linear map.

Lemma 3.4. *Let the notation be as above. Then $\Delta_{\mathrm{sub}} - \tilde{\Delta} \in \Gamma(\mathcal{D})$.*

Proof. Take a local orthonormal frame $\{Z_1, \dots, Z_d\}$ of $\mathcal{D}$ on a coordinate neighborhood $U \subset \mathcal{M}$. Then it is well known that

$$\Delta_{\mathrm{sub}} f = \sum_{i=1}^{d} \{Z_i^2 f + (\operatorname{div} Z_i) Z_i f\},$$

where div stands for the divergence with respect to the measure **vol**.

On the other hand, since $\mathrm{grad}_{\mathcal{D}} f = \sum_{i=1}^{d}(Z_i f)Z_i$, we see from (3.4) that

$$\tilde{\Delta} f = \sum_{i=1}^{d} Z_i^2 f + \sum_{i,j=1}^{d} \widehat{\omega}_i^j(Z_j)Z_i f$$

and hence

$$\Delta_{\mathrm{sub}} - \tilde{\Delta} = \sum_{i=1}^{d}(\mathrm{div}\, Z_i)Z_i - \sum_{i,j=1}^{d} \widehat{\omega}_i^j(Z_j)Z_i \in \Gamma(U;\mathcal{D}).$$

The left-hand side is a globally defined at most second-order operator. However, as the right-hand side shows, its second-order part vanishes. Hence, this is a globally defined first-order operator, i.e. a vector field. □

Lemma 3.5. *Let the notation be as above. Then we have*

$$\sum_{i=1}^{d} A_i^2(\pi^* f) = \pi^*(\tilde{\Delta} f), \quad f \in C^2(\mathcal{M}).$$

Here, $\pi^* f := f \circ \pi \in C^2(\mathcal{P})$ is the pullback of f by the projection $\pi\colon \mathcal{P} \to \mathcal{M}$.

Proof. We work with the local chart $(x^1, \ldots, x^n; \{e_{\alpha\beta}\}_{1\le\alpha,\beta\le n})$ on $\pi^{-1}(U) \cong U \times G$ as above. We see from the local expression (3.7) that, for $1 \le i, j \le d$,

$$A_i(\pi^* f) = \sum_{\gamma=1}^{d} e_{\gamma i}\pi^*(Z_\gamma f),$$

$$A_j A_i(\pi^* f) = \sum_{\gamma,\delta=1}^{d} e_{\delta j} e_{\gamma i}\pi^*(Z_\delta Z_\gamma f) - \sum_{\gamma,\delta,\varepsilon=1}^{d} \Gamma_{\delta\varepsilon}^{\gamma} e_{\varepsilon j} e_{\delta i}\pi^*(Z_\gamma f).$$

On the other hand, noting that $\{u\langle \mathbf{e}_1\rangle, \ldots, u\langle \mathbf{e}_d\rangle\}$ is an orthonormal basis at $\pi(u)$, we compute as follows:

$$\begin{aligned}\langle \nabla_{u\langle \mathbf{e}_j\rangle} \mathrm{grad}_{\mathcal{D}} f, u\langle \mathbf{e}_i\rangle\rangle_{\mathcal{D}} &= \sum_{\gamma,\delta,\varepsilon=1}^{d} e_{\varepsilon j} e_{\delta i}\langle \nabla_{Z_\varepsilon}(Z_\gamma f)Z_\gamma, Z_\delta\rangle_{\mathcal{D}} \\ &= \sum_{\delta,\varepsilon=1}^{d} e_{\varepsilon j} e_{\delta i} Z_\varepsilon Z_\delta f - \sum_{\gamma,\delta,\varepsilon=1}^{d} e_{\varepsilon j} e_{\delta i}\Gamma_{\delta\varepsilon}^{\gamma} Z_\gamma f \\ &= A_j A_i(\pi^* f)(u),\end{aligned}$$

where we used $\Gamma_{\delta\varepsilon}^{\gamma} = -\Gamma_{\gamma\varepsilon}^{\delta}$. Setting $i = j$ and summing them, we prove the lemma. □

Now we introduce an SDE on $\mathcal{P}$. For $V \in \Gamma(T\mathcal{M})$, we set $A_0 \in \Gamma(T\mathcal{P})$ as

$$A_0(u) = \ell_u \langle V_0(\pi(u)) \rangle, \quad u \in \mathcal{P}, \text{ where } V_0 := V + (\Delta_{\mathrm{sub}} - \tilde{\Delta})/2 \in \Gamma(T\mathcal{M}).$$

Let $(w_t)_{t\in[0,1]}$ be a standard d-dimensional Brownian motion and consider the following Stratonovich-type SDE for $u \in \mathcal{P}$ and $0 < \varepsilon \le 1$:

$$U_t^\varepsilon = \varepsilon \sum_{i=1}^d A_i(U_t^\varepsilon) \circ dw_t^i + \varepsilon^2 A_0(U_t^\varepsilon)\, dt, \quad U_0^\varepsilon = u. \tag{3.10}$$

From the scaling property of Brownian motion, the two processes (U_t^ε) and $(U^1_{\varepsilon^2 t})$ have the same law. We will write $X_t^\varepsilon := \pi(U_t^\varepsilon)$.

Lemma 3.6. *Let $x \in \mathcal{M}$. Choose $u \in \pi^{-1}(x)$ and consider SDE (3.10). Then the law of the process $(X_t^\varepsilon)_{t\in[0,1]}$ is independent of the choice of u and this process is a diffusion process on $\mathcal{M}$ associated with the generator $\varepsilon^2(\Delta_{\mathrm{sub}}/2 + V)$.*

Proof. For the same reason (the rotational invariance of (w_t)) as in the Riemannian case, the law of $(X_t^\varepsilon)_{t\in[0,1]}$ is independent of the choice of u. By the Itô formula, one can show that the generator of (U_t^ε) is $\varepsilon^2(\frac{1}{2}\sum_{i=1}^d A_i^2 + A_0)$. From this and Lemma 3.5, we can easily see that (X_t^ε) is still a diffusion process and its generator is $\varepsilon^2(\Delta_{\mathrm{sub}}/2 + V)$. □

Remark 3.7. For manifold-valued SDEs, the manifold need not be a submanifold of a Euclidean space. However, for manifold-valued RDEs and Malliavin calculus, the manifold is usually embedded into a Euclidean space. Therefore, in what follows we choose an embedding $\iota\colon \mathcal{P} \hookrightarrow \mathbb{R}^M$ and extend the vector fields A_i, $0 \le i \le d$, to C^∞ vector fields on $\mathbb{R}^M$ with compact support. In this way, we may view SDE (3.10) as an SDE on $\mathbb{R}^M$. For our purpose, any M and ι will do. (Recall that, thanks to the tubular neighborhood theorem, every smooth function or vector field on $\mathcal{P}$ extends to smooth function or vector field on $\mathbb{R}^M$ with compact support, respectively.)

§4. Preliminaries from stochastic analysis

In this section we recall several important probabilistic results which we will use in the proofs of our main results. Basically, all results in this section (except Lemma 4.3) are either known or easily derived from known facts.

§4.1. Elements of (manifold-valued) Malliavin calculus

We first recall Watanabe's theory of generalized Wiener functionals (i.e. Watanabe distributions) in Malliavin calculus. Most of the contents and the notation in this section are contained in [14, Sects. V.8–V.10] with trivial modifications. Also, [25, 23, 13, 22] are good textbooks on Malliavin calculus. For manifold-valued Malliavin calculus, see [28]. For basic results of quasi-sure analysis, we refer to [21, Chap. II].

Let $(\mathcal{W}, \mathcal{H}, \mu)$ be the classical d-dimensional Wiener space. (The results in this subsection also hold on any abstract Wiener space, however.) Let us recall the following:

(a) The basics of Sobolev spaces $\mathbf{D}_{p,r}(\mathcal{K})$ of $\mathcal{K}$-valued (generalized) Wiener functionals, where $p \in (1,\infty)$, $r \in \mathbb{R}$, and $\mathcal{K}$ is a real separable Hilbert space. As usual, we will use the spaces $\mathbf{D}_\infty(\mathcal{K}) = \bigcap_{k=1}^\infty \bigcap_{1<p<\infty} \mathbf{D}_{p,k}(\mathcal{K})$, $\widetilde{\mathbf{D}}_\infty(\mathcal{K}) = \bigcap_{k=1}^\infty \bigcup_{1<p<\infty} \mathbf{D}_{p,k}(\mathcal{K})$ of test functionals and the spaces $\mathbf{D}_{-\infty}(\mathcal{K}) = \bigcup_{k=1}^\infty \bigcup_{1<p<\infty} \mathbf{D}_{p,-k}(\mathcal{K})$, $\widetilde{\mathbf{D}}_{-\infty}(\mathcal{K}) = \bigcup_{k=1}^\infty \bigcap_{1<p<\infty} \mathbf{D}_{p,-k}(\mathcal{K})$ of Watanabe distributions as in [14]. When $\mathcal{K} = \mathbb{R}$, we simply write $\mathbf{D}_{p,r}$ etc.

(b) Meyer's equivalence of Sobolev norms. (See [14, Thm. 8.4]. A stronger version can be found in [25, Thm. 4.6].)

(c) Pullback $T \circ F = T(F) \in \widetilde{\mathbf{D}}_{-\infty}$ of tempered Schwartz distribution $T \in \mathcal{S}'(\mathbb{R}^n)$ on $\mathbb{R}^n$ by a non-degenerate Wiener functional $F \in \mathbf{D}_\infty(\mathbb{R}^n)$. (See [14, Sect. 5.9]. In fact, this is very strongly related to item (d) below.)

(d) A generalized version of the integration by parts formula (IbP formula) in the sense of Malliavin calculus for Watanabe distributions, which is given as follows (see [14, p. 377]):

For $F = (F^1, \ldots, F^n) \in \mathbf{D}_\infty(\mathbb{R}^n)$, we denote by $\sigma_F^{ij}(w) = \langle DF^i(w), DF^j(w)\rangle_{\mathcal{H}}$ the (i,j)-component of the Malliavin covariance matrix $(1 \le i, j \le n)$. We denote by $\gamma_F^{ij}(w)$ the (i,j)-component of the inverse matrix σ_F^{-1} (if the inverse exists in a certain sense). Recall that F is called non-degenerate in the sense of Malliavin if $(\det \sigma_F)^{-1} \in \bigcap_{1<p<\infty} L^p$. Note that $\sigma_F^{ij} \in \mathbf{D}_\infty$ and $D\gamma_F^{ij} = -\sum_{k,l} \gamma_F^{ik}(D\sigma_F^{kl})\gamma_F^{lj}$. Hence, the derivatives of γ_F^{ij} can be written in terms of the γ_F^{ij} and the derivatives of the σ_F^{ij}, which implies that $\gamma_F^{ij} \in \mathbf{D}_\infty$ too.

Suppose $G \in \mathbf{D}_\infty$ and $T \in \mathcal{S}'(\mathbb{R}^n)$. Then the following integration by parts holds:

$$\mathbb{E}[\partial_i T(F) \cdot G] = \mathbb{E}[T(F) \cdot \Phi_i(\,\cdot\,; G)], \tag{4.1}$$

where $\Phi_i(w; G) \in \mathbf{D}_\infty$ is given by

$$\Phi_i(w; G) = \sum_{j=1}^d D^*(\gamma_F^{ij} \cdot G \cdot DF^j)(w). \tag{4.2}$$

Note that $\mathbb{E}$ on both sides of (4.1) is in fact the generalized expectation, that is, the pairing of $\widetilde{\mathbf{D}}_{-\infty}$ and $\widetilde{\mathbf{D}}_{\infty}$. Here, D and D^* are the $\mathcal{H}$-derivative (i.e. the gradient operator in the sense of Malliavin calculus) and its adjoint (i.e. the divergence operator) respectively.

(e) If $\eta \in \mathbf{D}_{-\infty}$ satisfies that $\langle \eta, F\rangle \geq 0$ for every non-negative $F \in \mathbf{D}_{\infty}$, it is called a positive Watanabe distribution. According to Sugita's theorem [26], for every positive Watanabe distribution η, there uniquely exists a finite Borel measure μ_η on $\mathcal{W}$ such that

$$\langle \eta, F\rangle = \int_{\mathcal{W}} \widetilde{F}(w)\mu_\eta(dw), \quad F \in \mathbf{D}_{\infty}$$

holds, where $\widetilde{F}$ stands for ∞-quasi-continuous modification of F. If $\eta \in \mathbf{D}_{p,-k}$ is positive, then it holds that

$$\mu_\eta(A) \leq \|\eta\|_{p,-k} \operatorname{Cap}_{q,k}(A) \quad \text{for every Borel subset } A \subset \mathcal{W},$$

where $p, q \in (1,\infty)$ with $1/p + 1/q = 1$, $k \in \mathbb{N}_+$, and $\operatorname{Cap}_{q,k}$ stands for the (q,k)-capacity associated with $\mathbf{D}_{q,k}$. (For more details, see [21, Chap. II].)

We will also use a localized version of the Watanabe distribution theory, which can be found in [27, pp. 216–217]. (For proofs, see [17, Props. 3.1 and 3.2].)

Let $\rho > 0$, $\xi \in \mathbf{D}_{\infty}$, and $F \in \mathbf{D}_{\infty}(\mathbb{R}^n)$ and suppose that

$$\inf_{v\in\mathbb{S}^{n-1}} v^*\sigma_F v \geq \rho \quad \text{on } \bigl\{w \in \mathcal{W} \bigm| |\xi(w)| \leq 2\bigr\}, \tag{4.3}$$

where $\mathbb{S}^{n-1}$ is the unit ball of $\mathbb{R}^n$. Let $\chi\colon \mathbb{R} \to \mathbb{R}$ be a smooth function whose support is contained in $[-1,1]$. Then the following proposition holds (see [27, Prop. 6.1]).

Proposition 4.1. *Assume (4.3). For every $T \in \mathcal{S}'(\mathbb{R}^n)$, $\chi(\xi)\cdot T(F) \in \widetilde{\mathbf{D}}_{-\infty}$ can be defined in a unique way so that the following properties hold:*

(i) *If $T_k \to T \in \mathcal{S}'(\mathbb{R}^n)$ as $k \to \infty$, then $\chi(\xi)\cdot T_k(F) \to \chi(\xi)\cdot T(F) \in \widetilde{\mathbf{D}}_{-\infty}$.*

(ii) *If T is given by $g \in \mathcal{S}(\mathbb{R}^n)$, then $\chi(\xi)\cdot T(F) = \chi(\xi)g(F) \in \mathbf{D}_{\infty}$.*

We also provide an asymptotic theorem. It is a very special case of [27, Prop. 6.2]. Let $\{F_\varepsilon\}_{0\leq\varepsilon\leq 1} \subset \mathbf{D}_{\infty}(\mathbb{R}^n)$ and $\{\xi_\varepsilon\}_{0\leq\varepsilon\leq 1} \subset \mathbf{D}_{\infty}$ be families of Wiener functionals such that the following asymptotics hold:

$$F_\varepsilon = F_0 + O(\varepsilon) \quad \text{in } \mathbf{D}_{\infty}(\mathbb{R}^n) \text{ as } \varepsilon \searrow 0, \tag{4.4}$$

$$\xi_\varepsilon = \xi_0 + O(\varepsilon) \quad \text{in } \mathbf{D}_{\infty} \text{ as } \varepsilon \searrow 0. \tag{4.5}$$

Here, $O(\varepsilon)$ is the large Landau symbol. Recall that $\mathbf{D}_{\infty}$ and $\mathbf{D}_{\infty}(\mathbb{R}^n)$ are endowed with a natural topology as Fréchet spaces.

Proposition 4.2. *Assume* (4.4), (4.5), *and* $|\xi_0| \le 1/8$. *Moreover, assume that there exists* $\rho > 0$ *independent of* ε *such that* (4.3) *with* $F = F_\varepsilon$ *and* $\xi = \xi_\varepsilon$ *holds for every* $\varepsilon \in (0,1]$. *Let* $\chi\colon \mathbb{R} \to \mathbb{R}$ *be a smooth function whose support is contained in* $[-1,1]$ *such that* $\chi(x) = 1$ *if* $|x| \le 1/2$. *Then we have*

$$\lim_{\varepsilon \searrow 0} \chi(\xi_\varepsilon) \cdot T(F_\varepsilon) = T(F_0) \quad \text{in } \widetilde{\mathbf{D}}_{-\infty}.$$

More precisely, there exists $k \in \mathbb{N}_+$ *such that the above convergence takes place in* $\mathbf{D}_{p,-k}$ *for every* $p \in (1,\infty)$.

Let us quickly review manifold-valued Malliavin calculus. Malliavin calculus for SDEs on manifolds was developed by Taniguchi [28]. Roughly speaking, under suitable assumptions, almost all important results in the Euclidean case still hold true in the manifold case with natural modifications.

Let $\mathcal{N}$ be a compact manifold of dimension m, which is equipped with a smooth volume $\mathbf{vol}_{\mathcal{N}}$. (A measure on $\mathcal{N}$ is said to be a smooth volume if it is expressed on each coordinate chart as a strictly positive smooth density function times the Lebesgue measure.) Choose a Riemannian metric on $\mathcal{N}$ so that the determinant of the (deterministic) Malliavin covariance of $\mathcal{N}$-valued functionals is well defined. Any choice of the Riemannian metric and the smooth volume will do.

An $\mathcal{N}$-valued Wiener functional $F\colon \mathcal{W} \to \mathcal{N}$ is said to belong to $\mathbf{D}_{p,k}(\mathcal{N})$, $p \in (1,\infty)$ and $k \in \mathbb{N}+$ if $f(F) \in \mathbf{D}_{p,k}$ for every $f \in C^\infty(\mathcal{N})$. If $\iota\colon \mathcal{N} \to \mathbb{R}^M$ is an embedding, then $F \in \mathbf{D}_{p,k}(\mathcal{N})$ holds if and only if $\iota(F) \in \mathbf{D}_{p,k}(\mathbb{R}^M)$ since every $f \in C^\infty(\mathcal{N})$ extends to a smooth function on $\mathbb{R}^M$ with compact support. The same holds true for $F \in \mathbf{D}_\infty(\mathcal{N}) \coloneqq \bigcap_{k=1}^\infty \bigcap_{1<p<\infty} \mathbf{D}_{p,k}(\mathcal{N})$. For $F \in \bigcup_{1<p<\infty} \mathbf{D}_{p,1}(\mathcal{N})$, $D_h F(w) \in T_{F(w)}\mathcal{N}$. Hence, the Malliavin covariance $\sigma_F(w)$ in this case is a symmetric bilinear form on $T^*_{F(w)}\mathcal{N} \times T^*_{F(w)}\mathcal{N}$. Thanks to the Riemannian metric, $\det \sigma_F(w)$ can still be defined.

One of two main results in [28] is as follows. As in the Euclidean case, if $F \in \mathbf{D}_\infty(\mathcal{N})$ is non-degenerate in the sense of Malliavin, i.e. $(\det \sigma_F)^{-1} \in \bigcap_{1<p<\infty} L^p$, then the composition $T(F) = T \circ F \in \widetilde{\mathbf{D}}_{-\infty}$ is well defined as a Watanabe distribution for every distribution T on $\mathcal{N}$. Moreover, the law of F on $\mathcal{N}$ has a smooth density p_F function with respect to $\mathbf{vol}_{\mathcal{N}}$. In particular, $\delta_a(Y_1^\varepsilon)$ is a positive Watanabe distribution and $p_F(a) = \mathbb{E}[\delta_a(F)]$ for every $a \in \mathcal{N}$. One should note here that δ_a and $\delta_a(Y_1^\varepsilon)$ depend on the choice of $\mathbf{vol}_{\mathcal{N}}$. (Since any other smooth volume can be expressed as $\widehat{\mathbf{vol}}_{\mathcal{N}}(dy) = \rho(y)\mathbf{vol}_{\mathcal{N}}(dy)$ for some strictly positive smooth function ρ on $\mathcal{N}$, the delta function with respect to $\widehat{\mathbf{vol}}_{\mathcal{N}}$ is given by $\hat{\delta}_a = \rho(a)^{-1}\delta_a$.) The other main result in [28] is proving non-degeneracy for the projected solution of an SDE whose coefficient vector fields satisfy the partial Hörmander condition.

From here we consider SDE (3.10) (with $x \in \mathcal{M}$ and $u \in \pi^{-1}(x)$) and set $\mathcal{N}$ to be either $\mathcal{M}$ or $\mathcal{P}$. By Remark 3.7 we can easily see from the corresponding result in the Euclidean case that $U_t^\varepsilon \in \mathbf{D}_\infty(\mathcal{P})$ for every $(\varepsilon, t) \in [0,1]^2$ and $X_t^\varepsilon = \pi(U_t^\varepsilon) \in \mathbf{D}_\infty(\mathcal{M})$ and that, for every $f \in C^\infty(\mathcal{P})$, $p \in (1,\infty)$, $k \in \mathbb{N}$, the $\mathbf{D}_{p,k}$-norm of $f(U_t^\varepsilon)$ is bounded in $(\varepsilon, t) \in [0,1]^2$. In Lemma 3.1 we checked the partial Hörmander condition. Therefore, X_t^ε is non-degenerate for every $t, \varepsilon \in (0,1]$. More precisely, the following Kusuoka–Stroock estimate is known: there exist a constant $\nu > 0$ independent of p and a constant $C_p > 0$ such that, for every $1 < p < \infty$,

$$\|(\det \sigma_{X_1^\varepsilon})^{-1}\|_{L^p} \le C_p \varepsilon^{-\nu}, \quad \varepsilon \in (0,1]. \tag{4.6}$$

Combining this with Lemma 3.6, the transition probability of $\varepsilon^2(\Delta_{\mathrm{sub}}/2 + V)$-diffusion has a density $p_t^\varepsilon(x,a)$ with respect to $\mathbf{vol}(da)$, which is smooth in $a \in \mathcal{M}$. It holds that

$$p_t^\varepsilon(x,a) = p_{\varepsilon^2 t}^1(x,a) = \mathbb{E}[\delta_a(X_t^\varepsilon)]. \tag{4.7}$$

In the appendix we will show that $p_t^\varepsilon(x,a) > 0$ for all $\varepsilon, t \in (0,1]$ and $x, a \in \mathcal{M}$, which then enables us to define the pinned diffusion measure associated with $\varepsilon^2(\Delta_{\mathrm{sub}}/2 + V)$ from every x to every a. (We will later make sure that the measure actually exists.)

§4.2. Elements of rough path theory

In this subsection we recall the geometric rough path space with Hölder or Besov norm and quasi-sure properties of the rough path lift. For basic properties of geometric rough path space, we refer to [20, 9]. For the geometric rough path space with Besov norm, we refer to [9, Appx. A.2]. The quasi-sure properties of the rough path lift are summarized in [15]. In this paper we assume $\alpha \in (1/3, 1/2)$ for the Hölder parameter. We also assume that the Besov parameter $(\alpha, 4m)$ satisfies the conditions

$$\frac{1}{3} < \alpha < \frac{1}{2}, \quad m \in \mathbb{N}_+, \quad \alpha - \frac{1}{4m} > \frac{1}{3}, \quad 4m\Big(\frac{1}{2} - \alpha\Big) > 1. \tag{4.8}$$

We work in Lyons' original formulation of RDEs (see [20]), but we basically study the first level paths of solutions only. For brevity we will write $\lambda_t^\varepsilon := \varepsilon^2 t$ for $\varepsilon \in (0,1]$.

We denote by $G\Omega_\alpha^H(\mathbb{R}^d)$ the α-Hölder geometric rough path space over $\mathbb{R}^d$. A generic element of $G\Omega_\alpha^H(\mathbb{R}^d)$ is denoted by $\mathbf{w} = (\mathbf{w}^1, \mathbf{w}^2)$. For $\beta \in (0,1]$, let $C_0^{\beta-H}([0,1], \mathbb{R}^k)$ be the Banach space of all $\mathbb{R}^k$-valued β-Hölder continuous paths that start at 0. If $\alpha + \beta > 1$, then the Young pairing

$$G\Omega_\alpha^H(\mathbb{R}^d) \times C_0^{\beta-H}([0,1], \mathbb{R}^k) \ni (\mathbf{w}, \lambda) \mapsto (\mathbf{w}, \boldsymbol{\lambda}) \in G\Omega_\alpha^H(\mathbb{R}^{d+k})$$

is a well-defined, locally Lipschitz continuous map. (See [9, Sect. 9.4] for example.)

Now we consider a system of RDEs driven by the Young pairing $(\mathbf{w}, \boldsymbol{\lambda}) \in G\Omega^H_\alpha(\mathbb{R}^{d+1})$ of $\mathbf{w} \in G\Omega^H_\alpha(\mathbb{R}^d)$ and $\lambda \in C_0^{1-H}([0,1],\mathbb{R}^1)$. (The main example we have in mind is $\lambda_t = \mathrm{const} \times t$.) For vector fields $V_i \colon \mathbb{R}^n \to \mathbb{R}^n$ $(0 \le i \le d)$, consider

$$dx_t = \sum_{i=1}^d V_i(x_t)\, dw_t^i + V_0(x_t)\, d\lambda_t, \quad x_0 = x \in \mathbb{R}^n. \tag{4.9}$$

The RDEs for the Jacobian process and its inverse are given as

$$dJ_t = \sum_{i=1}^d \nabla V_i(x_t) J_t\, dw_t^i + \nabla V_0(x_t) J_t\, d\lambda_t, \qquad J_0 = \mathrm{Id}_n, \tag{4.10}$$

$$dK_t = -\sum_{i=1}^d K_t \nabla V_i(x_t)\, dw_t^i - K_t \nabla V_0(x_t)\, d\lambda_t, \quad K_0 = \mathrm{Id}_n. \tag{4.11}$$

Note that J, K, and ∇V_i are $\mathrm{Mat}(n,n)$-valued. Here, $\mathrm{Mat}(n,m)$ stands for the set of all real $n \times m$ matrices and Id_n stands for the identity matrix of size n.

For simplicity we assume that V_i, $0 \le i \le d$, is of C_b^∞, that is, when viewed as an $\mathbb{R}^n$-valued function, V_i is a bounded smooth function with bounded derivatives of all order. It is then known that a unique global solution of (4.9)–(4.11) exists for any $\mathbf{w}$ and λ . Moreover, Lyons' continuity theorem holds. In that case, the following map is continuous:

$$G\Omega^H_\alpha(\mathbb{R}^d) \times C_0^{1-H}([0,1],\mathbb{R}^1) \ni (\mathbf{w}, \lambda) \mapsto (\mathbf{x}, \mathbf{J}, \mathbf{K}) \in G\Omega^H_\alpha(\mathbb{R}^n \oplus \mathrm{Mat}(n,n)^{\oplus 2}).$$

The map $(\mathbf{w}, \lambda) \mapsto \mathbf{x}$ is denoted by

$$\Phi \colon G\Omega^H_\alpha(\mathbb{R}^d) \times C_0^{1-H}([0,1],\mathbb{R}^1) \to G\Omega^H_\alpha(\mathbb{R}^n).$$

(We adopt Lyons' formulation of RDEs as in [20]. So the initial values of the first level paths must be adjusted.) If $w \in C_0^{1-H}([0,1],\mathbb{R}^d)$ or $w \in \mathcal{H} = \mathcal{H}^d$ and $\mathbf{w}$ is its natural lift, then the path

$$t \mapsto (x + \mathbf{x}^1_{0,t}, \mathrm{Id} + \mathbf{J}^1_{0,t}, \mathrm{Id} + \mathbf{K}^1_{0,t}) \tag{4.12}$$

coincides with the solution of a system (4.9)–(4.11) of ODEs understood in the usual Riemann–Stieltjes sense. Recall that $\mathbf{x}^1_{0,t}$ is the first level path of $\mathbf{x}$ evaluated at $(0,t)$. Keep in mind that $(\mathrm{Id} + \mathbf{J}^1_{0,t})^{-1} = \mathrm{Id} + \mathbf{K}^1_{0,t}$ always holds.

When $\mathbf{w}$ is Brownian rough path $\mathbf{W}$, i.e. the natural lift of (w_t), and $\lambda = \lambda^1$, the process in (4.12) coincides μ-a.s. with the corresponding system of usual Stratonovich SDEs with drift. In this case, $x_t := x + \mathbf{x}^1_{0,t} \in \mathbf{D}_\infty(\mathbb{R}^n)$ for every t. If $G \colon \mathbb{R}^n \to \mathbb{R}^m$ is a smooth map with bounded derivatives of all order (≥ 1), then

$G(x_t) \in \mathbf{D}_\infty(\mathbb{R}^m)$ and for every $h \in \mathcal{H}$ and t, it holds that

$$D_h G(x_t) = (\nabla G)(x + \mathbf{x}^1_{0,t})(\mathrm{Id} + \mathbf{J}^1_{0,t}) \int_0^t (\mathrm{Id} + \mathbf{K}^1_{0,s}) \mathbf{V}(x + \mathbf{x}^1_{0,s})\, dh_s, \quad \text{a.s.} \tag{4.13}$$

(with $\mathbf{w} = \mathbf{W}$). Here, we view $\mathbf{V} := [V_1, \dots, V_d] \in \mathrm{Mat}(n, d)$ and $\nabla G \in \mathrm{Mat}(m, n)$.

We define a continuous function

$$\Gamma \colon G\Omega^H_\alpha(\mathbb{R}^d) \times C_0^{1-H}([0,1], \mathbb{R}^1) \times [0,1] \to \mathrm{Mat}(m, m) \tag{4.14}$$

as follows: set

$$\Gamma(\mathbf{w}, \lambda)_t = (\nabla G)(x + \mathbf{x}^1_{0,t})(\mathrm{Id} + \mathbf{J}^1_{0,t}) C(\mathbf{w}, \lambda)_t (\mathrm{Id} + \mathbf{J}^1_{0,t})^* (\nabla G)(x + \mathbf{x}^1_{0,t})^*,$$

where

$$C(\mathbf{w}, \lambda)_t := \int_0^t (\mathrm{Id} + \mathbf{K}^1_{0,s}) \mathbf{V}(x + \mathbf{x}^1_{0,s}) \mathbf{V}(x + \mathbf{x}^1_{0,s})^* (\mathrm{Id} + \mathbf{K}^1_{0,s})^*\, ds.$$

Here, the superscript $*$ stands for the transpose of a matrix. Then (4.13) implies that when $\mathbf{w} = \mathbf{W}$ and $\lambda = \lambda^1$, the Malliavin covariance matrix of $G(x + \mathbf{x}^1_{0,t})$ equals $\Gamma(\mathbf{W}, \lambda^1)_t$, μ-a.s. Similarly for $h \in \mathcal{H}$, the deterministic Malliavin covariance matrix of $G(x + \mathbf{x}^1_{0,t})$ equals $\Gamma(\mathbf{h}, \lambda^1)_t$, where $\mathbf{w} = \mathbf{h}$ is the natural rough path lift of h. (When G is the identity map, these formulas are well known. The general case is just a straightforward modification.)

For $(\alpha, 4m)$ that satisfies (4.8), $G\Omega^B_{\alpha,4m}(\mathbb{R}^d)$ denotes the geometric rough path space over $\mathbb{R}^d$ with $(\alpha, 4m)$-Besov norm. Recall that the distance on this space is given by

$$\begin{aligned} d(\mathbf{w}, \widehat{\mathbf{w}}) &= \|\mathbf{w}^1 - \widehat{\mathbf{w}}^1\|_{\alpha,4m-B} + \|\mathbf{w}^2 - \widehat{\mathbf{w}}^2\|_{2\alpha,2m-B} \\ &:= \sum_{i=1,2} \left(\iint_{0 \le s < t \le 1} \frac{|\mathbf{w}^i_{s,t} - \widehat{\mathbf{w}}^i_{s,t}|^{4m/i}}{|t-s|^{1+4m\alpha}}\, ds\, dt \right)^{i/4m}. \end{aligned}$$

By the Besov–Hölder embedding theorem for rough path spaces, there is a continuous embedding $G\Omega^B_{\alpha,4m}(\mathbb{R}^d) \hookrightarrow G\Omega^H_{\alpha-(1/4m)}(\mathbb{R}^d)$. If $\alpha < \alpha' < 1/2$, there is a continuous embedding $G\Omega^H_{\alpha'}(\mathbb{R}^d) \hookrightarrow G\Omega^B_{\alpha,4m}(\mathbb{R}^d)$. Basically, we will not write the first embedding explicitly. (For example, if we write $\Phi(\mathbf{w}, \lambda)$ for $(\mathbf{w}, \lambda) \in G\Omega^B_{\alpha,4m}(\mathbb{R}^d) \times C_0^{1-H}([0,1], \mathbb{R}^1)$, then it is actually the composition of the first embedding map above and Φ with respect to $\{\alpha - 1/(4m)\}$-Hölder topology.) It is known that the Young translation by $h \in \mathcal{H}$ works well on $G\Omega^B_{\alpha,4m}(\mathbb{R}^d)$ under (4.8). The map $(\mathbf{w}, h) \mapsto \tau_h(\mathbf{w})$ is continuous from $G\Omega^B_{\alpha,4m}(\mathbb{R}^d) \times \mathcal{H}$ to $G\Omega^B_{\alpha,4m}(\mathbb{R}^d)$, where $\tau_h(\mathbf{w})$ is the Young translation of $\mathbf{w}$ by h (see [15, Lem. 5.1]).

Now we review quasi-sure properties of the rough path lift map $\mathcal{L}$ from $\mathcal{W}$ to $G\Omega^B_{\alpha,4m}(\mathbb{R}^d)$. For $k = \mathbb{N}_+$ and $w \in \mathcal{W}$, we denote by $w(k)$ the kth dyadic piecewise linear approximation of w associated with the partition $\{j2^{-k} \mid 0 \le j \le 2^k\}$ of $[0,1]$. We denote the natural lift of $w(k)$ by $\mathcal{L}(w(k))$, which is defined by the Riemann–Stieltjes (or Young) integral. We set

$$\mathcal{Z}_{\alpha,4m} := \big\{ w \in \mathcal{W} \;\big|\; \{\mathcal{L}(w(k))\}_{k=1}^{\infty} \text{ is Cauchy in } G\Omega^B_{\alpha,4m}(\mathbb{R}^d) \big\}.$$

We define $\mathcal{L}\colon \mathcal{W} \to G\Omega^B_{\alpha,4m}(\mathbb{R}^d)$ by $\mathcal{L}(w) = \lim_{m\to\infty} \mathcal{L}(w(k))$ if $w \in \mathcal{Z}_{\alpha,4m}$ and we define $\mathcal{L}(w)$ to be the zero rough path if $w \notin \mathcal{Z}_{\alpha,4m}$. We will use this version of $\mathcal{L}$, which is Borel measurable, and write $\mathbf{W} := \mathcal{L}(w)$ (as before) when it is regarded as a rough-path-space-valued random variable defined on $\mathcal{W}$.

Note that $\mathcal{H}$ and $C_0^{\beta-H}([0,1],\mathbb{R}^d)$ with $\beta \in (1/2, 1]$ are subsets of $\mathcal{Z}_{\alpha,4m}$ and the two definitions of rough path lift coincide. (We will often write $\mathbf{h} = \mathcal{L}(h)$ for $h \in \mathcal{H}$.) Under the scalar multiplication (i.e. the dilation) and the Cameron–Martin translation, $\mathcal{Z}_{\alpha,4m}$ is left invariant. Moreover, $c\mathcal{L}(w) = \mathcal{L}(cw)$ and $\tau_h(\mathcal{L}(w)) = \mathcal{L}(w+h)$ for any $w \in \mathcal{Z}_{\alpha,4m}$, $c \in \mathbb{R}$, and $h \in \mathcal{H}$. It is known that $\mathcal{Z}^c_{\alpha,4m}$ is slim, that is, the (p,r)-capacity of this set is zero for any $p \in (1,\infty)$ and $r \in \mathbb{N}_+$. Therefore, from the viewpoint of quasi-sure analysis, the lift map $\mathcal{L}$ is well defined. Moreover, the map $\mathcal{W} \ni w \mapsto \mathcal{L}(w) \in G\Omega^B_{\alpha,4m}(\mathbb{R}^d)$ is ∞-quasi-continuous. (This kind of ∞-quasi-continuity was first shown in [1].)

Let $\varepsilon \in (0,1]$ be a small parameter. We now recall that the unique solution of an RDE driven by $\varepsilon\mathbf{W} = \mathcal{L}(\varepsilon w)$ coincides with that of the corresponding scaled Stratonovich SDE given as

$$dX^\varepsilon_t = \varepsilon \sum_{i=1}^d V_i(X^\varepsilon_t) \circ dw^i_t + \varepsilon^2 V_0(X^\varepsilon_t)\,dt, \quad X^\varepsilon_0 = x \in \mathbb{R}^n. \tag{4.15}$$

When necessary, we will write $X^\varepsilon_t = X^\varepsilon(t,x,w)$ or $X^\varepsilon(t,x)$. Then $X^\varepsilon(\,\cdot\,,x,w) = x + \Phi(\varepsilon\mathbf{W},\lambda^\varepsilon)^1$ holds μ-a.s., which means that the right-hand side is an ∞-quasi-continuous modification of the left-hand side as a $C^{\alpha-H}([0,1],\mathbb{R}^n)$-valued Wiener functional. Similarly, for every $t,\varepsilon \in (0,1]$, $\Gamma(\varepsilon\mathbf{W},\lambda^\varepsilon)_t = \varepsilon^{-2}\sigma_{X^\varepsilon_t}$ holds, μ-a.s. Therefore, not just X^ε_t itself, but also its Malliavin covariance matrix $\sigma_{X^\varepsilon_t}$ is a continuous function of a Brownian rough path.

The skeleton ODE (without drift) associated with small noise problems for the above SDE (4.15) is given as follows: for $h \in \mathcal{H}$,

$$d\phi_t = \sum_{i=1}^d V_i(\phi_t)\,dh^i_t, \quad \phi_0 = x \in \mathbb{R}^n. \tag{4.16}$$

We write the unique solution $\phi = \phi(h)$ when necessary (which equals $x + \Phi(\mathbf{h},0)^1$). The deterministic Malliavin covariance matrix of $\phi(h)_t$ at h is denoted by $\sigma_{\phi_t}(h)$.

As is well known, $\Gamma(\mathbf{h},0)_t = \sigma_{\phi_t}(h)$. Recall that $\det\sigma_{\phi_t}(h) > 0$ if and only if the tangent map of $\phi_t\colon \mathcal{H} \to \mathbb{R}^n$ at h is surjective.

Set $X^{\varepsilon,h} := X^\varepsilon(\,\cdot\,, x, w + (h/\varepsilon)) = x + \Phi(\tau_h(\varepsilon\mathbf{W}), \lambda^\varepsilon)^1$ for $h \in \mathcal{H}$. In other words, $X^{\varepsilon,h}$ uniquely solves the Stratonovich SDE

$$dX_t^{\varepsilon,h} = \sum_{i=1}^d V_i(X_t^\varepsilon) \circ d(\varepsilon w_t^i + h_t) + \varepsilon^2 V_0(X_t^\varepsilon)\,dt, \quad X_0^{\varepsilon,h} = x.$$

Small noise asymptotics of $X^{\varepsilon,h}$ has been extensively studied. One basic result is the following asymptotics:

$$X_1^{\varepsilon,h} = \phi(h)_1 + \varepsilon\Xi_1^h + O(\varepsilon^2) \quad \text{in } \mathbf{D}_\infty(\mathbb{R}^n) \text{ as } \varepsilon \searrow 0,$$

where Ξ_1^h is the element of the first-order Wiener chaos given by the Wiener integral

$$\Xi_t^h(w) = (\mathrm{Id} + \mathbf{J}(\mathbf{h},0)_{0,t}^1) \int_0^t (\mathrm{Id} + \mathbf{K}(\mathbf{h},0)_{0,s}^1)\mathbf{V}(\phi(h)_s)\,dw_s.$$

Hence, for $G\colon \mathbb{R}^n \to \mathbb{R}^m$ as above,

$$G(X_1^{\varepsilon,h}) = G(\phi(h)_1) + \varepsilon(\nabla G)(\phi(h)_1)\Xi_1^h + O(\varepsilon^2) \quad \text{in } \mathbf{D}_\infty(\mathbb{R}^m) \text{ as } \varepsilon \searrow 0.$$

Note that $(\nabla G)(\phi(h)_1)\Xi_1^h$ belongs to the first order Wiener chaos and therefore induces a mean-zero Gaussian measure on $\mathbb{R}^m$. Its covariance matrix equals $\Gamma(\mathbf{h},0)_1$, which in turn equals the deterministic Malliavin covariance matrix of $\mathcal{H} \ni k \mapsto G(\phi(k)_1)$ at h.

The skeleton ODE (with drift) associated with the above SDE (4.15) with $\varepsilon = 1$ is given as follows: for $h \in \mathcal{H}$,

$$d\zeta_t = \sum_{i=1}^d V_i(\zeta_t)\,dh_t^i + V_0(\zeta_t)\,dt, \quad \zeta_0 = x \in \mathbb{R}^n. \tag{4.17}$$

We write the unique solution $\zeta = \zeta(h)$ when necessary (which equals $x + \Phi(\mathbf{h},\lambda^1)$). The deterministic Malliavin covariance matrix of $\zeta(h)_t$ at h is denoted by $\sigma_{\zeta_t}(h)$. As is well known, $\Gamma(\mathbf{h},\lambda^1)_t = \sigma_{\zeta_t}(h)$.

Lemma 4.3. *Consider SDE (4.15) with $\varepsilon = 1$ and ODE (4.17). We assume that at $t \in (0,1]$, $\det\sigma_{X_t^1} > 0$ holds, μ-a.s. Then $\{\mathbf{h} = \mathcal{L}(h) \mid h \in \mathcal{H},\ \det\sigma_{\zeta_t}(h) > 0\}$ is dense in $G\Omega_\alpha^H(\mathbb{R}^d)$ for any $1/3 < \alpha < 1/2$.*

Proof. Take $m \in \mathbb{N}_+$ so large that $(\alpha + (1/4m), 4m)$ still satisfies (4.8) and set

$$A := \bigl\{w \in \mathcal{Z}_{\alpha+(1/4m),4m} \bigm| \det\sigma_{X_t^\varepsilon}(w) > 0\bigr\} \subset \mathcal{W}.$$

This subset is of full μ-measure and hence $\mathcal{L}(A)$ is of full measure with respect to the law of Brownian rough path. Note that $\mathcal{L}(A) \subset G\Omega_\alpha^H(\mathbb{R}^d)$ due to the Besov–Hölder embedding theorem. Thanks to the support theorem for Brownian rough path (see [9, Thm. 13.54]), $\mathcal{L}(A)$ must be dense in $G\Omega_\alpha^H(\mathbb{R}^d)$. For every $w \in A$, we have $\lim_{k\to\infty} \mathcal{L}(w(k)) = \mathcal{L}(w)$ in α-Hölder topology and therefore

$$\lim_{k\to\infty} \sigma_{\zeta_t}(w(k)) = \lim_{k\to\infty} \Gamma\big(\mathcal{L}(w(k)), \lambda\big)_t = \Gamma(\mathcal{L}(w), \lambda)_t = \sigma_{X_t^1}(w).$$

This implies that $\det \sigma_{\zeta_t}(w(k)) > 0$ for large enough k. This proves the lemma. $\square$

§5. Large deviations for rough path lift of positive Watanabe distributions

In this section we formulate an LDP for the rough path lifts of Watanabe's pullback of the delta functions, from which our main theorem (Theorem 2.1) easily follows.

For $x \in \mathcal{M}$, we take any $u \in \pi^{-1}(x)$ and consider SDE (3.10) driven by the canonical realization of d-dimensional Brownian motion $(w_t)_{t\in[0,1]}$. By (4.6), $\delta_a(X_t^\varepsilon) \in \widetilde{\mathbf{D}}_{-\infty}$ is a well-defined positive Watanabe distribution for every $a \in \mathcal{M}$. By the positivity of the heat kernel (which will be proved in the appendix), we see that $p_t^\varepsilon(x,a) = \mathbb{E}[\delta_a(X_t^\varepsilon)] > 0$ for all $t, \varepsilon \in (0,1]$ and $x, a \in \mathcal{M}$. By Sugita's theorem [26], the positive Watanabe distribution $\delta_a(X_1^\varepsilon)$ at time $t = 1$ is in fact a non-trivial finite Borel measure on $\mathcal{W}$, which will be denoted by $\theta_{u,a}^\varepsilon$.

Since $\mathcal{L}$ is defined outside a slim set, we can lift the measure $\theta_{u,a}^\varepsilon$ to a measure on $G\Omega_{\alpha,4m}^B(\mathbb{R}^d)$. We write $\nu_{u,a}^\varepsilon = (\varepsilon\mathcal{L})_*[\theta_{u,a}^\varepsilon]$. Here, $\varepsilon\mathcal{L}$ is the composition of $\mathcal{L}$ and the dilation by ε. Since the complement of $\mathcal{Z}_{\alpha,4m}$ is slim, $\nu_{u,a}^\varepsilon$ does not depend on how $\mathcal{L}$ is defined on this complement. We denote by $\hat{\theta}_{u,a}^\varepsilon$ and $\hat{\nu}_{u,a}^\varepsilon$ the normalized measure of $\theta_{u,a}^\varepsilon$ and $\nu_{u,a}^\varepsilon$, respectively. (Since the total mass of $\theta_{u,a}^\varepsilon$ or of $\nu_{u,a}^\varepsilon$ equals $\mathbb{E}[\delta_a(X_t^\varepsilon)] > 0$, this normalization is well defined.)

Let $\phi(h)$ be the solution of ODE (3.9) and write $\psi(h) = \pi(\phi(h))$. In what follows, we write $\mathcal{H} = \mathcal{H}^d$ for simplicity. We set

$$\mathcal{Q}^{u,a} = \{h \in \mathcal{H} \mid \psi(h)_1 = a\}. \tag{5.1}$$

By Chow–Rashevsky's theorem [24, Thm. 1.14], there exists an admissible path $(x_t)_{t\in[0,1]} \in \mathcal{H}_x(\mathcal{M}, \mathcal{D})$ such that $x_1 = a$. By Proposition 3.2, its anti-development belongs to $\mathcal{Q}^{u,a}$. This implies that $\mathcal{Q}^{u,a} \neq \emptyset$ for any u and a.

Define a rate function $I_1 : G\Omega_{\alpha,4m}^B(\mathbb{R}^n) \to [0,\infty]$ as

$$I_1(\mathbf{w}) = \begin{cases} \frac{1}{2}\|h\|_{\mathcal{H}}^2 & (\text{if } \mathbf{w} = \mathcal{L}(h) \text{ for some } h \in \mathcal{Q}^{u,a}), \\ \infty & (\text{otherwise}). \end{cases}$$

From the Schilder-type LDP for a Brownian rough path [9, Thm. 13.42], we can easily see that I_1 is good. Also define $\hat{I}_1(\mathbf{w}) = I_1(\mathbf{w}) - \min\{\|h\|^2_{\mathcal{H}}/2 \mid h \in \mathcal{Q}^{u,a}\}$, which is also good. From the goodness of I_1 and Proposition 3.2, we can easily see that the minimum above exists and equals $d_{\mathrm{SR}}(x,a)^2/2$.

Our main purpose in this section is to prove that $\{\nu^{\varepsilon}_{u,a}\}_{0<\varepsilon\le 1}$ satisfies an LDP of Schilder type on $G\Omega^B_{\alpha,4m}(\mathbb{R}^d)$ as $\varepsilon \searrow 0$. As we will see, our main result easily follows from the following theorem.

Theorem 5.1. *Let the notation be as above. Let $u \in \mathcal{P}$ and $a \in \mathcal{M}$ and assume (4.8) for the Besov parameter $(\alpha, 4m)$. Then the following hold:*

(i) *The family $\{\nu^{\varepsilon}_{u,a}\}_{0<\varepsilon\le 1}$ of finite measures is exponentially tight and satisfies an LDP on $G\Omega^B_{\alpha,4m}(\mathbb{R}^d)$ as $\varepsilon \searrow 0$ with speed ε^{-2} and good rate function I_1, that is, for every Borel set $A \subset G\Omega^B_{\alpha,4m}(\mathbb{R}^d)$, the following inequalities hold:*

$$
\begin{aligned}
-\inf_{\mathbf{w}\in A^{\circ}} I_1(\mathbf{w}) &\le \liminf_{\varepsilon\searrow 0} \varepsilon^2 \log \nu^{\varepsilon}_{u,a}(A^{\circ}) \\
&\le \limsup_{\varepsilon\searrow 0} \varepsilon^2 \log \nu^{\varepsilon}_{u,a}(\bar{A}) \le -\inf_{\mathbf{w}\in \bar{A}} I_1(\mathbf{w}).
\end{aligned}
$$

(ii) *The family $\{\hat{\nu}^{\varepsilon}_{u,a}\}_{0<\varepsilon\le 1}$ of probability measures is exponentially tight and satisfies an LDP on $G\Omega^B_{\alpha,4m}(\mathbb{R}^d)$ as $\varepsilon \searrow 0$ with speed ε^{-2} and good rate function $\hat{I}_1$.*

Since the whole set is both open and closed, Theorem 5.1(i) implies that

$$
\begin{aligned}
\lim_{t\searrow 0} t \log p^1_t(x,a) &= \lim_{\varepsilon\searrow 0} \varepsilon^2 \log \mathbb{E}[\delta_a(X^{\varepsilon}_1)] \\
&= \lim_{\varepsilon\searrow 0} \varepsilon^2 \log \theta^{\varepsilon}_{u,a}(\mathcal{W}) \\
&= \lim_{\varepsilon\searrow 0} \varepsilon^2 \log \nu^{\varepsilon}_{u,a}(G\Omega^B_{\alpha,4m}(\mathbb{R}^d)) \\
&= -\min\bigl\{\|h\|^2_{\mathcal{H}}/2 \bigm| h \in \mathcal{Q}^{x,a}\bigr\} = -d_{\mathrm{SR}}(x,a)^2/2.
\end{aligned}
$$

We have also used (4.7) above. Due to this Varadhan-type asymptotic formula, Theorem 5.1(ii) is immediate from (i). We will prove Theorem 5.1(i) in Sections 6 and 7.

Proof of Theorem 2.1. Now we consider $G\Omega^H_{\alpha}(\mathbb{R}^d)$ with $\alpha \in (1/3, 1/2)$. Due to the Besov–Hölder embedding, $\hat{\nu}^{\varepsilon}_{u,a}$ actually sits on this space and Theorem 5.1(ii) still holds even if $G\Omega^B_{\alpha,4m}(\mathbb{R}^d)$ is replaced by this space. Obviously, the family $\{\delta_{\lambda^{\varepsilon}}\}_{0<\varepsilon\le 1}$ is exponentially tight and satisfies an LDP on $C^{1-H}_0([0,1],\mathbb{R})$ with good rate function $+\infty \cdot \mathbf{1}_{\{0\}^c}$ with the convention that $\infty \cdot 0 = 0$. By a general result for LDPs for product measures (see [6, Exer. 4.2.7]), in which the exponential tightness plays a key role, $\{\hat{\nu}^{\varepsilon}_{u,a} \otimes \delta_{\lambda^{\varepsilon}}\}_{0<\varepsilon\le 1}$ satisfies an LDP on $G\Omega^H_{\alpha}(\mathbb{R}^d) \times$

$C_0^{1-H}([0,1],\mathbb{R})$ with good rate function which is defined for $(\mathbf{w},\lambda)$ as

$$\hat{I}_1(\mathbf{w}) + \infty \cdot \mathbf{1}_{\{0\}^c}(\lambda) = \begin{cases} \frac{1}{2}\|h\|_{\mathcal{H}}^2 & (\text{if } \mathbf{w} = \mathcal{L}(h) \text{ for some } h \in \mathcal{Q}^{u,a} \text{ and } \lambda = 0), \\ \infty & (\text{otherwise}). \end{cases}$$

Choose an embedding as in Remark 3.7 and consider RDE (4.9) with $V_i = A_i$, $x = u$, and $n = M$. Define a continuous map $\Psi\colon G\Omega_\alpha^H(\mathbb{R}^d) \times C_0^{1-H}([0,1],\mathbb{R}) \to C_x([0,1],\mathcal{M})$ by $\Psi(\mathbf{w},\lambda)_t = \pi(x + \Phi(\mathbf{w},\lambda)^1_{0,t})$. Then $\Psi(\varepsilon\mathbf{W},\lambda^\varepsilon)$ is an ∞-quasi-continuous modification of $X^\varepsilon = \pi(U^\varepsilon)$, where U^ε solves SDE (3.10). In what follows we use this version of X^ε. Note that $\psi(h) = \Psi(\mathcal{L}(h),0)$ for $h \in \mathcal{H}$.

Now we claim that the law of Ψ under $\hat{\nu}^\varepsilon_{u,a} \otimes \delta_{\lambda^\varepsilon}$ is the pinned $\varepsilon^2(\Delta_{\mathrm{sub}}/2+V)$-diffusion measure $\mathbb{Q}^\varepsilon_{x,a}$ from $x = \pi(u)$ to a. Let $k \geq 1$, $G \in C^\infty(\mathcal{M}^k)$, and $0 = t_0 < t_1 < \cdots < t_k < t_{k+1} = 1$ be arbitrary. Then we have

$$\begin{aligned} &\int G(\Psi(\mathbf{w},\lambda)_{t_1},\dots,\Psi(\mathbf{w},\lambda)_{t_k})\hat{\nu}^\varepsilon_{u,a} \otimes \delta_{\lambda^\varepsilon}(d\mathbf{w}\, d\lambda) \\ &\quad = \int G(\Psi(\varepsilon\mathbf{W},\lambda^\varepsilon)_{t_1},\dots,\Psi(\varepsilon\mathbf{W},\lambda^\varepsilon)_{t_k})\hat{\nu}^\varepsilon_{u,a}(dw) \\ &\quad = p_1^\varepsilon(x,a)^{-1}\,\mathbb{E}[G(X^\varepsilon_{t_1},\dots,X^\varepsilon_{t_k})\delta_a(X^\varepsilon_1)] \\ &\quad = p_1^\varepsilon(x,a)^{-1}\int_{\mathcal{M}^k} G(x_1,\dots,x_k)\prod_{i=0}^k p^\varepsilon_{t_{i+1}-t_i}(x_i,x_{i+1})\prod_{i=1}^k \mathrm{vol}(dx_i) \end{aligned}$$

as desired. Here, we set $x_0 = x$ and $x_{k+1} = a$ for simplicity. This proves our claim. Note that this argument also proves the existence of the pinned diffusion measure $\mathbb{Q}^\varepsilon_{x,a}$.

By the above fact and Lyons' continuity theorem, we can use the contraction principle ([6, Thm. 4.2.1]) to prove that $\{\mathbb{Q}^\varepsilon_{x,a}\}_{0<\varepsilon\leq 1}$ satisfies an LDP with good rate function. Since ψ is a bijection that preserves the energy (Proposition 3.2), the rate function J is given by (2.3). This completes the proof of Theorem 2.1. □

§6. Lower estimate

In this section we prove the lower estimate of LDP in Theorem 5.1(i). Take any $u \in \mathcal{P}$ and $a \in \mathcal{M}$. For this u, $\phi(h)$ denotes the unique solution of ODE (3.9) and $\psi(h)$ is its projection on $\mathcal{M}$. The subset $\mathcal{Q}^{u,a} \subset \mathcal{H}$ is defined in (5.1).

Lemma 6.1. *For every $h \in \mathcal{Q}^{u,a}$ there exists a sequence $\{h_j\}_{j=1}^\infty$ in $\mathcal{Q}^{u,a}$ which satisfies the following conditions:*

(i) $\lim_{j\to\infty}\|h_j - h\|_{\mathcal{H}} = 0$.

(ii) *For all j, $D\psi(h_j)_1\colon \mathcal{H} \to T_a\mathcal{M}$ is surjective, where $D\psi(h_j)_1$ denotes the tangent map of $\mathcal{H} \ni k \mapsto \psi(k)_1 \in \mathcal{M}$ at h_j.*

(iii) *For all j, $\langle h_j, \bullet \rangle_{\mathcal{H}} \in \mathcal{W}^*$, that is, this linear functional on $\mathcal{H}$ extends to a bounded linear functional on $\mathcal{W}$.*

Proof. First we will find $\{h_j\}_{j=1}^{\infty}$ which satisfies (i) and (ii) only. Let $\{Z_1, \ldots, Z_d\}$ be an orthonormal frame of $\mathcal{D}$ on a coordinate neighborhood U of $x = \pi(u)$. We view U as an open subset of $\mathbb{R}^n$ and extend Z_i, $1 \le i \le d$, as a smooth vector field on $\mathbb{R}^n$ with compact support. We denote by $\mathcal{H}_\tau$, $\tau \in (0,1)$, the Hilbert space of $\mathbb{R}^d$-valued Cameron–Martin paths defined on the time interval $[0, \tau]$.

Consider the following ODE on $\mathbb{R}^n$ driven by a Cameron–Martin path k:

$$dx(k)_t = \sum_{i=1}^{d} Z_i(x(k)_t)\, dk_t^i, \quad x(k)_0 = x.$$

The following fact was proved in [16, Sect. 3]: For every j large enough, there exists a Cameron–Martin path $k_j \in \mathcal{H}_{1/j}$ such that $|k'_{j,t}| \le 1$ for almost all $t \in [0, 1/j]$ and $Dx(k_j)_{1/j}$ is surjective. (Precisely, $\{Z_i\}$ is assumed to satisfy the bracket generating condition at every point of $\mathbb{R}^n$ in [16], while the condition is assumed only on U here. However, this difference does not matter at all since when j is large enough, the Cameron–Martin norm of $x(k_j)$ is small enough and therefore $x(k_j)$ stays inside U anyway.)

Then, since $\{Z_i\}$ are orthonormal, $|x(k_j)'_t| \le 1$ for almost all $t \in [0, 1/j]$. Denote by $l_j := \psi^{-1}(x(k_j)) \in \mathcal{H}_{1/j}$ the anti-development of $x(k_j)$; then $|l'_{j,t}| \le 1$ for almost all $t \in [0, 1/j]$ too. Define $h_j \in \mathcal{H}$ by

$$h_{j,t} := \begin{cases} l_{j,t} & \text{on } t \in [0, 1/j], \\ l_{j,(2/j)-t} & \text{on } t \in [1/j, 2/j], \\ h_\tau \text{ with } \tau = \dfrac{t-(2/j)}{1-(2/j)} & \text{on } t \in [2/j, 1]. \end{cases}$$

Since ODE (3.9) has no drift term, we can easily see that $\phi(h_j)_{2/j} = u$ and $\phi(h_j)_1 = \phi(h)_1$. In particular, $h_j \in \mathcal{Q}^{u,a}$. It is routine to check (i) (see [16] for a proof for instance).

Next we show (ii). Fix j and consider the admissible path $\psi(h_j)\colon [0,1] \to \mathcal{M}$. We can find a partition of $0 = s_0 < s_1 < \cdots < s_N = 1$ of $[0,1]$ such that, for all $1 \le r \le N$, $\psi(h_j) \restriction_{[s_{r-1}, s_r]}$ is contained in a coordinate neighborhood U_r on which an orthonormal frame $\{Z_1^{(r)}, \ldots, Z_d^{(r)}\}$ of $\mathcal{D}$ can be chosen. We may assume $s_1 = 1/j$, $U_1 = U$, and $Z_i^{(1)} = Z_i$ for all $1 \le i \le d$.

Obviously, there exists a unique Cameron–Martin path $k^{(r)}\colon [s_{r-1}, s_r] \to \mathbb{R}^d$ such that $\psi(h_j) \restriction_{[s_{r-1}, s_r]}$ satisfies the following ODE with the initial condition

$x^{(r)}_{s_{r-1}} = \psi(h_j)_{s_{r-1}}$:

$$dx^{(r)}_t = \sum_{i=1}^{d} Z^{(r)}_i(x^{(r)}_t)\, dk^{(r),i}_t \quad \text{on } [s_{r-1}, s_r]. \tag{6.1}$$

When the initial value of ODE (6.1) is replaced by $\xi^{(r-1)}$, we write $\mathrm{End}^{(r)}(\xi^{(r-1)}) := x^{(r)}_{s_r}$. When $\xi^{(r-1)}$ is close enough to $\psi(h_j)_{s_{r-1}}$, $\mathrm{End}^{(r)}(\xi^{(r-1)})$ is well defined. By the theory of flow of diffeomorphisms for ODEs, $\mathrm{End}^{(r)}$ is a local diffeomorphism from a neighborhood of $\psi(h_j)_{s_{r-1}}$ to a neighborhood of $\psi(h_j)_{s_r}$. (Note that here and in what follows, $k^{(r)}$ is fixed.) Hence, $\mathrm{End} := \mathrm{End}^{(N)} \circ \cdots \circ \mathrm{End}^{(2)}$ is a local diffeomorphism from a neighborhood of $\psi(h_j)_{s_1}$ to a neighborhood of $\psi(h_j)_1 = a$. In particular, the tangent map of End is bijective from $T_b\mathcal{M}$ to $T_a\mathcal{M}$, where we write $b := x(k_j)_{s_1} = \psi(h_j)_{s_1}$.

Now consider $\psi(h_j) \restriction_{[0,s_1]} = x(k_j)$. Since $Dx(k_j)_{s_1}$ is surjective, for every $v \in T_b\mathcal{M}$ there exist a sufficiently small $\varepsilon_0 > 0$ and a C^1-curve $(-\varepsilon_0, \varepsilon_0) \ni \varepsilon \mapsto k^\varepsilon_j \in \mathcal{H}_{1/j}$ such that $k^0_j = k_j$ and $(d/d\varepsilon)|_{\varepsilon=0} x(k^\varepsilon_j)_{s_1} = v$. (Below, ε_0 may change from line to line.)

Take $u \in T_a\mathcal{M}$ arbitrarily and let v be the unique element of $T_b\mathcal{M}$ which corresponds to u through the bijection of the derivative map of End at b. For this v, we take k^ε_j as above. Define h^ε_j to be the unique element in $\mathcal{H}$ such that (1) $\psi(h^\varepsilon_j)$ coincides with $x(k^\varepsilon_j)$ on $[0, s_1]$ and (2) $\psi(h^\varepsilon_j) \restriction_{[s_{r-1},s_r]}$ solves ODE (6.1) for all $2 \le r \le N$. By way of construction, $h^0_j = h_j$ and $(d/d\varepsilon)|_{\varepsilon=0} \psi(h^\varepsilon_j)_1 = u$. Thus, we have shown (ii).

Finally, we deal with (iii). Since $h \mapsto \psi(h)_1$ is Fréchet-C^1 and the inclusion $\mathcal{W}^* \hookrightarrow \mathcal{H}^* \cong \mathcal{H}$ is continuous and dense, we may use a topological lemma [15, Lem. 7.3]. It implies that we can find $\hat{h}_j \in \mathcal{Q}^{u,a}$ which satisfies the following conditions for sufficiently large j: (1) $\|h_j - \hat{h}_j\|_{\mathcal{H}} \le 1/j$, (2) $D\psi(\hat{h}_j)_1 \colon \mathcal{H} \to T_a\mathcal{M}$ is surjective, and (3) $\langle \hat{h}_j, \bullet \rangle_{\mathcal{H}} \in \mathcal{W}^*$. Hence, $\{\hat{h}_j\}_{j=1}^\infty$ is the desired sequence. (Precisely, [15, Lem. 7.3] is for $\mathbb{R}^n$-valued Fréchet-C^1 maps. However, it still holds for our manifold-valued case without modification since information outside a sufficiently small neighborhood of a is not used in its proof.) $\square$

For $R > 0$, we set

$$\widehat{B}_R = \bigl\{ \mathbf{w} \in G\Omega^B_{\alpha,4m}(\mathbb{R}^d) \bigm| \|\mathbf{w}^1\|^{4m}_{\alpha,4m-B} + \|\mathbf{w}^2\|^{2m}_{2\alpha,2m-B} < R^{4m} \bigr\} \tag{6.2}$$

and set $\widehat{B}_R(\mathbf{h}) = \tau_h(\widehat{B}_R)$, where τ_h is the Young translation by $h \in \mathcal{H}$ on $G\Omega^B_{\alpha,4m}(\mathbb{R}^d)$. Since τ_h is a homeomorphism from $G\Omega^B_{\alpha,4m}(\mathbb{R}^d)$ to itself, $\{\widehat{B}_R(\mathbf{h}) \mid R > 0\}$ forms a fundamental system of open neighborhoods around $\mathbf{h} = \mathcal{L}(h)$.

Proposition 6.2. *Assume that $h \in \mathcal{Q}^{u,a}$ satisfies that $D\psi(h)_1 \colon \mathcal{H} \to T_a\mathcal{M}$ is surjective and that $\langle h, \bullet\rangle_{\mathcal{H}} \in \mathcal{W}^*$. Then there exists a constant $c = c(h) > 0$ independent of R such that*

$$\liminf_{\varepsilon \searrow 0} \varepsilon^2 \log \nu^{\varepsilon}_{u,a}(\widehat{B}_R(\mathbf{h})) \geq -\frac{1}{2}\|h\|^2_{\mathcal{H}} - cR \tag{6.3}$$

holds for every sufficiently small $R > 0$.

Proof. In this proof, $R \in (0, R_0)$ and $\varepsilon \in (0, \varepsilon_0)$, where $R_0 > 0$ and $\varepsilon_0 > 0$ are sufficiently small constants and may vary from line to line.

Take a coordinate neighborhood $\widetilde{U}$ of a. We also view $\widetilde{U}$ as a bounded open subset of $\mathbb{R}^n$. By applying a dilation on $\mathbb{R}^n$ to $\widetilde{U}$ if necessary we may also assume that δ_a on $\mathcal{M}$ (with respect to **vol**) corresponds to the usual δ_a on $\mathbb{R}^n$ (with respect to the standard Lebesgue measure). We also take another open subset U so that $a \in U \subset \overline{U} \subset \widetilde{U}$.

For $F \in \mathbf{D}_{2,1}(\mathcal{M})$ that takes values in $\widetilde{U}$ (or $F \in \mathbf{D}_{2,1}(\mathcal{M})$ restricted to a subset of $\{F \in \widetilde{U}\}$), there are two Malliavin covariance matrices. One is the original one defined with respect to the Riemannian metric $\mathcal{M}$, while the other is the standard one for $\mathbb{R}^n$-valued Wiener functionals via the inclusion $\widetilde{U} \subset \mathbb{R}^n$. Since there is a constant $C = C(\widetilde{U}) > 0$ such that

$$C^{-1} \det \sigma_F(w) \leq \det \tilde{\sigma}_F(w) \leq C \det \sigma_F(w) \quad \text{on } \{w \in \mathcal{W} \mid F(w) \in \widetilde{U}\}, \tag{6.4}$$

either one of the two works. In what follows we will use the standard one for $\mathbb{R}^n$-valued Wiener functionals and denote it by σ_F again by slightly abusing the notation.

For $b \in \pi^{-1}(a)$, there exist a smooth map $\hat{\pi} \colon \mathcal{P} \to \mathbb{R}^n$ and a open neighborhood V of b such that $V \subset \pi^{-1}(\widetilde{U})$ and $\pi \restriction_V \equiv \hat{\pi} \restriction_V$. We can extend $\hat{\pi}$ again so that it is a smooth function $\hat{\pi} \colon \mathbb{R}^M \to \mathbb{R}^n$ with compact support. (Recall the embedding $\iota \colon \mathcal{P} \hookrightarrow \mathbb{R}^M$ in Remark 3.7.) We write $\widehat{X}^{\varepsilon}_t = \hat{\pi}(U^{\varepsilon}_t)$. When we need to specify the dependency on u and w, we write $\widehat{X}^{\varepsilon}_t = \widehat{X}^{\varepsilon}(t, u, w)$. (We will use the same notation for X^{ε}_t and U^{ε}_t too.)

Let $\chi \colon \mathbb{R} \to \mathbb{R}$ be as in Proposition 4.2. Moreover, we assume that χ is even and non-increasing on $[0, \infty)$ so that χ takes values in $[0, 1]$. Take any $f \in C^{\infty}(\mathbb{R}^n, [0, \infty))$ whose support is contained in the unit ball and set $f_j = j^n f(j\,\cdot)$. Then $\lim_{j\to\infty} f_j = \delta_0$ in $\mathcal{S}'(\mathbb{R}^n)$. If we set $f^{\varepsilon,a}_j = \varepsilon^{-n} f_j((\cdot - a)/\varepsilon)$, then $\lim_{j\to\infty} f^{\varepsilon,a}_j = \delta_a$ in $\mathcal{S}'(\mathbb{R}^n)$. There exists $j_0 > 0$ such that the support of $f^{\varepsilon,a}_j$ is contained in U for every $j \geq j_0$ and $\varepsilon \in (0, 1]$. In that case, $f^{\varepsilon,a}_j$ can also be viewed as a function on $\mathcal{M}$. We will assume $j \geq j_0$ and set $b := \phi(h)_1 \in \mathcal{P}$.

Then it holds that

$$
\begin{aligned}
\nu^{\varepsilon}_{u,a}(\widehat{B}_R(\mathbf{h})) &= \int I_{\widehat{B}_R(\mathbf{h})}(\mathbf{w})\nu^{\varepsilon}_{u,a}(d\mathbf{w}) = \int I_{\widehat{B}_R}(\tau_{-h}(\mathbf{w}))\nu^{\varepsilon}_{u,a}(d\mathbf{w}) \\
&= \int I_{\widehat{B}_R}(\tau_{-h}(\varepsilon\mathbf{W}))\theta^{\varepsilon}_{u,a}(dw) \\
&\geq \int \chi\Big(\frac{\|\tau_{-h}(\varepsilon\mathbf{W})^1\|^{4m}_{\alpha,4m-B} + \|\tau_{-h}(\varepsilon\mathbf{W})^2\|^{2m}_{2\alpha,2m-B}}{R^{4m}}\Big)\theta^{\varepsilon}_{u,a}(dw) \\
&= \mathbb{E}\Big[\chi\Big(\frac{\|\tau_{-h}(\varepsilon\mathbf{W})^1\|^{4m}_{\alpha,4m-B} + \|\tau_{-h}(\varepsilon\mathbf{W})^2\|^{2m}_{2\alpha,2m-B}}{R^{4m}}\Big)\delta_a(X^{\varepsilon}_1)\Big] \\
&= \lim_{j\to\infty}\mathbb{E}\Big[\chi\Big(\frac{\|\tau_{-h}(\varepsilon\mathbf{W})^1\|^{4m}_{\alpha,4m-B} + \|\tau_{-h}(\varepsilon\mathbf{W})^2\|^{2m}_{2\alpha,2m-B}}{R^{4m}}\Big)f^{\varepsilon,a}_j(X^{\varepsilon}_1)\Big] \\
&= \lim_{j\to\infty}\mathbb{E}\Big[\chi\Big(\frac{\|\tau_{-h}(\varepsilon\mathbf{W})^1\|^{4m}_{\alpha,4m-B} + \|\tau_{-h}(\varepsilon\mathbf{W})^2\|^{2m}_{2\alpha,2m-B}}{R^{4m}}\Big)f^{\varepsilon,a}_j(\widehat{X}^{\varepsilon}_1)\Big] \\
&= e^{-\|h\|^2_{\mathcal{H}}/2\varepsilon^2} \\
&\quad\times \lim_{j\to\infty}\mathbb{E}\Big[e^{-\langle h,w\rangle/\varepsilon}\chi\Big(\frac{\|(\varepsilon\mathbf{W})^1\|^{4m}_{\alpha,4m-B} + \|(\varepsilon\mathbf{W})^2\|^{2m}_{2\alpha,2m-B}}{R^{4m}}\Big) \\
&\qquad\qquad \times \varepsilon^{-n} f_j\Big(\varepsilon^{-1}\Big[\widehat{X}^{\varepsilon}\Big(1,u,w+\frac{h}{\varepsilon}\Big)-a\Big]\Big)\Big] \\
&\geq e^{-\|h\|^2_{\mathcal{H}}/2\varepsilon^2}e^{-cR/\varepsilon^2}\varepsilon^{-n} \\
&\quad\times \lim_{j\to\infty}\mathbb{E}\Big[\chi\Big(\frac{\|(\varepsilon\mathbf{W})^1\|^{4m}_{\alpha,4m-B} + \|(\varepsilon\mathbf{W})^2\|^{2m}_{2\alpha,2m-B}}{R^{4m}}\Big) \\
&\qquad\qquad \times f_j\Big(\varepsilon^{-1}\Big[\widehat{X}^{\varepsilon}\Big(1,u,w+\frac{h}{\varepsilon}\Big)-a\Big]\Big)\Big].
\end{aligned}
\tag{6.5}
$$

Note that when $R > 0$ and $\varepsilon > 0$ are sufficiently small, U^{ε}_1 is close enough to b and therefore we have $X^{\varepsilon}_1 = \widehat{X}^{\varepsilon}_1$. We used the Cameron–Martin formula too. We now check the last inequality. From the assumption that $\langle h,\bullet\rangle_{\mathcal{H}} \in \mathcal{W}^*$ and the fact that the $(\alpha,4m)$-Besov norm (of the first level path) is stronger than the usual sup-norm, we have $|\langle h,w\rangle_{\mathcal{H}}| \leq cR/\varepsilon$ for a certain constant $c > 0$ if $\|(\varepsilon\mathbf{W})^1\|_{\alpha,4m-B} \leq R$.

Set

$$
\begin{aligned}
\xi_{\varepsilon} &= R^{-4m}(\|(\varepsilon\mathbf{W})^1\|^{4m}_{\alpha,4m-B} + \|(\varepsilon\mathbf{W})^2\|^{2m}_{2\alpha,2m-B}), \\
F_{\varepsilon} &= \varepsilon^{-1}\Big[\widehat{X}^{\varepsilon}\Big(1,x,w+\frac{h}{\varepsilon}\Big)-a\Big].
\end{aligned}
$$

Let Γ be as in (4.14) (with $(x_t) = (U_t)$, $G = \hat{\pi}$, and $m = M$). Then the Malliavin covariance of F_{ε} equals $\Gamma(\tau_h(\varepsilon\mathbf{W}),\lambda^{\varepsilon})_1$, which tends to $\Gamma(\mathbf{h},0)_1$ as $(\varepsilon\mathbf{W},\lambda) \to$

$(\mathbf{0}, 0)$. Note that $\Gamma(\mathbf{h}, 0)_1$ is the deterministic Malliavin covariance of ψ_1 at h. As is well known, $D\psi(h)_1$ is surjective if and only if $\Gamma(\mathbf{h}, 0)_1$ is a strictly positive symmetric matrix. Hence, when $R_0 > 0$ and $\varepsilon_0 > 0$ are sufficiently small, there exists a constant $\rho > 0$ such that condition (4.3) holds for every $\varepsilon \in (0, \varepsilon_0]$ and $R \in (0, R_0]$ (ρ does not depend on ε or R). Now we can apply Proposition 4.1 to the right-hand side of (6.5) to obtain

$$\nu^{\varepsilon}_{x,a}(\widehat{B}_R(\mathbf{h})) \geq e^{-\|h\|^2_{\mathcal{H}}/2\varepsilon^2} e^{-cR/\varepsilon^2} \varepsilon^{-n}\, \mathbb{E}[\chi(\xi_\varepsilon)\delta_0(F_\varepsilon)].$$

It suffices to prove that $\lim_{\varepsilon \searrow 0} \mathbb{E}[\chi(\xi_\varepsilon)\delta_0(F_\varepsilon)]$ exists and is strictly positive. Since χ is constant near the origin, (4.5) clearly holds with $\xi_0 = 0$. We will check (4.4).

Since (U^{ε}_t) is viewed as the solution of the vector-space-valued SDE with small noise, its asymptotic behavior as $\varepsilon \searrow 0$ is well known:

$$U^{\varepsilon}\Big(1, u, w + \frac{h}{\varepsilon}\Big) = b + \varepsilon\eta_1 + O(\varepsilon^2) \quad \text{in } \mathbf{D}_\infty(\mathbb{R}^M),$$

where $\eta_1 = \eta_1(w)$ is a certain element of the first-order Wiener chaos (which can actually be written explicitly as a Wiener integral). It is easy to see from this that

$$F_\varepsilon = \nabla\hat{\pi}(b)\langle\eta_1\rangle + O(\varepsilon) \quad \text{in } \mathbf{D}_\infty(\mathbb{R}^n).$$

Then $\nabla\hat{\pi}(b)\langle\eta_1\rangle$ is also an element of the first order Wiener chaos and therefore induces a Gaussian measure of mean zero on $\mathbb{R}^n$. Since its covariance matrix equals $\Gamma(\mathbf{h}, 0)_1$, the Gaussian measure is non-degenerate and hence its probability density function is strictly positive at the origin. From this and Proposition 4.2 we see that

$$\lim_{\varepsilon \searrow 0} \mathbb{E}[\chi(\xi_\varepsilon)\delta_0(F_\varepsilon)] = \mathbb{E}[\delta_0(\nabla\hat{\pi}(b)\langle\eta_1\rangle)] \in (0, \infty).$$

This completes the proof of Proposition 6.2. □

Proof of the lower estimate in Theorem 5.1(i). Let $O \subset G\Omega^B_{\alpha,4m}(\mathbb{R}^d)$ be an open set with $\inf_{\mathbf{w}\in O} I_1(\mathbf{w}) < \infty$. Then we see from Lemma 6.1 that for every $\kappa > 0$ we can find $h \in \mathcal{Q}^{u,a}$ such that (1) $\mathbf{h} \in O$, (2) $0 \leq (\|h\|^2_{\mathcal{H}}/2) - \inf_{\mathbf{w}\in O} I_1(\mathbf{w}) < \kappa$ and (3) h satisfies the assumption of Proposition 6.2. Using Proposition 6.2, we have

$$\liminf_{\varepsilon \searrow 0} \varepsilon^2 \log \nu^{\varepsilon}_{u,a}(O) \geq \liminf_{\varepsilon \searrow 0} \varepsilon^2 \log \nu^{\varepsilon}_{u,a}(\widehat{B}_R(\mathbf{h})) \geq -\inf_{\mathbf{w}\in O} I_1(\mathbf{w}) - cR - \kappa$$

for every sufficiently small $R > 0$. Letting $R \searrow 0$ and then $\kappa \searrow 0$, we have the desired lower estimate. □

§7. Upper estimate

In this section we prove the upper estimate of LDP in Theorem 5.1(i). By a standard argument in large deviation theory, we can easily prove it by using

Propositions 7.3 and 7.5, which will be given below. (That $\nu^{\varepsilon}_{u,a}$ may not be a probability measure is irrelevant in this part.) In this section the positive constant ν varies from line to line.

The key to proving the upper estimate is a localized version of the IbP formula. This type of IbP formula is not new. For example, it was used in the proofs of Propositions 4.1 and 4.2. A quite similar argument appeared in the manifold-valued Malliavin calculus in [28]. Concerning this, see also [30].

We consider the projected process $X^{\varepsilon} = \pi(U^{\varepsilon})$ defined by SDE (3.10). Let $a \in U \subset \overline{U} \subset \widetilde{U}$ as in (6.4). By taking $\widetilde{U} \subset \mathbb{R}^n$ slightly smaller if necessary, we can extend the coordinate functions on $\widetilde{U} \ni x \mapsto x^i \in \mathbb{R}$, $1 \le i \le n$, to a smooth function on $\mathcal{M}$, which is denoted by β^i. If we set $\widehat{X}^{\varepsilon}_1 = \{\beta^i(X^{\varepsilon}_1)\}_{i=1}^n$, then $\widehat{X}^{\varepsilon}_1 \in \mathbf{D}_{\infty}(\mathbb{R}^n)$ and $\widehat{X}^{\varepsilon}_1 = X^{\varepsilon}_1$ on $\{w \mid X^{\varepsilon}_1(w) \in \widetilde{U}\}$. It should be recalled that the Sobolev norm $\|\widehat{X}^{\varepsilon}_1\|_{p,k}$ is bounded in $\varepsilon \in (0,1]$ for every $1 < p < \infty$ and $k \in \mathbb{N}$.

Lemma 7.1. *Let the notation be as above and write $F_{\varepsilon} := \widehat{X}^{\varepsilon}_1$ for notational simplicity. Suppose that $\eta_1, \eta_2 \colon \mathbb{R}^n \to [0,1]$ are smooth functions with compact support in $\widetilde{U}$ such that $\eta_2 \equiv 1$ on the support of η_1. Then, for every $f \in \mathcal{S}(\mathbb{R}^n)$, $G \in \mathbf{D}_{p,k}$ $(1 < p < \infty, k \in \mathbb{N}_+)$, and $1 \le i \le n$, the following assertions hold true:*

(i) $\gamma_{F_{\varepsilon}} \eta_1(F_{\varepsilon}) \in \mathbf{D}_{\infty}(\mathrm{Mat}(n,n))$. *Here, $\gamma_{F_{\varepsilon}}$ is the inverse of the Malliavin covariance matrix $\sigma_{F_{\varepsilon}}$.*

(ii) *If we set*

$$\Phi^{\eta_1}_i(\,\cdot\,;G) = \sum_{j=1}^{d} D^*(\gamma^{ij}_{F_{\varepsilon}} \eta_1(F_{\varepsilon}) \cdot G \cdot DF^j_{\varepsilon}),$$

then $\Phi^{\eta_1}_i(\,\cdot\,;G) \in \mathbf{D}_{p',k-1}$ for every $p' \in (1,p)$. Moreover, there exist a positive constant $c = c_{p,p'}$ and ν such that

$$\|\Phi^{\eta_1}_i(\,\cdot\,;G)\|_{p',k-1} = c\varepsilon^{-\nu}\|G\|_{p,k}, \quad 0 < \varepsilon \le 1. \tag{7.1}$$

Here, $c = c_{p,p'}$ and ν are independent of ε and G. Moreover, ν does not depend on (p,p') either.

(iii) *We have*

$$\mathbb{E}[\partial_i f(F_{\varepsilon})\eta_1(F_{\varepsilon})G] = \mathbb{E}[f(F_{\varepsilon})\eta_2(F_{\varepsilon})\Phi^{\eta_1}_i(\,\cdot\,;G)]. \tag{7.2}$$

Proof. By (4.6) and (6.4), there exist positive constants C_p, ν such that

$$\mathbb{E}[(\det \sigma_{F_{\varepsilon}})^{-p}\mathbf{1}_{\{F_{\varepsilon}\in\widetilde{U}\}}]^{1/p} \le C_p\varepsilon^{-\nu}, \quad 0 < \varepsilon \le 1. \tag{7.3}$$

Here, C_p and ν are independent of ε. Moreover, ν does not depend on p either.

Now we prove (i). It may be heuristically obvious, but we must take care of the possibility that $\gamma_{F_{\varepsilon}}$ is not defined outside $\{F_{\varepsilon} \in \widetilde{U}\}$. For $m \in \mathbb{N}_+$ we set $\sigma^m_{F_\varepsilon} := \sigma_{F_\varepsilon} + m^{-1}\mathrm{Id}_n$, where Id_n stands for the identity matrix of size n. Then $(\det \sigma^m_{F_\varepsilon})^{-1} \le m^n$, a.s. and its inverse $\gamma^m_{F_\varepsilon} := (\sigma^m_{F_\varepsilon})^{-1}$ exists. Moreover, $(\det \sigma^m_{F_\varepsilon})^{-1} \nearrow (\det \sigma_{F_\varepsilon})^{-1} \in [0, \infty]$, a.s. Note that $\gamma^m_{F_\varepsilon}$ and $(\det \sigma^m_{F_\varepsilon})^{-1}$ are both $\mathbf{D}_\infty$-functionals defined on $\mathcal{W}$. Recall that

$$\gamma^m_{F_\varepsilon} = (\det \sigma^m_{F_\varepsilon})^{-1} \times [\text{the adjugate matrix of } \sigma^m_{F_\varepsilon}].$$

From this and (7.3), we can easily see that $(\gamma^m_{F_\varepsilon})^{ij}\eta_1(F_\varepsilon) \to \gamma^{ij}_{F_\varepsilon}\eta_1(F_\varepsilon)$ in L^p as $m \to \infty$ $(1 < p < \infty)$.

Now we calculate the first-order derivative. Note that

$$\begin{aligned} D\{(\gamma^m_{F_\varepsilon})^{ij}\eta_1(F_\varepsilon)\} = &-\sum_{k,l}(\gamma^m_{F_\varepsilon})^{ik} \cdot D\sigma^{kl}_{F_\varepsilon} \cdot (\gamma^m_{F_\varepsilon})^{lj} \cdot \eta_1(F_\varepsilon) \\ &+ \sum_{l}(\gamma^m_{F_\varepsilon})^{ij} \cdot \partial_l\eta_1(F_\varepsilon) \cdot DF^l_\varepsilon. \end{aligned}$$

For the same reason as above, this belongs to $\mathbf{D}_\infty(\mathcal{H})$ and converges in $L^p(\mathcal{H})$, $1 < p < \infty$, as $m \to \infty$. The closability of D implies that

$$D\{\gamma^{ij}_{F_\varepsilon}\eta_1(F_\varepsilon)\} = -\sum_{k,l}\gamma^{ik}_{F_\varepsilon} \cdot D\sigma^{kl}_{F_\varepsilon} \cdot \gamma^{lj}_{F_\varepsilon} \cdot \eta_1(F_\varepsilon) + \sum_{l}\gamma^{ij}_{F_\varepsilon} \cdot \partial_l\eta_1(F_\varepsilon) \cdot DF^l_\varepsilon \tag{7.4}$$

and $\|\gamma^{ij}_{F_\varepsilon}\eta_1(F_\varepsilon)\|_{p,1} = O(\varepsilon^{-\nu})$ for every $1 < p < \infty$. Repeating essentially the same argument for higher-order derivatives, we can prove that $\|\gamma^{ij}_{F_\varepsilon}\eta_1(F_\varepsilon)\|_{p,k} \le O(\varepsilon^{-\nu})$ for every $1 < p < \infty$ and $k \in \mathbb{N}_+$ (if we adjust the value of $\nu > 0$). Thus, we have shown (i).

To prove (ii), just recall that D^* is a bounded linear map from $\mathbf{D}_{p,k}(\mathcal{H})$ to $\mathbf{D}_{p,k-1}$ for every p and k (see [14, p. 365] for instance).

Finally, we prove (iii). From a well-known formula for D^*, we see that

$$\Phi^{\eta_1}_i(\,\cdot\,; G) = -\sum_{j=1}^{d}\{\langle D(\gamma^{ij}_{F_\varepsilon} \cdot \eta_1(F_\varepsilon) \cdot G), DF^j_\varepsilon\rangle_{\mathcal{H}} + \gamma^{ij}_{F_\varepsilon} \cdot \eta_1(F_\varepsilon) \cdot G \cdot LF^j_\varepsilon\},$$

where $L = -D^*D$ is the Ornstein–Uhlenbeck operator. Since this vanishes outside $\{F_\varepsilon \in \mathrm{supp}(\eta_1)\}$, we have $\Phi^{\eta_1}_i(\,\cdot\,; G) = \eta_2(F_\varepsilon)\Phi^{\eta_1}_i(\,\cdot\,; G)$. As in the proof for the standard IbP formula in (4.2), we see from the definition of D^* that

$$\mathbb{E}[\partial_i f(F_\varepsilon)\eta_1(F_\varepsilon)G] = \mathbb{E}\left[\left\langle Df(F_\varepsilon), \sum_{j=1}^{d}\gamma^{ij}_{F_\varepsilon}DF^j_\varepsilon\right\rangle_{\mathcal{H}}\eta_1(F_\varepsilon)G\right] = \mathbb{E}[f(F_\varepsilon)\Phi^{\eta_1}_i(\,\cdot\,; G)].$$

This completes the proof of Lemma 7.1. □

Take a smooth function $\eta_j \colon \mathbb{R}^n \to [0,1]$, $1 \le j \le 2n+1$, with compact support in $\widetilde{U}$, with the following properties: (1) $\eta_1 \equiv 1$ on U, (2) $\eta_{j+1} \equiv 1$ on the support of η_j for all $1 \le j \le 2n$. We write $\eta := \{\eta_i\}_{i=1}^{2n+1}$.

Let $\beta := (i_1, i_2, \ldots, i_s)$ be a multi-index of length at most $2n$, that is, $s \le 2n$ and $1 \le i_1, \ldots, i_s \le n$. For $\beta = (i_1)$, just set $\Phi^\eta_\beta(\,\cdot\,; G) = \Phi^{\eta_1}_{i_1}(\,\cdot\,; G)$. For $\beta = (i_1, i_2)$, set $\Phi^\eta_\beta(\,\cdot\,; G) = \Phi^{\eta_2}_{i_2}(\,\cdot\,, \Phi^{\eta_1}_{i_1}(\,\cdot\,; G))$. When $s \ge 3$, write $\beta' = (i_1, \ldots, i_{s-1})$ and recursively set $\Phi^\eta_\beta(\,\cdot\,; G) = \Phi^{\eta_s}_{i_s}(\,\cdot\,, \Phi^\eta_{\beta'}(\,\cdot\,; G))$.

Proposition 7.2. *There exists a constant $\nu > 0$ such that*

$$\|\delta_a(X^\varepsilon_1)\|_{2,-2n} = O(\varepsilon^{-\nu}) \quad \text{as } \varepsilon \searrow 0.$$

Proof. If $\{g_k\}_{k\in\mathbb{N}}$ is a sequence of smooth functions supported in U such that $g_k \to \delta_a$ in $\mathcal{S}'(\mathbb{R}^n)$ as $k \to \infty$, then we see from the results in [28] that

$$\begin{aligned} \mathbb{E}[\delta_a(X^\varepsilon_1)G] &= \lim_{k\to\infty} \mathbb{E}[g_k(X^\varepsilon_1)G] \\ &= \lim_{k\to\infty} \mathbb{E}[g_k(X^\varepsilon_1)\eta_1(X^\varepsilon_1)G] = \lim_{k\to\infty} \mathbb{E}[g_k(F_\varepsilon)\eta_1(F_\varepsilon)G] \end{aligned} \tag{7.5}$$

for every $G \in \mathbf{D}_\infty$, where we write $F_\varepsilon = \widehat{X}^\varepsilon_1$ again.

Next recall that if we set $f(x) := \prod_{j=1}^n (x_j - a_j)^+$ and $\beta = (1, 1, 2, 2, \ldots, n, n)$, then $\partial^\beta f = \delta_a$ in the distribution sense. Take $\kappa > 0$ so small that $\prod_{j=1}^n [a_j - \kappa, a_j + \kappa] \subset U$ holds. For this κ, we can find a sequence $\{\lambda_k\}_{k=1}^\infty$ of smooth and non-decreasing functions on $\mathbb{R}$ such that (1) $\lambda_k(z)$ coincides with $z^+ := z \vee 0$ outside $[-\kappa, \kappa]$ for all k, (2) $\lambda_k(z)$ converges to z^+ uniformly on $[-\kappa, \kappa]$ as $k \to \infty$. Set $f_k(x) = \prod_{j=1}^n \lambda_k(x_j - a_j)$. Then f_k is smooth on $\mathbb{R}^n$ and $\lim_{k\to\infty} f_k = f$ uniformly on $\mathbb{R}^n$. Note that $\lim_{k\to\infty} \partial^\beta f_k = \delta_a$ in the distribution sense and $\partial^\beta f_k \equiv 0$ outside U. So we may take $g_k = \partial^\beta f_k$ in (7.5) above.

By (7.2) in Lemma 7.1 and the way $\Phi^\eta_\beta(\,\cdot\,; G)$ is defined, we have

$$\mathbb{E}[\partial^\beta f_k(F_\varepsilon)\eta_1(F_\varepsilon)G] = \mathbb{E}[f_k(F_\varepsilon)\eta_{2n+1}(F_\varepsilon)\Phi^\eta_\beta(\,\cdot\,; G)].$$

From this and (7.5), we have

$$\mathbb{E}[\delta_a(X^\varepsilon_1)G] = \mathbb{E}[f(F_\varepsilon)\eta_{2n+1}(F_\varepsilon)\Phi^\eta_\beta(\,\cdot\,; G)]. \tag{7.6}$$

Note that $|f(F_\varepsilon)\eta_{2n+1}(F_\varepsilon)|$ is dominated by a polynomial in $|F_\varepsilon|$. Therefore, $\|f(F_\varepsilon)\eta_{2n+1}(F_\varepsilon)\|_{L^p}$ is bounded in ε for every $1 < p < \infty$. By using (7.2) and (7.1) in Lemma 7.1 repeatedly, we can easily show that

$$\|\Phi^\eta_\beta(\,\cdot\,; G)\|_{L^{3/2}} = c\varepsilon^{-\nu}\|G\|_{2,2n}, \quad 0 < \varepsilon \le 1,$$

if the value of $\nu > 0$ is adjusted. Therefore, the right-hand side of (7.6) is also dominated by $c\varepsilon^{-\nu}\|G\|_{2,2n}$, where c is independent of ε and k. This proves the proposition. □

Proposition 7.3. *The family* $\{\nu^{\varepsilon}_{u,a}\}_{0<\varepsilon\le 1}$ *is exponentially tight on* $G\Omega^{B}_{\alpha,4m}(\mathbb{R}^d)$*, that is, for every* $M \in (0,\infty)$*, there exists a compact set* $K = K_M$ *such that*

$$\limsup_{\varepsilon\searrow 0} \varepsilon^2 \log \nu^{\varepsilon}_{u,a}(K^c) \le -M.$$

Proof. Let $\widehat{B}_R = \widehat{B}^{\alpha,4m}_R$ be a "ball of radius $R>0$" as in (6.2). It is shown in [15, Lem. 7.6] that there exists a constant $c = c_{\alpha,4m} > 0$ independent of R such that

$$\mathrm{Cap}_{2,2n}\big(\{w \in W \mid \mathcal{L}(w) \in (\widehat{B}^{\alpha,4m}_R)^c\}\big) \le e^{-cR^2} \quad \text{for sufficiently large } R>0.$$

Take $\kappa>0$ so small that the Besov parameter $(\alpha+\kappa, 4m)$ still satisfies (4.8). It is well known that $G\Omega^{B}_{\alpha+\kappa,4m}(\mathbb{R}^d)$ is embedded in $G\Omega^{B}_{\alpha,4m}(\mathbb{R}^d)$ and every bounded subset of the former is precompact in the latter.

By the way $\nu^{\varepsilon}_{u,a}$ is defined, it holds that

$$\begin{aligned}\nu^{\varepsilon}_{u,a}((\widehat{B}^{\alpha+\kappa,4m}_R)^c) &= \theta^{\varepsilon}_{u,a}\big(\{w \in \mathcal{W} \mid \varepsilon\mathcal{L}(w) \in (\widehat{B}^{\alpha+\kappa,4m}_R)^c\}\big)\\ &\le \|\delta_a(X^{\varepsilon}_1)\|_{2,-2n}\,\mathrm{Cap}_{2,2n}\big(\{w\in W \mid \mathcal{L}(w) \in (\widehat{B}^{\alpha+\kappa,4m}_{R/\varepsilon})^c\}\big)\\ &\le C\varepsilon^{-\nu}e^{-cR^2/\varepsilon^2} \quad \text{for sufficiently large } R>0.\end{aligned}$$

Here we used the inequality in item (e) in Section 4.1 and Proposition 7.2. Hence, we have

$$\limsup_{\varepsilon\searrow 0} \varepsilon^2 \log \nu^{\varepsilon}_{u,a}((\widehat{B}^{\alpha+\kappa,4m}_R)^c) \le -cR^2.$$

For given $M>0$, we choose R so large that $cR^2 \ge M$ holds. Since $\widehat{B}^{\alpha+\kappa,4m}_R$ is precompact in $(\alpha,4m)$-Besov topology, the proof is completed. ◻

For $R>0$ and $\mathbf{w} \in G\Omega^{B}_{\alpha,4m}(\mathbb{R}^d)$, we set

$$\begin{aligned}\overline{B}_R(\mathbf{w}) = \big\{\mathbf{v} \in G\Omega^{B}_{\alpha,4m}(\mathbb{R}^d) \mid \|\mathbf{v}^1 - \mathbf{w}^1\|^{4m}_{\alpha,4m-B}&\\ + \|\mathbf{v}^2 - \mathbf{w}^2\|^{2m}_{2\alpha,2m-B} \le R^{4m}\big\}.&\end{aligned} \tag{7.7}$$

Clearly, $\{\overline{B}_R(\mathbf{w}) | R>0\}$ forms a fundamental system of neighborhoods around $\mathbf{w}$. Set

$$\Xi_\varepsilon = \|(\varepsilon\mathbf{W})^1 - \mathbf{w}^1\|^{4m}_{\alpha,4m-B} + \|(\varepsilon\mathbf{W})^2 - \mathbf{w}^2\|^{2m}_{2\alpha,2m-B}.$$

Then, for every $\varepsilon \in (0,1]$ and $\mathbf{w}$, $\Xi_\varepsilon = \Xi_\varepsilon(\,\cdot\,,\mathbf{w})$ is ∞-quasi continuous and belongs to $\mathbf{D}_\infty$. Moreover, $\{\Xi_\varepsilon\}_\varepsilon$ is bounded in ε in $\mathbf{D}_{p,k}$ for every $1<p<\infty$ and $k\in\mathbb{N}$. For a smooth, non-increasing function $\chi\colon [0,\infty) \to [0,1]$ such that $\chi \equiv 1$ on $[0,1]$ and $\chi \equiv 0$ on $[2,\infty)$, $\chi(\Xi_\varepsilon/R^{4m}) \in \mathbf{D}_\infty$ satisfies $\mathbf{1}_{\overline{B}_R(\mathbf{w})} \circ (\varepsilon\mathcal{L}) \le \chi(\Xi_\varepsilon/R^{4m}) \le \mathbf{1}_{\overline{B}_{2R}(\mathbf{w})} \circ (\varepsilon\mathcal{L})$ quasi-surely.

Let us recall that the law of $\varepsilon\mathbf{W} = \varepsilon\mathcal{L}(w)$ satisfies the standard version of Schilder-type LDP on $G\Omega^B_{\alpha,4m}(\mathbb{R}^d)$ with good rate function I, where

$$I(\mathbf{w}) = \begin{cases} \frac{1}{2}\|h\|^2_{\mathcal{H}} & (\text{if } \mathbf{w} = \mathcal{L}(h) \text{ for some } h \in \mathcal{H}), \\ \infty & (\text{otherwise}). \end{cases}$$

By a general result on LDPs (see [6, Lem. 4.1.6] for example), we then have

$$\begin{aligned} &\lim_{R\searrow 0} \limsup_{\varepsilon\searrow 0} \varepsilon^2 \log \mu\bigl(\{w \in \mathcal{W} \mid \varepsilon\mathbf{W} \in \overline{B}_R(\mathbf{w})\}\bigr) \\ &\quad \le -\lim_{R\searrow 0} \inf\{I(\mathbf{v}) \mid \mathbf{v} \in \overline{B}_R(\mathbf{w})\} = -I(\mathbf{w}), \quad \mathbf{w} \in G\Omega^B_{\alpha,4m}(\mathbb{R}^d). \end{aligned} \tag{7.8}$$

Lemma 7.4. *Let the notation be as above. For every $p \in (1,\infty)$ and $k \in \mathbb{N}$, it holds that*

$$\lim_{R\searrow 0} \limsup_{\varepsilon\searrow 0} \varepsilon^2 \log \|\chi(\Xi_\varepsilon(\,\cdot\,,\mathbf{w})/R^{4m})\|^p_{p,k} \le -I(\mathbf{w}), \quad \mathbf{w} \in G\Omega^B_{\alpha,4m}(\mathbb{R}^d).$$

Proof. We will write $\Xi_\varepsilon = \Xi_\varepsilon(\,\cdot\,,\mathbf{w})$. By Meyer's equivalence, it is enough to estimate $\sum_{j=0}^k \|D^j\chi(\Xi_\varepsilon/R^{4m})\|^p_{L^p}$, where D is the $\mathcal{H}$-derivative. Hence, it amounts to computing

$$\max_{0\le j\le k} \limsup_{\varepsilon\searrow 0} \varepsilon^2 \log \|D^j\chi(\Xi_\varepsilon/R^{4m})\|^p_{L^p}.$$

For $j = 0$, we can easily see that $\|\chi(\Xi_\varepsilon/R^{4m})\|^p_{L^p} \le \mu(\{\varepsilon\mathbf{W} \in \overline{B}_{2R}(\mathbf{w})\})$. For $j = 1$, we have $D\chi(\Xi_\varepsilon/R^{4m}) = \chi'(\Xi_\varepsilon/R^{4m})(D\Xi_\varepsilon)R^{-4m}$. Hence, for every $q \in (1,\infty)$ and $R > 0$, there exists a positive constant $C = C_{q,R}$ (independent of ε) such that

$$\begin{aligned} \|D\chi(\Xi_\varepsilon/R^{4m})\|^p_{L^p} &\le \mathbb{E}[|\chi'(\Xi_\varepsilon/R^{4m})|^{pq}]^{1/q}\, \mathbb{E}[\|D\Xi_\varepsilon\|^r_{\mathcal{H}} R^{-4mr}]^{1/r} \\ &\le C\mu(\{\varepsilon\mathbf{W} \in \overline{B}_{2R}(\mathbf{w})\})^{1/q}, \end{aligned}$$

where $1/q + 1/r = 1$. By repeating similar computations we have, for all $0 \le j \le k$, that

$$\|D^j\chi(\Xi_\varepsilon/R^{4m})\|^p_{L^p} \le C\mu(\{\varepsilon\mathbf{W} \in \overline{B}_{2R}(\mathbf{w})\})^{1/q}.$$

Using (7.8), we have $\lim_{R\searrow 0} \limsup_{\varepsilon\searrow 0} \varepsilon^2 \log \|\chi(\Xi_\varepsilon/R^{4m})\|^p_{p,k} \le -(1/q)I(\mathbf{w})$. Letting $q \searrow 1$, we finish the proof. □

Proposition 7.5. *Let the notation be as above. Then we have*

$$\lim_{R\searrow 0} \limsup_{\varepsilon\searrow 0} \varepsilon^2 \log \nu^\varepsilon_{u,a}(\overline{B}_R(\mathbf{w})) \le -I_1(\mathbf{w}), \quad \mathbf{w} \in G\Omega^B_{\alpha,4m}(\mathbb{R}^d). \tag{7.9}$$

Proof. Write $\hat{a} := \Psi(\mathbf{w},0)_1$ for simplicity. First we consider the case $a \neq \hat{a}$. Let U and $\widehat{U}$ (with $U \cap \widehat{U} = \emptyset$) be a neighborhood of a and $\hat{a}$, respectively. By Lyons' continuity theorem, there exists $R_0 > 0$ and $\varepsilon_0 > 0$ such that $\Psi(\mathbf{v}, \lambda^\varepsilon)_1 \in \widehat{U}$ if $\mathbf{v} \in \overline{B}_{2R_0}(\mathbf{w})$ and $\varepsilon \in (0, \varepsilon_0)$. Suppose that $\{g_k\}_{k=1}^\infty$ is a sequence of smooth functions with supports in U that approximates δ_a. Then we have

$$\begin{aligned}\nu^\varepsilon_{u,a}(\overline{B}_R(\mathbf{w})) &\le \mathbb{E}[\chi(\Xi_\varepsilon/R^{4m})\delta_a(X^\varepsilon_1)] = \lim_{k\to\infty} \mathbb{E}[\chi(\Xi_\varepsilon/R^{4m}) g_k(X^\varepsilon_1)] \\ &\le \lim_{k\to\infty} \mathbb{E}[\mathbf{1}_{\{\varepsilon\mathbf{W}\in\overline{B}_{2R}(\mathbf{w})\}}\, g_k(\Psi(\varepsilon\mathbf{W}, \lambda^\varepsilon)_1)] = 0\end{aligned}$$

if $\varepsilon \in (0, \varepsilon_0)$ and $R \in (0, R_0)$. Thus, we have shown (7.9) for this case.

Next we consider the case $a = \hat{a}$. As in the proof of Proposition 7.2, we use the localized IbP formula. Let $\eta := \{\eta_i\}_{i=1}^{2n+1}$ as in Proposition 7.2. Then we can see that, for every $p \in (1, \infty)$, there exist positive constants c and ν (independent of ε) such that

$$\begin{aligned}\nu^\varepsilon_{u,a}(\overline{B}_R(\mathbf{w})) &\le \mathbb{E}[\chi(\Xi_\varepsilon/R^{4m})\delta_a(X^\varepsilon_1)] \\ &= \mathbb{E}[f(F_\varepsilon)\eta_{2n+1}(F_\varepsilon)\Phi^\eta_\beta(\,\cdot\,; \chi(\Xi_\varepsilon/R^{4m}))] \\ &\le c\varepsilon^{-\nu}\|\chi(\Xi_\varepsilon/R^{4m})\|_{p,2n}.\end{aligned}$$

Here we used (7.1) and (7.2) in Lemma 7.1 repeatedly. Using Lemma 7.4 and then letting $p \searrow 1$, we have $\lim_{R\searrow 0}\limsup_{\varepsilon\searrow 0} \varepsilon^2 \log \nu^\varepsilon_{u,a}(\overline{B}_R(\mathbf{w})) \le -I(\mathbf{w})$ in this case. Thus, we have proved (7.9). □

Appendix. Positivity of the heat kernel

Let $p^\varepsilon_t(x, a) = \mathbb{E}[\delta_a(X^\varepsilon_t)]$ be the heat kernel (or the density function) on $\mathcal{M}$ associated with the $\varepsilon^2(\Delta_{\mathrm{sub}}/2 + V)$-diffusion process. Here, (X^ε_t) is the projection of the solution of SDE (3.10) with the starting point u. The skeleton ODE with drift which corresponds to SDE (3.10) with $\varepsilon = 1$ is given as

$$d\bar{\phi}(h)_t = \sum_{i=1}^d A_i(\bar{\phi}(h)_t)\, dh^i_t + A_0(\bar{\phi}(h)_t)\, dt, \quad \bar{\phi}(h)_0 = u. \tag{A.1}$$

We write $\bar{\psi}(h)_t = \pi(\bar{\phi}(h)_t)$.

The purpose of this appendix is to verify that

$$p^\varepsilon_t(x, a) > 0 \quad \text{for every } x, a \in \mathcal{M} \text{ and } \varepsilon, t \in (0, 1]. \tag{A.2}$$

By the scaling property $p^\varepsilon_t(x, a) = p^1_{\varepsilon^2 t}(x, a)$, it is enough to prove (A.2) for $\varepsilon = 1$. This kind of positivity for the density of Wiener functionals has been well studied (see [2] and Remark A.2 below). According to these results, Lemma A.1 below implies (A.2).

Lemma A.1. *Let $x, a \in \mathcal{M}$ and $u \in \pi^{-1}(x)$. Then, for every $\tau \in (0,1]$, there exists $h \in \mathcal{H}$ such that $\overline{\psi}(h)_\tau = a$ and $D\overline{\psi}(h)_\tau : \mathcal{H} \to T_a\mathcal{M}$ is surjective.*

Proof. It is sufficient to prove the case $\tau = 1$. The general case can be done with trivial modifications.

We now make two simple remarks. First, let $\{Z_1, \ldots, Z_d\}$ be an orthonormal frame of $\mathcal{D}$ on a coordinate neighborhood $U \subset \mathcal{M}$. Consider the following ODE on U driven by a Cameron–Martin path h:

$$d\zeta(h)_t = \sum_{i=1}^{d} Z_i(\zeta(h)_t)\, dh_t^i + V_0(\zeta(h)_t)\, dt. \tag{A.3}$$

Then $\zeta = \zeta(h)$ is of finite energy and satisfies $\zeta_t' - V_0(\zeta_t) \in \mathcal{D}_{\zeta_t}$ for almost all t. Conversely, if ζ is a path on $\mathcal{M}$ of finite energy such that $\zeta_t' - V_0(\zeta_t) \in \mathcal{D}_{\zeta_t}$ for almost all t, then there uniquely exists a Cameron–Martin path h such that $\zeta = \zeta(h)$. In this sense, we have a one-to-one correspondence $h \leftrightarrow \zeta(h)$.

Second, consider $\zeta \in \mathcal{H}_x(\mathcal{M}, T\mathcal{M})$ such that $\zeta_t' - V_0(\zeta_t) \in \mathcal{D}_{\zeta_t}$ for almost all $t \in [0,1]$. We denote by ξ the horizontal lift of ζ with $\xi_0 = u$. Since A_0 is the horizontal lift of V_0, $\xi_t' - A_0(\xi_t) \in \mathcal{K}_{\xi_t}$ for almost all $t \in [0,1]$. Since $\{A_1(u), \ldots, A_d(u)\}$ forms a linear basis of $\mathcal{K}_u$ for all $u \in \mathcal{P}$, there uniquely exists $h \in \mathcal{H}$ such that $\xi = \overline{\phi}(h)$. Conversely, if $\overline{\phi}(h)$ solves ODE (A.1), then $\zeta := \overline{\psi}(h)$ satisfies $\zeta_t' - V_0(\zeta_t) \in \mathcal{D}_{\zeta_t}$ for almost all t. In this way, we have a one-to-one correspondence $h \leftrightarrow \zeta$.

By using Proposition 3.3 and Lemma 4.3, we can now prove Lemma A.1 in a similar way to the proof of Lemma 6.1.

Let U be a coordinate neighborhood of x and consider ODE (A.3) with $\zeta(h)_0 = x$ on U. By extending the coefficient vector fields with compact support, we also view (A.3) as an ODE on $\mathbb{R}^n$. Since these vector fields satisfy Hörmander's condition at x, we can use Lemma 4.3. It implies that there exists a Cameron–Martin path $k : [0, 1/2] \to \mathbb{R}^d$ such that (1) the tangent map of $\zeta(\,\cdot\,)_{1/2}$ is surjective at k and (2) $[0, 1/2] \ni t \mapsto \zeta(k)_t$ stays inside U. By Proposition 3.3, there exists a finite energy path $\eta : [1/2, 1] \to \mathcal{M}$ such that $\eta_{1/2} = \zeta(k)_{1/2}$, $\eta_1 = a$, and $\eta_t' - V_0(\eta_t) \in \mathcal{D}_{\eta_t}$ for almost all t.

Define $\zeta \in \mathcal{H}_x(\mathcal{M}, T\mathcal{M})$ to be the concatenation of $\zeta(k)$ and η. Then $\zeta_1 = a$ and $\zeta_t' - V_0(\zeta_t) \in \mathcal{D}_{\zeta_t}$ for almost all t. The corresponding $h \in \mathcal{H}$ is the desired element. (The proof of the surjectivity of $D\overline{\psi}(h)_1$ is essentially the same as in the proof of Lemma 6.1 and is therefore omitted.) □

Remark A.2. Precisely speaking, the positivity theorem for the density in [2] is for SDEs on a Euclidean space. But, after slightly modifying it, one can verify that it still holds for SDEs on a compact manifold.

Acknowledgements

This research is partly supported by JSPS KAKENHI (grant no. 20H01807).

References

[1] S. Aida, Vanishing of one-dimensional L^2-cohomologies of loop groups, J. Funct. Anal. **261** (2011), 2164–2213. Zbl 1236.58020 MR 2824575

[2] S. Aida, S. Kusuoka, and D. Stroock, On the support of Wiener functionals, in *Asymptotic problems in probability theory: Wiener functionals and asymptotics (Sanda/Kyoto, 1990)*, Pitman Research Notes in Mathematics Series 284, Longman Scientific & Technical, Harlow, 1993, 3–34. Zbl 0790.60047 MR 1354161

[3] I. Bailleul, Large deviation principle for bridges of sub-Riemannian diffusion processes, in *Séminaire de Probabilités XLVIII*, Lecture Notes in Mathematics 2168, Springer, Cham, 2016, 189–198. Zbl 1372.60027 MR 3618130

[4] F. Baudoin, Q. Feng, and M. Gordina, Integration by parts and quasi-invariance for the horizontal Wiener measure on foliated compact manifolds, J. Funct. Anal. **277** (2019), 1362–1422. Zbl 1427.53028 MR 3969194

[5] I. Beschastnyi, K. Habermann, and A. Medvedev, Cartan connections for stochastic developments on sub-Riemannian manifolds, J. Geom. Anal. **32** (2022), art. no. 13, 37 pp. Zbl 1515.58009 MR 4349457

[6] A. Dembo and O. Zeitouni, *Large deviations techniques and applications*, 2nd ed., Applications of Mathematics (New York) 38, Springer, New York, 1998. Zbl 0896.60013 MR 1619036

[7] K. D. Elworthy, Y. Le Jan, and X.-M. Li, Equivariant diffusions on principal bundles, in *Stochastic analysis and related topics in Kyoto*, Advanced Studies in Pure Mathematics 41, Mathematical Society of Japan, Tokyo, 2004, 31–47. Zbl 1071.58030 MR 2083702

[8] K. D. Elworthy, Y. Le Jan, and X.-M. Li, *The geometry of filtering*, Frontiers in Mathematics, Birkhäuser, Basel, 2010. Zbl 1219.58017 MR 2731662

[9] P. K. Friz and N. B. Victoir, *Multidimensional stochastic processes as rough paths*, Cambridge Studies in Advanced Mathematics 120, Cambridge University Press, Cambridge, 2010. Zbl 1193.60053 MR 2604669

[10] E. Grong and A. Thalmaier, Curvature-dimension inequalities on sub-Riemannian manifolds obtained from Riemannian foliations: Part I, Math. Z. **282** (2016), 99–130. Zbl 1361.53028 MR 3448376

[11] E. P. Hsu, *Stochastic analysis on manifolds*, Graduate Studies in Mathematics 38, American Mathematical Society, Providence, RI, 2002. Zbl 0994.58019 MR 1882015

[12] P. Hsu, Brownian bridges on Riemannian manifolds, Probab. Theory Related Fields **84** (1990), 103–118. Zbl 0698.58052 MR 1027823

[13] Y. Hu, *Analysis on Gaussian spaces*, World Scientific, Hackensack, NJ, 2017. Zbl 1386.60005 MR 3585910

[14] N. Ikeda and S. Watanabe, *Stochastic differential equations and diffusion processes*, 2nd ed., North-Holland Mathematical Library 24, North-Holland, Amsterdam; Kodansha, Tokyo, 1989. Zbl 0684.60040 MR 1011252

[15] Y. Inahama, Large deviation principle of Freidlin–Wentzell type for pinned diffusion processes, Trans. Amer. Math. Soc. **367** (2015), 8107–8137. Zbl 1338.60083 MR 3391911

[16] Y. Inahama, Large deviations for rough path lifts of Watanabe's pullbacks of delta functions, Int. Math. Res. Not. IMRN **2016** (2016), 6378–6414. Zbl 1404.60110 MR 3579967

[17] Y. Inahama and S. Taniguchi, Short time full asymptotic expansion of hypoelliptic heat kernel at the cut locus, Forum Math. Sigma **5** (2017), art. no. e16, 74 pp. Zbl 1369.60040 MR 3669328

[18] H. Kondo and S. Taniguchi, A construction of diffusion processes associated with sub-Laplacian on CR manifolds and its applications, J. Math. Soc. Japan **69** (2017), 111–125. Zbl 1362.32024 MR 3597549

[19] H. Kunita, Supports of diffusion processes and controllability problems, in *Proceedings of the International Symposium on Stochastic Differential Equations (Res. Inst. Math. Sci., Kyoto Univ., Kyoto, 1976)*, Wiley-Interscience, John Wiley & Sons, New York-Chichester-Brisbane, 1978, 163–185. Zbl 0409.60063 MR 536011

[20] T. J. Lyons, M. Caruana, and T. Lévy, *Differential equations driven by rough paths*, Lecture Notes in Mathematics 1908, Springer, Berlin, 2007. Zbl 1176.60002 MR 2314753

[21] P. Malliavin, *Stochastic analysis*, Grundlehren der mathematischen Wissenschaften 313, Springer, Berlin, 1997. Zbl 0878.60001 MR 1450093

[22] H. Matsumoto and S. Taniguchi, *Stochastic analysis*, Japanese ed., Cambridge Studies in Advanced Mathematics 159, Cambridge University Press, Cambridge, 2017. Zbl 1375.60011 MR 3586279

[23] D. Nualart, *The Malliavin calculus and related topics*, 2nd ed., Probability and Its Applications (New York), Springer, Berlin, 2006. Zbl 1099.60003 MR 2200233

[24] L. Rifford, *Sub-Riemannian geometry and optimal transport*, SpringerBriefs in Mathematics, Springer, Cham, 2014. Zbl 1454.49003 MR 3308395

[25] I. Shigekawa, *Stochastic analysis*, Translations of Mathematical Monographs 224, American Mathematical Society, Providence, RI, 2004. Zbl 1064.60003 MR 2060917

[26] H. Sugita, Positive generalized Wiener functions and potential theory over abstract Wiener spaces, Osaka J. Math. **25** (1988), 665–696. Zbl 0737.46038 MR 969026

[27] S. Takanobu and S. Watanabe, Asymptotic expansion formulas of the Schilder type for a class of conditional Wiener functional integrations, in *Asymptotic problems in probability theory: Wiener functionals and asymptotics (Sanda/Kyoto, 1990)*, Pitman Research Notes in Mathematics Series 284, Longman Scientific & Technical, Harlow, 1993, 194–241. Zbl 0801.60020 MR 1354169

[28] S. Taniguchi, Malliavin's stochastic calculus of variations for manifold-valued Wiener functionals and its applications, Z. Wahrsch. Verw. Gebiete **65** (1983), 269–290. Zbl 0507.60068 MR 722132

[29] A. Thalmaier, Geometry of subelliptic diffusions, in *Geometry, analysis and dynamics on sub-Riemannian manifolds. Vol. II*, EMS Series of Lectures in Mathematics, European Mathematical Society, Zürich, 2016, 85–169. Zbl 1516.35200 MR 3642644

[30] N. Yoshida, Asymptotic expansions of maximum likelihood estimators for small diffusions via the theory of Malliavin–Watanabe, Probab. Theory Related Fields **92** (1992), 275–311. Zbl 0767.60035 MR 1165514