Available online at www.sciencedirect.com

ScienceDirect

Stochastic Processes and their Applications 121 (2011) 759–792

stochastic processes and their applications

www.elsevier.com/locate/spa

Small-time kernel expansion for solutions of stochastic differential equations driven by fractional Brownian motions

Fabrice Baudoin*, Cheng Ouyang

Department of Mathematics, Purdue University, West Lafayette, IN 47907, United States

Received 20 May 2010; received in revised form 2 November 2010; accepted 16 November 2010
Available online 5 December 2010

Abstract

The goal of this paper is to show that under some assumptions, for a d-dimensional fractional Brownian motion with Hurst parameter $H > 1/2$, the density of the solution of the stochastic differential equation

$$X_t^x = x + \sum_{i=1}^{d} \int_0^t V_i(X_s^x) \mathrm{d}B_s^i,$$

admits the following asymptotics at small times:

$$p(t; x, y) = \frac{1}{(t^H)^d} \mathrm{e}^{-\frac{d^2(x,y)}{2t^{2H}}} \left(\sum_{i=0}^{N} c_i(x, y) t^{2iH} + O(t^{2(N+1)H}) \right).$$

MSC: 28D05; 60D58

Keywords: Fractional Brownian motion; Small times expansion; Laplace method; Stochastic differential equation

* Corresponding author.
E-mail addresses: fbaudoin@math.purdue.edu, fbaudoin@purdue.edu (F. Baudoin), couyang@math.purdue.edu (C. Ouyang).

0304-4149/$ - see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.spa.2010.11.011

1. Introduction

In this paper, we are interested in the study of stochastic differential equations on $\mathbb{R}^d$ at small times:

$$X_t^x = x + \sum_{i=1}^d \int_0^t V_i(X_s^x) \mathrm{d}B_s^i \tag{1.1}$$

where V_i's are C^∞-bounded vector fields on $\mathbb{R}^d$ and B is a d-dimensional fractional Brownian motion with Hurst parameter $H > 1/2$. Since $H > 1/2$, the integrals $\int_0^t V_i(X_s^x)\mathrm{d}B_s^i$ are understood in the sense of Young's integration (see [23,24]), and it is known (see e.g. [20]) that an equation like (1.1) has one and only one solution. Throughout our discussion, we make the following two-part assumption.

Assumption 1.1.

- A1: For every $x \in \mathbb{R}^d$, the vectors $V_1(x), \dots, V_d(x)$ form a basis of $\mathbb{R}^d$.
- A2: There exist smooth and bounded functions ω_{ij}^l such that

  $$[V_i, V_j] = \sum_{l=1}^d \omega_{ij}^l V_l,$$

  and

  $$\omega_{ij}^l = -\omega_{il}^j.$$

The first assumption means that the vector fields form an elliptic differential system. As a consequence of the work of Baudoin and Hairer [6], it is known that the law of $X_t, t > 0$, admits therefore a smooth density $p(t; x, y)$ with respect to Lebesgue measure (also see [21]). The second assumption is of geometric nature and actually means that the Levi-Civita connection associated with the Riemannian structure given by the vector fields V_i is

$$\nabla_X Y = \frac{1}{2}[X, Y].$$

In a Lie group structure, this is equivalent to the fact that the Lie algebra is of compact type, or in other words that the adjoint representation is unitary. Our main result is the following:

Theorem 1.2. *Under the above assumption, in a neighborhood V of x, the density function $p(t; x, y)$ of X_t^x in (1.1) has the following asymptotic expansion near $t = 0$:*

$$p(t; x, y) = \frac{1}{(t^H)^d} \mathrm{e}^{-\frac{d^2(x,y)}{2t^{2H}}} \left(\sum_{i=0}^N c_i(x, y) t^{2iH} + r_{N+1}(t, x, y) t^{2(N+1)H} \right), \quad y \in V.$$

Here $c_0(x, y) > 0$ and $d(x, y)$ is the Riemannian distance between x and y determined by the vector fields $V_1, \dots, V_d$. Moreover, we can chose V such that $c_i(x, y)$ are C^∞ in $V \times V \subset \mathbb{R}^d \times \mathbb{R}^d$, and for all multi-indices α and β

$$\sup_{t \le t_0} \sup_{(x,y) \in V \times V} |\partial_x^\alpha \partial_y^\beta \partial_t^k r_{N+1}(t, x, y)| < \infty$$

for some $t_0 > 0$.

As a first corollary of the above theorem, we observe that it implies the Varadhan type asymptotics

$$\lim_{t\to 0} t^{2H} \ln p(t; x, y) = -\frac{d^2(x, y)}{2}.$$

For $H = 1/2$, which corresponds to the Brownian motion case, the above theorem admits numerous proofs. The first proofs were analytic and based on the parametrix method. Such methods do not apply in the present framework since the Markov property for X_t^x is lost whenever $H > 1/2$. However, in the seminal works [1,2], Azencott introduced probabilistic methods to prove the result. These methods introduced by Azencott were then further developed by Ben Arous and Léandre in [7–9,15], in order to cover the case of hypoelliptic heat kernels. In this work, we follows Ben Arous' approach [8], the strategy of which is sketched as follows.

The first idea is to consider the scaled stochastic differential equation

$$\mathrm{d}X_t^\varepsilon = \varepsilon \sum_{i=1}^n V_i(X_t^\varepsilon)\mathrm{d}B_t^i, \quad \text{with } X_0^\varepsilon = x_0.$$

We observe that there exist neighborhoods U and V of x_0 and a bounded smooth function $F(x, y, z)$ on $U \times V \times \mathbb{R}^d$ such that:

(1) For any $(x, y) \in U \times V$ the infimum

$$\inf\left\{F(x, y, z) + \frac{d(x, z)^2}{2}, z \in \mathbb{R}^n\right\} = 0$$

is attained at the unique point y.

(2) For each $(x, y) \in U \times V$, there exists a ball centered at y with radius r independent of x, y such that $F(x, y, \cdot)$ is a constant outside of the ball.

So, denoting by $p_\varepsilon(x_0, y)$ the density of X_1^ε, by the Fourier inversion formula we have

$$\begin{aligned} p_\varepsilon(x_0, y)\mathrm{e}^{-\frac{F(x_0,y,y)}{\varepsilon^2}} &= \frac{1}{(2\pi)^d}\int \mathrm{e}^{-\mathrm{i}\zeta\cdot y}\mathrm{d}\zeta \int \mathrm{e}^{\mathrm{i}\zeta\cdot z}\mathrm{e}^{-\frac{F(x_0,y,z)}{\varepsilon^2}} p_\varepsilon(x_0, z)\mathrm{d}z \\ &= \frac{1}{(2\pi\varepsilon)^d}\int \mathrm{d}\zeta\,\mathbb{E}\left(\mathrm{e}^{\frac{\mathrm{i}\zeta\cdot(X_1^\varepsilon - y)}{\varepsilon}}\mathrm{e}^{-\frac{F(x_0,y,X_1^\varepsilon)}{\varepsilon^2}}\right). \end{aligned}$$

Thus, the asymptotics of $p_t(x_0, y)$ may be understood from the asymptotics when $\varepsilon \to 0$ of

$$J_\varepsilon(x_0, y) = \mathbb{E}\left(\mathrm{e}^{\frac{\mathrm{i}\zeta\cdot(X_1^\varepsilon - y)}{\varepsilon}}\mathrm{e}^{-\frac{F(x_0,y,X_1^\varepsilon)}{\varepsilon^2}}\right).$$

Then, by using the Laplace method on the Wiener space based on the large deviation principle, we get an expansion in powers of ε of $J_\varepsilon(x_0, y)$ which leads to the expected asymptotics for the density function.

Finally, let us explain where assumption A2 is needed, which is also a major difference from the classical case. This assumption essentially means that the derivative of the Itô map is unitary (see Assumption $\mathcal{H}$, Main theorem, page 278 in [17], for a precise meaning in a Lie group framework for the Brownian motion case).

It particularly implies that the Riemannian distance is the control distance associated with the equation that we consider. More precisely, denote by $\Phi_.(x, k)$ the solution of the ordinary differential equation

$$x_t = x + \sum_{i=1}^{d} \int_0^t V_i(x_s) \mathrm{d}k_s^i$$

where k is an element from the Cameron–Martin space $\mathscr{H}_H$ of the underlying fractional Brownian motion B. Define

$$D^2(x, y) = \inf_{k \in \mathscr{H}_H, \Phi_1(x,k)=y} \|k\|^2_{\mathscr{H}_H}.$$

The assumption A2 implies (see Proposition 4.3 for more details)

$$D^2(x, y) = d^2(x, y).$$

Here $d(x, y)$ is the Riemannian distance. When assumption A2 is dropped, it is clear from the above that one needs to construct a bounded smooth function $F(x, y, z)$ on $U \times V \times \mathbb{R}^d$ associated with $D^2(x, y)$ (instead of $d^2(x, y)$ as in the classical case). On the other hand, the original construction of F in the classical case depends on some smoothness of $d^2(x, y)$, which is not clear for $D^2(x, y)$ in our case. This causes a technical difficulty in the present work.

Another reason that we want to work with $d(x, y)$ is that, in this case, our result would be useful for estimating probability in a small ball of x, the starting point of the process X. Without assumption A2, one should expect to replace $d^2(x, y)$ by $D^2(x, y)$ in the result of Theorem 1.2 (assuming that the technical difficulty specified above could be overcome). But in this case we expect $D(x, y)$ not to induce a metric in $\mathbb{R}^d$ in general, and it is not clear to the authors at this moment whether $D(x, y)$ is comparable to $d(x, y)$ or not.

The rest of this paper is organized as follows. In a preliminary section we recall some known facts about fractional Brownian motion and equations driven by it. In the second section we show how the Laplace method may be carried out in the fractional Brownian motion case and finally in the third section, which is the heart of the present paper, we prove Theorem 1.2. We move the proofs of some technical lemmas to the Appendix.

Remark 1.3. Under the framework of this present work, the Laplace method can be obtained in the general hypoelliptic case and without imposing the structure equations on the vector fields in Assumption 1.1. These two assumptions are imposed only to obtain the correct Riemannian distance in the kernel expansion.

Remark 1.4. When $H > 1/2$, to obtain a short-time asymptotic formula for the density of the solution to Eq. (1.1) but with drift, one needs to work on a version of Laplace method with fractional powers of ε, which will be very onerous and tedious in computation.

Remark 1.5. When the present work was almost completed, we noticed that a proof for the Laplace method for stochastic differential equations driven by fractional Brownian motion with Hurst parameter $1/3 < H < 1/2$ became available from Inahama [13] on the mathematics arXiv.

2. Preliminaries

2.1. *Stochastic differential equations driven by fractional Brownian motions*

We consider the Wiener space of continuous paths

$$\mathbb{W}^{\otimes d} = \left(\mathcal{C}([0, T], \mathbb{R}^d), (\mathcal{B}_t)_{0 \le t \le T}, \mathbb{P} \right)$$

where:

(1) $\mathcal{C}([0,T],\mathbb{R}^d)$ is the space of continuous functions $[0,T]\to\mathbb{R}^d$;
(2) $(\beta_t)_{t\geq 0}$ is the coordinate process defined by $\beta_t(f)=f(t)$, $f\in\mathcal{C}([0,T],\mathbb{R}^d)$;
(3) $\mathbb{P}$ is the Wiener measure;
(4) $(\mathcal{B}_t)_{0\leq t\leq T}$ is the ($\mathbb{P}$-completed) natural filtration of $(\beta_t)_{0\leq t\leq T}$.

A d-dimensional fractional Brownian motion with Hurst parameter $H\in(0,1)$ is a Gaussian process

$$B_t=(B_t^1,\dots,B_t^d),\quad t\geq 0,$$

where $B^1,\dots,B^d$ are d independent centered Gaussian processes with covariance function

$$R(t,s)=\frac{1}{2}\left(s^{2H}+t^{2H}-|t-s|^{2H}\right).$$

It can be shown that such a process admits a continuous version whose paths are Hölder p continuous, $p<H$. Throughout this paper, we will always consider the 'regular' case, $H>1/2$. In this case the fractional Brownian motion can be constructed on the Wiener space by a Volterra type representation (see [10]). Namely, under the Wiener measure, the process

$$B_t=\int_0^t K_H(t,s)\mathrm{d}\beta_s,\quad t\geq 0 \tag{2.1}$$

is a fractional Brownian motion with Hurst parameter H, where

$$K_H(t,s)=c_H s^{\frac{1}{2}-H}\int_s^t (u-s)^{H-\frac{3}{2}}u^{H-\frac{1}{2}}\mathrm{d}u,\quad t>s$$

and c_H is a suitable constant.

Denote by $\mathcal{E}$ the set of step functions on $[0,T]$. Let $\mathcal{H}$ be the Hilbert space defined as the closure of $\mathcal{E}$ with respect to the scalar product

$$\langle \mathbf{1}_{[0,t]},\mathbf{1}_{[0,s]}\rangle_{\mathcal{H}}=R_H(t,s).$$

The isometry K_H^* from $\mathcal{H}$ to $L^2([0,T])$ is given by

$$(K_H^*\varphi)(s)=\int_s^T\varphi(t)\frac{\partial K_H}{\partial t}(t,s)\mathrm{d}t.$$

Moreover, for any $\varphi\in L^2([0,T])$ we have

$$\int_0^T\varphi(s)\mathrm{d}B_s=\int_0^T(K_H^*\varphi)(s)\mathrm{d}\beta_s.$$

We consider the following stochastic differential equation:

$$X_t^x=x+\int_0^t V_0(X_s^x)\mathrm{d}s+\sum_{i=1}^d\int_0^t V_i(X_s^x)\mathrm{d}B_s^i \tag{2.2}$$

where the V_i's are C^∞ vector fields on $\mathbb{R}^d$ with bounded derivatives to any order and B is the d-dimensional fractional Brownian motion defined by (2.1). The existence and uniqueness of solutions for such equations have been widely studied and are known to hold in this framework.

2.1.1. Pathwise estimates

Let us have $1/2 < \lambda < H$ and denote by $C^\lambda(0, T; \mathbb{R}^d)$ the space of λ-Hölder continuous functions equipped with the λ-Hölder norm

$$\|f\|_{\lambda,T} := \|f\|_\infty + \sup_{0\le s<t\le T} \frac{|f(t) - f(s)|}{(t-s)^\lambda},$$

where $\|f\|_\infty := \sup_{t\in[0,T]} |f(t)|$.

The following remarks will be useful later.

Remark 2.1.

1. It is clear that if $f_1,\ f_2 \in C^\lambda$, then $f_1 f_2 \in C^\lambda$ with $\|f_1 f_2\|_{\lambda,t} \le \|f_1\|_{\lambda,t} \|f_2\|_{\lambda,t}$. Therefore, polynomials of elements in C^λ are still in C^λ. It is also clear that whenever φ is a Lipschitz function and $f \in C^\lambda$, we have $\varphi(f) \in C^\lambda$.
2. Let $f \in C^\lambda(0, T; \mathbb{R}^d)$ and $g : [0, T] \to \mathcal{M}_{n\times d}$ be a matrix-valued function and suppose $g \in C^\lambda$. By standard argument (see [16] for instance),
$$\int_0^\cdot g_s \, \mathrm{d}f_s \in C^\lambda(0, T; \mathbb{R}^n)$$
with
$$\left\| \int_0^\cdot g_s \, \mathrm{d}f_s \right\|_{\lambda,T} \le C \|g\|_{\lambda,T} \|f\|_{\lambda,T}.$$
In the above, C is a constant only depending on λ and T.

Lemma 2.2 (*Hu and Nualart [12]*). *Consider the stochastic differential equation (1.2), and assume that $\mathbb{E}(|X_0|^p) < \infty$ for all $p \ge 2$. If the derivatives of the V_i's are bounded and Hölder continuous of order $\lambda > 1/H - 1$, then*

$$\mathbb{E}\left(\sup_{0\le t\le T} |X_t|^p \right) < \infty$$

for all $p \ge 2$. If furthermore the V_i's are bounded and $\mathbb{E}(\exp(\lambda|X_0|^q)) < \infty$ for any $\lambda > 0$ and $q < 2H$, then

$$\mathbb{E}\left(\exp \lambda \left(\sup_{0\le t\le T} |X_t|^q \right) \right) < \infty$$

for any $\lambda > 0$ and $q < 2H$.

2.2. The Cameron–Martin theorem for fBm

Consider the classical Cameron–Martin space $\mathscr{H} = \{h \in P_o(\mathbb{R}^d) : \|h\|_{\mathscr{H}} < \infty\}$, where

$$\|h\|_{\mathscr{H}} = \left(\int_0^T |\dot{h}_s|^2 \mathrm{d}s \right)^{\frac{1}{2}}.$$

The Cameron–Martin space for the fractional Brownian motion B is

$$\mathscr{H}_H = K_H(\mathscr{H}),$$

where the map $K_H : \mathscr{H} \to \mathscr{H}_H$ is given by

$$(K_H h)_t = \int_0^t K_H(t,s)\dot{h}_s \mathrm{d}s, \quad \text{for all } h \in \mathscr{H}.$$

The inner product on $\mathscr{H}_H$ is defined by

$$\langle k_1, k_2 \rangle_{\mathscr{H}_H} = \langle h_1, h_2 \rangle_{\mathscr{H}}, \quad k_i = K_H h_i,\ i = 1, 2.$$

Hence K_H is an isometry between $\mathscr{H}$ and $\mathscr{H}_H$.

Remark 2.3. It can be shown that when $\gamma \in \mathscr{H}_H$, γ is H-Hölder continuous.

The following Cameron–Martin theorem is known (see [10]).

Theorem 2.4 (*Cameron–Martin Theorem for fBm*)**.** *Let $B^k = B + k$ be the shifted fractional Brownian motion, where $k \in \mathscr{H}_H$ is a Cameron–Martin path. The law $\mathbb{P}_H^k$ of B^k and the law $\mathbb{P}_H$ of B are mutually absolutely continuous. Furthermore, the Radon–Nikodym derivative is given by*

$$\frac{\mathrm{d}\mathbb{P}_H^k}{\mathrm{d}\mathbb{P}_H} = \exp\left[-\int_0^T (K_H^*)^{-1}(\dot{h})_s \mathrm{d}B_s - \frac{1}{2}\|k\|_{\mathscr{H}_H}^2\right].$$

In the above, $h = (K_H)^{-1}k$ and the integral with respect to B is understood as Young's integral.

2.3. The large deviation principle for fBm

The following large deviation principle for stochastic differential equations driven by fractional Brownian motion is a special case of Proposition 19.14 in [11] (see also [18]).

Proposition 2.5. *Fix $\lambda \in (1/2, H)$. Let X^ε be the solution to the following stochastic differential equations driven by fBm B:*

$$X_t^\varepsilon = x_0 + \int_0^t V_0(X_s)\mathrm{d}s + \sum_{i=1}^d \varepsilon \int_0^t V_i(X_s)\mathrm{d}B_s^i \tag{2.3}$$

where the V_i's are C^∞ vector fields on $\mathbb{R}^d$ with bounded derivatives to any order. The process X^ε satisfies a large deviation principle, in λ-Hölder topology, with a good rate function given by

$$\Lambda(\phi) = \inf\{\bar{\Lambda}(\gamma) : \phi = I(\gamma)\}$$

where I is the Itô map given by (2.3) with ε being replaced by 1, and $\bar{\Lambda}$ is given by

$$\bar{\Lambda}(\gamma) = \begin{cases} \frac{1}{2}\|\gamma\|_{\mathscr{H}_H}^2 & \text{if } \gamma \in \mathscr{H}_H, \\ +\infty & \text{otherwise.} \end{cases}$$

3. The Laplace method

Consider the following stochastic differential equation driven by fractional Brownian motion on $\mathbb{R}^d$:

$$X_t^\varepsilon = x_0 + \int_0^t V_0(X_s)\mathrm{d}s + \sum_{i=1}^d \varepsilon \int_0^t V_i(X_s)\mathrm{d}B_s^i.$$

For the convenience of our discussion, in what follows, we write the above equation in the following form:

$$X_t^\varepsilon = x + \varepsilon \int_0^t \sigma(X_s^\varepsilon)\mathrm{d}B_s + \int_0^t b(\varepsilon, X_s^\varepsilon)\mathrm{d}s,$$

where σ is a smooth $d \times d$ matrix and b a smooth function from $\mathbb{R}^+ \times \mathbb{R}^d$ to $\mathbb{R}^d$. We also assume that σ and b have bounded derivatives to any order.

Fix $1/2 < \lambda < H$. Let F and f be two bounded infinitely Fréchet differentiable functionals on $C^\lambda(0, T; \mathbb{R}^d)$ with bounded derivatives (as linear operators) to any order. We are interested in studying the asymptotic behavior of

$$J(\varepsilon) = \mathbb{E}\big[f(X^\varepsilon)\exp\{-F(X^\varepsilon)/\varepsilon^2\}\big], \quad \text{as } \varepsilon \downarrow 0.$$

For each $k \in \mathscr{H}_H$, denote by $\Phi(k)$ (which is the deterministic Itô map) the solution to the following deterministic differential equation:

$$\mathrm{d}u_t = \sigma(u_t)\mathrm{d}k_t + b(0, u_t)\mathrm{d}t, \quad \text{with } u_0 = x. \tag{3.1}$$

Throughout our discussion we make the following assumptions:

Assumption 3.1.

- H1: $F + \Lambda$ attains its minimum for a finite number of paths $\phi_1, \phi_2, \dots, \phi_n$ on $P(\mathbb{R}^d)$.
- H2: For each $i \in \{1, 2, \dots, n\}$, we have $\phi_i = \Phi(\gamma_i)$ and γ_i is a non-degenerate minimum of the functional $F \circ \Phi + 1/2\|\cdot\|^2_{\mathscr{H}_H}$, i.e.,

$$\forall k \in \mathscr{H}_H - \{0\}, \quad d^2(F \circ \Phi + 1/2\|\cdot\|^2_{\mathscr{H}_H})(\gamma_i)k^2 > 0.$$

The following theorem is the main result of this section.

Theorem 3.2. *Under the assumptions* H1 *and* H2 *above, we have*

$$J(\varepsilon) = \mathrm{e}^{-\frac{a}{\varepsilon^2}}\mathrm{e}^{-\frac{c}{\varepsilon}}\left(\alpha_0 + \alpha_1\varepsilon + \cdots + \alpha_N\varepsilon^N + O(\varepsilon^{N+1})\right).$$

Here

$$a = \inf\{F + \Lambda(\phi), \phi \in P(\mathbb{R}^d)\} = \inf\{F \circ \Phi(k) + 1/2|k|^2_{\mathscr{H}_H},\ k \in \mathscr{H}_H\}$$

and

$$c = \inf\big\{\mathrm{d}F(\phi_i)Y_i,\ i \in \{1, 2, \dots, n\}\big\},$$

where Y_i is the solution of

$$\mathrm{d}Y_i(s) = \partial_x\sigma(\phi_i(s))Y_i(s)\mathrm{d}\gamma_i(s) + \partial_\varepsilon b(0, \phi_i(s))\mathrm{d}s + \partial_x b(0, \phi_i(s))Y_i(s)\mathrm{d}s$$

with $Y_i(0) = 0$.

Lemma 3.3. *Let Φ be defined as above; we have*

$$\Lambda(\phi) = \inf\left\{\frac{1}{2}\|k\|^2_{\mathscr{H}_H}, \phi = \Phi(k),\ k \in \mathscr{H}_H\right\}.$$

Moreover, if $\Lambda(\phi) < \infty$, there exists a unique $k \in \mathscr{H}_H$ such that $\Phi(k) = \phi$ and $\Lambda(\phi) = 1/2\|k\|^2_{\mathscr{H}_H}$.

Proof. The first statement is apparent. For the second statement, we only need to notice that if

$$\phi = \Phi(k_1) = \Phi(k_2), \quad k_1, k_1 \in \mathscr{H}_H,$$

then

$$\int_0^t \sigma(\phi_s) d(k_1 - k_2)_s = 0, \quad t \in [0, T],$$

which implies that $k_1 = k_2$, since we assume that columns of σ are linearly independent. The proof is therefore completed. □

Lemma 3.4. *Under assumption* H1*, we have*

$$a \stackrel{\text{def}}{=} \inf\{F + \Lambda(\phi), \phi \in P(\mathbb{R}^d)\} = \inf\left\{F \circ \Phi(k) + \frac{1}{2}\|k\|^2_{\mathscr{H}_H}, k \in \mathscr{H}_H\right\},$$

and the minimum is attained for n paths $\gamma_1, \gamma_2, \ldots, \gamma_n \in \mathscr{H}_H$ such that

$$\Phi(\gamma_i) = \phi_i$$

and

$$\frac{1}{2}\|\gamma_i\|^2_{\mathscr{H}_H} = \Lambda(\Phi(\gamma_i)).$$

Proof. This is a direct corollary of Lemma 3.3. □

Assumption H2 has a simple interpretation as follows. Let γ be one of the γ_i's above. Define a bounded self-adjoint operator on $\mathscr{H}$ by

$$\mathrm{d}^2 F \circ \Phi(\gamma)(K_H h^1, K_H h^2) = (A h^1, h^2)_{\mathscr{H}}, \quad \text{for } h^1, h^2 \in \mathscr{H}.$$

Lemma 3.5. *The bounded self-adjoint operator A is Hilbert–Schmidt.*

Proof. The proof is similar to that in [7] but with slight modification. Thus we only sketch the proof here. In what follows, k always denotes an element in $\mathscr{H}_H$ and $h = K_H^{-1} k$ its corresponding element in $\mathscr{H}$.

For any $k^1, k^2 \in \mathscr{H}_H$, we have

$$\begin{aligned}\mathrm{d}^2 F \circ \Phi(\gamma)(K_H h^1, K_H h^2) &= \mathrm{d}^2 F \circ \Phi(\gamma)(k^1, k^2)\\ &= \mathrm{d}^2 F(\mathrm{d}\Phi(\gamma) k^1, \mathrm{d}\Phi(\gamma) k^2) + \mathrm{d}F(\phi)(\mathrm{d}^2\Phi(\gamma)(k^1, k^2)).\end{aligned}$$

Let

$$\phi = \Phi(\gamma) \quad \text{and} \quad \chi(k) = \mathrm{d}\Phi(\gamma) k.$$

It can be shown (cf. [7]) that

$$\begin{aligned}&\mathrm{d}\phi_t = \sigma(\phi_t)\mathrm{d}\gamma_t + b(0, \phi_t)\mathrm{d}t, \quad \text{with } \phi_0 = x,\\ &\mathrm{d}\chi_t = \sigma(\phi_t)\mathrm{d}k_t + \partial_x\sigma(\phi_t)\chi_t\mathrm{d}\gamma_t + \partial_x b(0, \phi_t)\chi_t\mathrm{d}t, \quad \text{with } \chi_0 = 0,\end{aligned}$$

and

$$\begin{aligned}\mathrm{d}^2\Phi(\gamma)(k^1, k^2)(t) = &\int_0^1 Q(t, s)\partial_x\sigma(\phi_s)\big(\chi(k^1)_s \mathrm{d}k_s^2 + \chi(k^2)_s \mathrm{d}k_s^1\big)\\ &+ \int_0^t \partial_{xx}^2\sigma(\phi_s)\big(\chi(k^1)_s, \chi(k^2)_s\big)\mathrm{d}\gamma_s + \int_0^t \partial_{xx}^2 b(0, \phi_s)\big(\chi(k^1)_s, \chi(k^2)_s\big)\mathrm{d}s.\end{aligned}$$

Here $Q(t,s)$ takes the form

$$Q(t,s) = \partial_x \phi_t(x) \partial_x \phi_s(x)^{-1}.$$

Moreover, we have

$$\begin{aligned} \chi_t(k) &= \int_0^t Q(t,s)\sigma(\phi_s)\mathrm{d}k_s \\ &= \int_0^t \left(\int_u^t Q(t,s)\sigma(\phi_s) \frac{\partial K_H(s,u)}{\partial s}\mathrm{d}s \right) \dot{h}_u \mathrm{d}u. \end{aligned} \tag{3.2}$$

Set

$$\begin{aligned} V(h^1,h^2)(t) &= \int_0^t Q(t,s)\partial_x\sigma(\phi_s)\big(\chi(K_H h^1)_s d(K_H h^2)_s + \chi(K_H h^2)_s d(K_H h^1)_s\big) \\ &= \int_0^t Q(t,s)\partial_x\sigma(\phi_s)\big(\chi(k^1)_s \mathrm{d}k_s^2 + \chi(k^2)_s \mathrm{d}k_s^1\big) \\ &= \int_0^t \int_u^t Q(t,s)\partial_x\sigma(\phi_s) \frac{\partial K_H(s,u)}{\partial s}\big(\chi(k^1)_s h_u^2 + \chi(k^2)_s h_u^1\big)\mathrm{d}s\mathrm{d}u. \end{aligned} \tag{3.3}$$

Define a bounded self-adjoint operator $\tilde{A}$ from $\mathscr{H}$ to $\mathscr{H}$ by

$$\mathrm{d}F(\phi)(V(h^1,h^2)) = (\tilde{A}h^1, h^2)_{\mathscr{H}}.$$

We conclude that $\tilde{A}$ is Hilbert–Schmidt since, by (3.2) and (3.3), it is defined from an L^2 kernel. Therefore, to complete the proof, it suffices to show that $A - \tilde{A}$ is Hilbert–Schmidt.By the same argument as in [7], we only need to show that

$$\|\mathrm{d}\Phi(\gamma)K_H h\|_\infty = \|\chi(K_H h)\|_\infty \le C\|h\|_\infty, \quad \text{for all } h \in \mathscr{H}.$$

Indeed, by an easy application of the Gronwall inequality to the equation for χ, we have

$$\|\mathrm{d}\Phi(\gamma)(K_H h)\|_\infty \le \|K_H h\|_\infty.$$

Moreover, since

$$(K_H h)_t = \int_0^t K_H(t,s)\dot{h}_s \mathrm{d}s,$$

and noting that $\partial K_H(t,s)/\partial s \in L^1$, we have

$$|K_H h|_t \le \left| \int_0^t K_H(t,s)\dot{h}_s\mathrm{d}s \right| = \left| \int_0^t h_s \frac{\partial K_H(t,s)}{\partial s}\mathrm{d}s \right| \le \|h\|_\infty \int_0^t \left| \frac{\partial K_H(t,s)}{\partial s} \right| \mathrm{d}s.$$

The proof is completed. □

From the above lemma, assumption H2 simply means that the smallest eigenvalue of A is attained and is strictly greater that -1.

3.1. Localization around the minimum

By the large deviation principle, the sample paths that contribute to the asymptotics of $J(\varepsilon)$ lie in the neighborhoods of the minimizers of $F + \Lambda$. More precisely:

Lemma 3.6. *For $\rho > 0$, denote by $B(\phi_i, \rho)$ the open ball (under λ-Hölder topology) centered at ϕ_i with radius ρ. There exist $d > a$ and $\varepsilon_0 > 0$ such that for all $\varepsilon \leq \varepsilon_0$*

$$\left| J(\varepsilon) - \mathbb{E}\left[f(X_T^\varepsilon) \mathrm{e}^{-F(X_T^\varepsilon)/\varepsilon^2}, X^\varepsilon \in \bigcup_{1 \leq i \leq n} B(\phi_i, \rho) \right] \right| \leq \mathrm{e}^{-d/\varepsilon^2}.$$

Proof. This is a consequence of the large deviation principle. □

Assume that $n = 1$, i.e., $F + \Lambda$ attains its minimum on only one path ϕ. Let

$$J_\rho(\varepsilon) = \mathbb{E}\left[f(X_T^\varepsilon) \mathrm{e}^{-F(X_T^\varepsilon)/\varepsilon^2}, X^\varepsilon \in B(\phi, \rho) \right].$$

The above lemma tells us that to study the asymptotic behavior of $J(\varepsilon)$ as $\varepsilon \downarrow 0$, it is sufficient to study that of $J_\rho(\varepsilon)$.

3.2. Stochastic Taylor expansion and Laplace approximation

In this section, we prove an asymptotic expansion for $J_\rho(\varepsilon)$.

Let ϕ be the unique path that minimizes $F + \Lambda$. There exists a $\gamma \in \mathscr{H}_H$ such that

$$\phi = \Phi(\gamma), \quad \text{and} \quad \Lambda(\phi) = \frac{1}{2} \|\gamma\|_{\mathscr{H}_H}^2,$$

and for all $k \in \mathscr{H}_H - \{0\}$,

$$d^2 \left(F \circ \Phi + \frac{1}{2} \| \ \|_{\mathscr{H}_H}^2 \right) (\gamma) k^2 > 0.$$

Let

$$\chi(k) = \mathrm{d}\Phi(\gamma)k \quad \text{and} \quad \psi(k, k) = \mathrm{d}^2 \Phi(\gamma)(k, k).$$

We have

$$\mathrm{d}\chi_t = \sigma(\phi_t)\mathrm{d}k_t + \partial_x \sigma(\phi_t)\chi_t \mathrm{d}\gamma_t + \partial_x b(0, \phi_t)\chi_t \mathrm{d}t, \tag{3.4}$$

and

$$\begin{aligned} \mathrm{d}\psi_t &= 2\partial_x \sigma(\phi_t)\chi_t \mathrm{d}k_t + \partial_{xx}^2 \sigma(\phi_t)\chi_t^2 \mathrm{d}\gamma_t + \partial_x \sigma(\phi_t)\psi_t \mathrm{d}\gamma_t + \partial_{xx}^2 b(0, \phi_t)\chi_t^2 \mathrm{d}t \\ &\quad + \partial_x b(0, \phi_t)\psi_t \mathrm{d}t. \end{aligned} \tag{3.5}$$

Here $\chi_0 = \phi_0 = 0$. These formulas will be useful later.

Consider the following stochastic differential equation:

$$Z_t^\varepsilon = x + \int_0^t \sigma(Z_s^\varepsilon)(\varepsilon \mathrm{d}B_s + \mathrm{d}\gamma_s) + \int_0^t b(\varepsilon, Z_s^\varepsilon)\mathrm{d}s.$$

It is clear that $Z^0 = \phi$. Define $Z_t^{m,\varepsilon} = \partial_\varepsilon^m Z_t^\varepsilon$ and consider the Taylor expansion with respect to ε near $\varepsilon = 0$; we obtain

$$Z^\varepsilon = \phi + \sum_{j=0}^N \frac{g_j \varepsilon^j}{j!} + \varepsilon^{N+1} R_{N+1}^\varepsilon,$$

where $g_j = Z^{j,0}$. Explicitly, we have

$$\mathrm{d}g_1(s) = \sigma(\phi_s)\mathrm{d}B_s + \partial_x\sigma(\phi_s)g_1(s)\mathrm{d}\gamma_s + \partial_x b(0,\phi_s)g_1(s)\mathrm{d}s + \partial_\varepsilon b(0,\phi_s)\mathrm{d}s.$$

Like for the Brownian motion case, we have the following estimates, the proof of which is postponed to the Appendix.

Lemma 3.7. *For any $t \in [0,T]$, there exists a constant $C > 0$ such that for r large enough we have*

$$\mathbb{P}\{\|g_1\|_{\lambda,t} \geq r\} \leq \exp\left\{-\frac{Cr^2}{t^{2H}}\right\}$$
$$\mathbb{P}\{\|g_2\|_{\lambda,t} \geq r\} \leq \exp\left\{-\frac{Cr}{t^{2H}}\right\}$$

and on $\{t \leq T^\varepsilon\}$, where T^ε is the first exist time of Z^ε from $B(\phi,\rho)$, we have

$$\|\varepsilon R_1^\varepsilon\|_{\lambda,t} \leq \rho,$$
$$\mathbb{P}\{\|\varepsilon R_2^\varepsilon\|_{\lambda,t} \geq r; t \leq T^\varepsilon\} \leq \exp\left\{-\frac{Cr^2}{\rho t^{2H}}\right\},$$
$$\mathbb{P}\{\|\varepsilon R_3^\varepsilon\|_{\lambda,t} \geq r; t \leq T^\varepsilon\} \leq \exp\left\{-\frac{Cr}{\rho t^{2H}}\right\}.$$

Let $\theta(\varepsilon) = F(Z^\varepsilon)$. By Taylor expansion of $\theta(\varepsilon)$ with respect to ε, we obtain

$$\theta(\varepsilon) = \theta(0) + \varepsilon\theta'(0) + \varepsilon^2 U(\varepsilon).$$

Here

$$U(\varepsilon) = \int_0^1 (1-v)\theta''(\varepsilon v)\mathrm{d}v, \quad \text{and} \quad \theta(0) = F(\phi).$$

Lemma 3.8. *With the above notation, we have*

$$\theta'(0) = \mathrm{d}F(\phi)g_1 = -\int_0^T \left((K_H^*)^{-1}(K_H^{-1}\dot{\gamma})\right)_s \mathrm{d}B_s + \mathrm{d}F(\phi)Y.$$

Here Y is the solution of

$$\mathrm{d}Y_s = \partial_x\sigma(\phi_s)Y_s\mathrm{d}\gamma_s + \partial_\varepsilon b(0,\phi_s)\mathrm{d}s + \partial_x b(0,\phi_s)Y_s\mathrm{d}s, \qquad Y(0) = 0.$$

Proof. By an easy application of Gronwall's inequality to (3.4), we have for any $k \in \mathscr{H}_H$,

$$\|\mathrm{d}\Phi(\gamma)k\|_\infty \leq C\|k\|_\infty \tag{3.6}$$

for some positive constant C. Therefore, $\mathrm{d}\Phi(\gamma)$ can be extended continuously to an operator on $P(\mathbb{R}^d)$. We have

$$g_1 = \mathrm{d}\Phi(\gamma)B + Y.$$

On the other hand, since γ is a critical point of $F \circ \Phi + 1/2\| \cdot \|^2_{\mathscr{H}_H}$ and noting that $\|k\|_{\mathscr{H}_H} = \|K_H^{-1}k\|_{\mathscr{H}}$, we have

$$\begin{aligned} \mathrm{d}F(\phi)(\mathrm{d}\Phi(\gamma)k) &= -\int_0^T (\dot{K_H^{-1}}\gamma)_s (\dot{K_H^{-1}}k)_s \mathrm{d}s \\ &= -\int_0^T \left((K_H^*)^{-1}(\dot{K_H^{-1}}\gamma)\right)_s \mathrm{d}k_s \end{aligned} \tag{3.7}$$

for all $k \in \mathscr{H}_H$. The second equation above can be seen as follows. Define

$$h = K_H^{-1}k.$$

We have

$$\begin{aligned} \int_0^T \left((K_H^*)^{-1}(\dot{K_H^{-1}}\gamma)\right)_s \mathrm{d}k_s &= \int_0^T \left((K_H^*)^{-1}(\dot{K_H^{-1}}\gamma)\right)_s \int_0^s \frac{\partial K_H}{\partial s}(s,u)\dot{h}_u \mathrm{d}u\mathrm{d}s \\ &= \int_0^T \dot{h}_u \int_u^T \left((K_H^*)^{-1}(\dot{K_H^{-1}}\gamma)\right)_s \frac{\partial K_H}{\partial s}(s,u)\mathrm{d}s \\ &= \int_0^T \dot{h}_u (\dot{K_H^{-1}}\gamma)_u \mathrm{d}u \\ &= \int_0^T (\dot{K_H^{-1}}\gamma)_s (\dot{K_H^{-1}}k)_s \mathrm{d}s. \end{aligned}$$

From (3.6) and (3.7) we conclude that the path $(K_H^*)^{-1}(\dot{K_H^{-1}}\gamma)$ has bounded variation and hence, by passing to limit, we obtain

$$\mathrm{d}F(\phi)(\mathrm{d}\Phi(\gamma)B) = -\int_0^T \left((K_H^*)^{-1}(\dot{K_H^{-1}}\gamma)\right)_s \mathrm{d}B_s.$$

The proof is completed. □

Now, by Theorem 2.4 we have

$$\begin{aligned} J_\rho(\varepsilon) &= \mathbb{E}\Bigg[f(Z^\varepsilon)\exp\left(-\frac{F(Z^\varepsilon)}{\varepsilon^2}\right) \\ &\quad \times \exp\left(-\frac{1}{\varepsilon}\int_0^T \left((K_H^*)^{-1}(\dot{K_H^{-1}}\gamma)\right)_s \mathrm{d}B_s - \frac{\|\gamma\|^2_{\mathscr{H}_H}}{2\varepsilon^2}\right); Z^\varepsilon \in B(\phi,\rho)\Bigg] \\ &= \mathbb{E}\big[V(\varepsilon); Z^\varepsilon \in B(\phi,\rho)\big] \exp\left[-\frac{1}{\varepsilon^2}\left(F(\phi) + \frac{1}{2}\|\gamma\|^2_{\mathscr{H}_H}\right)\right] \exp\left[-\frac{\mathrm{d}F(\phi)Y}{\varepsilon}\right] \\ &= \mathbb{E}\big[V(\varepsilon); Z^\varepsilon B(\phi,\rho)\big] \exp\left[-\frac{a}{\varepsilon^2}\right] \exp\left[-\frac{\mathrm{d}F(\phi)Y}{\varepsilon}\right]. \end{aligned}$$

In the above

$$V(\varepsilon) = f(Z^\varepsilon)\mathrm{e}^{-U(\varepsilon)}.$$

To prove the Laplace approximation, it now suffices to estimate $\mathbb{E}\big[V(\varepsilon); Z^\varepsilon \in B(\phi,\rho)\big]$. For this purpose, we need the following two technical lemmas.

Lemma 3.9. *Let*

$$\theta(\varepsilon) = F(Z^\varepsilon) = \theta(0) + \varepsilon\theta'(0) + \varepsilon^2 U(\varepsilon)$$

where

$$U(\varepsilon) = \int_0^1 (1-v)\theta''(\varepsilon v)\mathrm{d}v, \quad \text{and} \quad \theta(0) = F(\phi).$$

There exist $\beta > 0$ *and* $\varepsilon_0 > 0$ *such that*

$$\sup_{0\le\varepsilon\le\varepsilon_0} \mathbb{E}\left(\mathrm{e}^{-(1+\beta)U(\varepsilon)}; t \le T^\varepsilon\right) < \infty.$$

Proof. See the Appendix. □

Lemma 3.10. *For all* $m > 0$ *and* $p \ge 2$, *there exists an* $\varepsilon_0 > 0$ *such that*

$$\sup_{\varepsilon\le\varepsilon_0} \mathbb{E}\left(\sup_{t\in[0,1]} |\partial_\varepsilon^m Z_t^\varepsilon|^p\right) < \infty.$$

Proof. This is a consequence of Lemma 2.2. □

Define $V^{(m)}(\varepsilon) = \partial_\varepsilon^m V(\varepsilon)$. By Lemmas 3.9 and 3.10, one can show that

$$\mathbb{E}|V^{(m)}(0)|^p < \infty, \quad \text{for all } p > 1, \ m > 0.$$

Consider the stochastic Taylor expansion for $V(\varepsilon)$

$$V(\varepsilon) = \sum_{m=0}^N \frac{\varepsilon^m V^{(m)}(0)}{m!} + \varepsilon^{N+1} S_{N+1}^\varepsilon$$

where

$$S_{N+1}^\varepsilon = \int_0^1 \frac{V^{(N+1)}(\varepsilon v)(1-v)^N}{N!}\mathrm{d}v.$$

It can be shown, again by Lemmas 3.9 and 3.10 (cf. [7]), that

$$\sup_{0\le\varepsilon\le\varepsilon_0} \mathbb{E}\left[|S_{N+1}^\varepsilon|; Z^\varepsilon \in B(\phi,\rho)\right] < \infty.$$

Thus we conclude that

$$\mathbb{E}\left[V(\varepsilon); Z^\varepsilon \in B(\phi,\rho)\right] = \sum_{m=0}^N \alpha_m \varepsilon^m + O(\varepsilon^{N+1}).$$

Moreover, one can show that

$$\alpha_m = \frac{\mathbb{E}V^{(m)}(0)}{m!}.$$

4. Short-time expansion for the transition density

We now arrive at the heart of our study and are interested in obtaining a short-time expansion for the density function of X_t, where

$$\mathrm{d}X_t = \sum_{i=1}^{d} V_i(X_t)\mathrm{d}B_t^i, \quad X_0 = x. \tag{4.1}$$

Here the V_i's are C^∞ vector fields on $\mathbb{R}^d$ with bounded derivatives to any order. Recall Assumption 1.1; our main result of this paper is the following.

Theorem 4.1. *Fix $x \in \mathbb{R}^d$. Assume that Assumption 1.1 is satisfied; then in a neighborhood V of x, the density function $p(t; x, y)$ of X_t in (4.1) has the following asymptotic expansion near $t = 0$:*

$$p(t; x, y) = \frac{1}{(t^H)^d} \mathrm{e}^{-\frac{d^2(x,y)}{2t^{2H}}} \left(\sum_{i=0}^{N} c_i(x, y)t^{2iH} + r_{N+1}(t, x, y)t^{2nH} \right), \quad y \in V.$$

Here $d(x, y)$ is the Riemannian distance between x and y determined by $V_1, \ldots, V_d$. Moreover, we can choose V such that $c_i(x, y)$ are C^∞ in $V \times V \subset \mathbb{R}^d \times \mathbb{R}^d$, and for all multi-indices α and β,

$$\sup_{t \le t_0} \sup_{(x,y) \in V \times V} |\partial_x^\alpha \partial_y^\beta \partial_t^k r_{N+1}(t, x, y)| < \infty$$

for some $t_0 > 0$.

Once the Laplace approximation in the previous section is obtained, the proof of the above theorem is actually quite standard and follows closely the argument given in, for instance, [8]. Thus, for most of the lemmas in what follows, we only outline the proofs, but stress the main differences from the Brownian motion case.

4.1. Preliminaries in differential geometry

The vector fields $V_1, V_2, \ldots, V_d$ on $\mathbb{R}^d$ determine a natural Riemannian metric $g = (g_{ij})$ on $\mathbb{R}^d$ under which $V_1(x), V_2(x), \ldots, V_d(x)$ form an orthonormal frame at each point $x \in \mathbb{R}^d$. More explicitly, let σ be the $d \times d$ matrix formed by

$$\sigma(x) = (V_1(x), V_2(x), \ldots, V_d(x)).$$

Denote by Γ the matrix inverse of $\sigma\sigma^*$. Then the Riemannian metric g is given by

$$g_{ij} = \Gamma_{ij}, \quad 1 \le i,\ j \le d.$$

Throughout our discussion, we denote by M the Riemannian manifold $\mathbb{R}^d$ equipped with the metric g specified above. The Riemannian distance between any two points x, y on M is denoted by $d(x, y)$. We recall that

$$d(x, y) = \inf_{\gamma \in \mathcal{C}(x,y)} \int_0^1 \sqrt{g_{\gamma(s)}(\gamma'(s), \gamma'(s))}\mathrm{d}s$$

where $\gamma \in \mathcal{C}(x, y)$ denotes the set of absolutely continuous curves $\gamma : [0, 1] \to \mathbb{R}^d$ such that $\gamma(0) = x, \gamma(1) = y$.

More analytically, this distance may also be defined as

$$d(x,y)=\sup\left\{f(x)-f(y),\, f\in C_b^\infty(\mathbb{R}^d),\, \sum_{i=1}^d (V_i f)^2\le 1\right\},$$

where $C_b^\infty(\mathbb{R}^d)$ denotes the set of smooth and bounded functions on $\mathbb{R}^d$. Since the vector fields $V_1,\dots,V_d$ are Lipschitz, it is well-known that this distance is complete and that the Hopf–Rinow theorem holds (that is, closed balls are compact).

Due to the second part of Assumption 1.1, the geodesics are easily described. If $k:\mathbb{R}_{\ge 0}\to\mathbb{R}$ is an α-Hölder path with $\alpha>1/2$ such that $k(0)=0$, we denote by $\Phi(x,k)$ the solution of the ordinary differential equation

$$x_t=x+\sum_{i=1}^d\int_0^t V_i(x_s)\mathrm{d}k_s^i.$$

Whenever there is no confusion, we always suppress the starting point x and denote it simply by $\Phi(k)$ as before.

Lemma 4.2. *$\Phi(x,k)$ is a geodesic if and only if $k(t)=tu$ for some $u\in\mathbb{R}^d$.*

Proof. It is well-known that geodesics c are smooth and are solutions of the equation

$$\nabla_{c'}c'=0,$$

where ∇ is the Levi-Civita connection. Therefore, in order for $\Phi(k)$ to be a geodesic, we see first that k needs to be smooth and then that

$$\nabla_{\sum_{i=1}^d V_i(x_s)\dot{k}_s^i}\sum_{i=1}^d V_i(x_s)\dot{k}_s^i=0.$$

Now, due to the structure equations

$$[V_i,V_j]=\sum_{l=1}^d \omega_{ij}^l V_l,$$

the Christoffel symbols of the connection are given by

$$\Gamma_{ij}^l=\frac{1}{2}\left(\omega_{ij}^l+\omega_{li}^j+\omega_{lj}^i\right)=\frac{1}{2}\omega_{ij}^l.$$

So the equation of the geodesics may be rewritten as

$$\sum_{l=1}^d\frac{\mathrm{d}^2k_s^l}{\mathrm{d}s^2}V_l(x_s)+\sum_{i,j,l=1}^d\omega_{ij}^l\dot{k}_s^i\dot{k}_s^j V_l(x_s)=0.$$

Due to the skew symmetry $\omega_{ij}^l=-\omega_{ji}^l$ we get

$$\frac{\mathrm{d}^2k_s^l}{\mathrm{d}s^2}=0,$$

which leads to the expected result. □

As a consequence of the previous lemma, we then have the following key result:

Proposition 4.3. *Let $T > 0$. For $x, y \in \mathbb{R}^d$,*

$$\inf_{k \in \mathscr{H}_H, \Phi_T(x,k)=y} \|k\|^2_{\mathscr{H}_H} = \frac{d^2(x,y)}{T^{2H}}.$$

Proof. In a first step we prove

$$\frac{d^2(x,y)}{T^{2H}} \le \inf_{k \in \mathscr{H}_H, \Phi_T(x,k)=y} \|k\|^2_{\mathscr{H}_H}.$$

Let $k \in \mathscr{H}_H$ such that $\Phi_0(k) = x$, $\Phi_T(k) = y$. Denote by z the solution of the equation

$$\mathrm{d}z_t = \sum_{i=1}^d V_i(z_t)\mathrm{d}k^i_t, \quad 0 \le t \le T.$$

We have therefore

$$z_0 = x, \quad z_T = y.$$

Let now $f \in C_b^\infty(\mathbb{R}^d)$ such that $\sum_{i=1}^d (V_i f)^2 \le 1$. By the change of variable formula, we get

$$f(y) - f(x) = \sum_{i=1}^d \int_0^T V_i f(z_t)\mathrm{d}k^i_t.$$

Since $k \in \mathscr{H}_H$, we can find h in the Cameron–Martin space of the Brownian motion such that

$$k_t = \int_0^t K_H(t,s)\dot{h}_s \mathrm{d}s.$$

Integrating by parts, we have then

$$\int_0^T V_i f(z_t)\mathrm{d}k^i_t = \int_0^T \left(\int_s^T \frac{\partial K_H}{\partial t}(t,s) V_i f(z_t)\mathrm{d}t \right) \dot{h}^i_s \mathrm{d}s.$$

Therefore from the Cauchy–Schwarz inequality, the isometry between $\mathscr{H}$ and $\mathscr{H}_H$ and the fact that $\sum_{i=1}^d (V_i f)^2 \le 1$, we deduce that

$$(f(y) - f(x))^2 \le R(T,T)\|\dot{h}\|^2_{L^2([0,1])} = T^{2H}\|k\|^2_{\mathscr{H}_H}.$$

Thus

$$\frac{d^2(x,y)}{T^{2H}} \le \inf_{k \in \mathscr{H}_H, \Phi_T(x,k)=y} \|k\|^2_{\mathscr{H}_H}.$$

We now prove the converse inequality.

We first assume that y is close enough to x that there exist $(y_1, \ldots, y_d) \in \mathbb{R}^d$ that satisfy

$$y = \exp\left(\sum_{i=1}^d y_i V_i \right)(x).$$

Let

$$k^i_t = \frac{\int_0^t K_H(t,s)K_H(T,s)\mathrm{d}s}{T^{2H}} y_i = \frac{R(t,T)}{T^{2H}} y_i.$$

In that case, it is easily seen that

$$\Phi(k)(t) = \exp\left(\sum_{i=1}^{d} \frac{R(t,T)}{T^{2H}} y_i V_i\right)(x).$$

In particular,

$$\Phi_0(k) = x, \qquad \Phi_1(k) = y.$$

Moreover,

$$\|k\|^2_{\mathcal{H}_H} = \frac{\sum_{i=1}^{d} y_i^2}{T^{2H}} = \frac{d^2(x,y)}{T^{2H}}.$$

As a consequence

$$\inf_{k\in\mathcal{H}_H,\,\Phi_T(x,k)=y} \|k\|^2_{\mathcal{H}_H} \le \frac{d^2(x,y)}{T^{2H}}.$$

If y is not close to x, we just have to pick a sequence $x_0 = x, \dots, x_m = y$ such that

$$d(x_i, x_{i+1}) \le \varepsilon$$

and

$$d(x,y) = \sum_{i=0}^{m-1} d(x_i, x_{i+1}),$$

where ε is small enough. $\square$

The second key point is the following:

Theorem 4.4. *Fix $x_0 \in M$. Let F be a C^∞ function on M. There exists a neighborhood V of x_0 such that if $y_0 \in V$ is a non-degenerate minimum of*

$$F(y) + \frac{d^2(x_0,y)}{2},$$

then there exists a unique $k^0 \in \mathcal{H}_H$ such that (a) $\Phi_1(x_0,k^0) = y_0$, (b) $d(x_0,y_0) = \|k^0\|_{\mathcal{H}_H}$, *and* (c) k^0 *is a non-degenerate minimum of the functional: $k \to F(\Phi_1(x_0,k)) + 1/2\|k\|^2_{\mathcal{H}_H}$ on $\mathcal{H}_H$.*

Proof. The first two statements are clear from Proposition 4.3. We only need to prove (c). To simplify notation, let

$$G(k) = F(\Phi_1(x_0,k)) + \frac{1}{2}\|k\|^2_{\mathcal{H}_H}.$$

Consider

$$w(u) = G(k^0 + uk),$$

and

$$v(u) = F(\Phi_1(x_0, k^0 + uk)) + \frac{1}{2} d^2(x_0, \Phi_1(x_0, k^0 + uk)).$$

It is clear that

$$w(u) \geq v(u), \qquad w(0) = v(0) \quad \text{and} \quad w'(0) = v'(0) = 0.$$

Thus

$$\mathrm{d}^2 G(k_0)(k,k) = w''(0) \geq v''(0) = \left(F + \frac{1}{2} d(x_0,\cdot)^2\right)''(y_0)\left(\mathrm{d}\Phi_1(k^0)k\right)^2.$$

When $k \notin \mathrm{Ker}(\mathrm{d}\Phi_1(x_0,k^0))$, we certainly have

$$\mathrm{d}^2 G(k^0)(k,k) > 0.$$

In the case where $k \in \mathrm{Ker}(\mathrm{d}\Phi_1(x_0,k^0))$, it is clear that we can assume that $\|k\|_{\mathscr{H}_H} = 1$. From Proposition 4.3 we first note that

$$k_t^0 = R(t,1) y_0, \quad t \in [0,1]. \tag{4.2}$$

Since $k \in \mathrm{Ker}(\mathrm{d}\Phi_1(x_0,k^0))$, we can consider a variation k^u of k^0 such that $\Phi_1(x_0,k^u) = y_0$, $u \in (-\varepsilon,\varepsilon)$. Without loss of generality, we may assume that k^u takes the form

$$k^u = k^0 + uk + u^2 k^1$$

for some $k^1 \in \mathscr{H}_H$. In what follows, we show that there exists a neighborhood V of x_0 such that for all $y_0 \in V$, we have

$$\left.\frac{\mathrm{d}^2}{\mathrm{d}u^2}\right|_{u=0} \|k^u\|^2_{\mathscr{H}_H} > 0.$$

Indeed,

$$\begin{aligned}\left.\frac{\mathrm{d}^2}{\mathrm{d}u^2}\right|_{u=0} \|k^u\|^2_{\mathscr{H}_H} &= 2\|k\|^2_{\mathscr{H}_H} + 4\langle k^0, k^1\rangle_{\mathscr{H}_H} \\ &= 2\|k\|^2_{\mathscr{H}_H} + 4k_1^1 \cdot y_0.\end{aligned}$$

For the second equation above, we used (4.2). On the other hand, since $\Phi_1(x_0,k^u) = y_0$ for all $u \in (-\varepsilon,\varepsilon)$ we have

$$\mathrm{d}\Phi_1(k^0)k = 0, \quad \text{and} \quad \mathrm{d}^2\Phi_1(k^0)(k,k) + 2\mathrm{d}\Phi_1(k^0)(k^1) = 0. \tag{4.3}$$

From Lemma 3.5 we have

$$|\mathrm{d}^2\Phi_1(k^0)(k,k)| \leq M\|k\|^2_{\mathscr{H}_H} \tag{4.4}$$

for some constant $M > 0$. By computation, we know that

$$\mathrm{d}\chi_t = \sigma(\phi_t)\mathrm{d}k_t^1 + \partial_x\sigma(\phi_t)\chi_t \mathrm{d}k_t^0 \tag{4.5}$$

where $\chi = \mathrm{d}\Phi(k^0)(k^1)$ and $\phi = \Phi(k^0)$. In particular, when $y_0 = x_0$ the above equation becomes (since $k^0 \equiv 0$ in this case)

$$\mathrm{d}\chi_t = \sigma(x_0)\mathrm{d}k_t^1.$$

Therefore when $x_0 = y_0$

$$|\chi_1| \geq C_1|k_1^1|, \quad \text{for all } k^1 \in \mathscr{H}_H$$

for some constant $C_1 > 0$. It is clear that, by Eq. (4.5) and a continuity argument, we can find a neighborhood V of x_0 such that for all $y_0 \in V$ we have

$$|\chi_1| \geq C|k_1^1| - \delta \quad \text{for all } k^1 \in \mathscr{H}_H \tag{4.6}$$

for some $0 < \delta < 1$ and some constant $C > 0$ depending on V. Now by the second equation in (4.3), and inequalities (4.4) and (4.6) we have

$$2(C|k_1^1| - \delta) \leq 2|\chi_1| = |\mathrm{d}^2 \Phi_1(k^0)(k,k)| \leq M\|k\|^2_{\mathscr{H}_H}. \tag{4.7}$$

Hence in V we have

$$\begin{aligned} \left.\frac{\mathrm{d}^2}{\mathrm{d}u^2}\right|_{u=0} \|k^u\|^2_{\mathscr{H}_H} &= 2\|k\|^2_{\mathscr{H}_H} + 4k_1^1 \cdot y_0 \\ &\geq 2\|k\|^2_{\mathscr{H}_H} - 4|k_1^1|\,|y_0| \\ &\geq 2\|k\|^2_{\mathscr{H}_H} - \frac{2|y_0|M}{C}\|k\|^2_{\mathscr{H}_H} - \frac{4|y_0|\delta}{C}. \end{aligned}$$

Now we only need to choose V even smaller (so that $|y_0|$ is small) to guarantee that the above is non-negative. □

Remark 4.5. In the above lemma, it is clear that we can choose the neighborhood V of x_0 such that for any $x \in V$, if $y \in V$ is a non-degenerate minimum of $F(y) + d(x,y)^2/2$, then the three properties in the lemma are fulfilled.

4.2. Asymptotics of the density function

Consider

$$\mathrm{d}X_t^\varepsilon = \varepsilon \sum_{i=1}^d V_i(X_t^\varepsilon)\mathrm{d}B_t^i \quad \text{with } X_0^\varepsilon = x.$$

Before applying the Laplace approximation to X_t^ε, we need the following lemma which gives us the correct functionals F and f.

Lemma 4.6. *Let V be as in Remark 4.5. There exists a bounded smooth function $F(x,y,z)$ on $V \times V \times M$ such that:*

(1) *For any $(x,y) \in V \times V$ the infimum*

$$\inf\left\{F(x,y,z) + \frac{d(x,z)^2}{2}, z \in M\right\} = 0$$

is attained at the unique point y. Moreover, it is a non-degenerate minimum.

(2) *For each $(x,y) \in V \times V$, there exists a ball centered at y with radius r independent of x, y such that $F(x,y,\cdot)$ is a constant outside of the ball.*

Proof. See Lemma 3.8 in [8]. □

Let F be as in the above lemma and $p_\varepsilon(x,y)$ the density function of X_1^ε. By the inversion of the Fourier transformation we have

$$p_\varepsilon(x,y)\mathrm{e}^{-\frac{F(x,y,y)}{\varepsilon^2}} = \frac{1}{(2\pi)^d}\int \mathrm{e}^{-\mathrm{i}\zeta\cdot y}\mathrm{d}\zeta \int \mathrm{e}^{\mathrm{i}\zeta\cdot z}\mathrm{e}^{-\frac{F(x,y,z)}{\varepsilon^2}} p_\varepsilon(x,z)\mathrm{d}z$$

$$= \frac{1}{(2\pi\varepsilon)^d} \int e^{-i\frac{\zeta \cdot y}{\varepsilon}} d\zeta \int e^{i\frac{\zeta \cdot z}{\varepsilon}} e^{-\frac{F(x,y,z)}{\varepsilon^2}} p_\varepsilon(x,z) dz$$
$$= \frac{1}{(2\pi\varepsilon)^d} \int d\zeta \, \mathbb{E}_x \left(e^{\frac{i\zeta\cdot(X_1^\varepsilon - y)}{\varepsilon}} e^{-\frac{F(x,y,X_1^\varepsilon)}{\varepsilon^2}} \right).$$

It is clear that by applying Laplace approximation to the expectation in the last equation above and switching the order of integration (with respect to ζ) and summation, we obtain an asymptotic expansion for the density function $p_\varepsilon(x, y)$. On the other hand, we cannot apply the Laplace method here directly since we need a uniform control in x and y. Also we need to show that the use of Fourier inversion is legitimate.

To make the above prior computation rigorous, we modify the Laplace method in the previous section as follows.

First note that by Theorem 4.4, Assumption 3.1 is satisfied. Consider

$$dZ_t^\varepsilon(x, y) = \sum_{i=1}^d V_i\big(Z_t^\varepsilon(x, y)\big)\big(\varepsilon dB_t^i + d\gamma_t^i(x, y)\big), \quad \text{with } Z_0^\varepsilon(x, y) = x.$$

In the above $(x, y) \in V \times V$ and $\gamma(x, y)$ is the unique path in $\mathscr{H}_H$ such that $\Phi_1(x, \gamma(x, y)) = y$ and $\|\gamma(x, y)\|_{\mathscr{H}_H} = d(x, y)$.

Lemma 4.7. *Let $Z_t^\varepsilon(x, y)$ be the process defined above; then $Z_t^\varepsilon(x, y)$ is C^∞ in (ε, x, y). Moreover, there exists an $\varepsilon_0 > 0$ such that*

$$\sup_{\varepsilon \le \varepsilon_0} \sup_{x,y \in V \times V} \sum_{j=0}^n \mathbb{E}\left(\sup_{t\in[0,1]} \|D^j(\partial_x^\alpha \partial_y^\beta \partial_\varepsilon^m Z_t^\varepsilon(x, y))\|_{HS}^p \right) < \infty.$$

Here m, n are non-negative integers, $p \ge 2$ and $\alpha \in \{1, 2, \ldots, d\}^k$, $\beta \in \{1, 2, \ldots, d\}^l$ are multiple indices.

Proof. The first statement is clear. The second statement is a consequence of Lemma 2.2. □

Now consider the stochastic Taylor expansion for Z^ε:

$$Z_t^\varepsilon = \phi_t(x, y) + \sum_{j=1}^N \frac{g_t^k(x, y)\varepsilon^k}{k!} + R_t^{N+1}(\varepsilon, x, y)\varepsilon^{N+1}. \tag{4.8}$$

Here

$$\phi(x, y) = \Phi(x, \gamma(x, y)),$$

and

$$R_t^{N+1}(\varepsilon, x, y) = \int_0^1 \partial_\varepsilon^{N+1} Z_t^\varepsilon(x, y) \frac{(1 - v)^N}{N!} dv.$$

Let

$$\theta(\varepsilon, x, y) = F(x, y, Z_1^\varepsilon(x, y)).$$

We have

$$\theta(\varepsilon, x, y) = \theta(0, x, y) + \varepsilon\partial_\varepsilon\theta(0, x, y) + \varepsilon^2 U(\varepsilon, x, y)$$

where

$$U(\varepsilon, x, y) = \int_0^1 \partial_\varepsilon^2 \theta(\varepsilon, x, y)(1 - v)\mathrm{d}v.$$

By our choice of Z^ε, it is clear that

$$\theta(0, x, y) = F(x, y, \phi_1(x, y)) = F(x, y, y). \tag{4.9}$$

Lemma 3.8 gives us

$$\partial_\varepsilon \theta(0, x, y) = -\int_0^1 (K_H^*)^{-1}(K_H^{-1}\dot{\gamma}(x, y))_s \mathrm{d}B_s. \tag{4.10}$$

Thus applying the Cameron–Martin theorem for fBm (Theorem 2.4), we have

$$\begin{aligned}
&\mathbb{E}_x \exp\left(\frac{\mathrm{i}\zeta \cdot (X_1^\varepsilon - y)}{\varepsilon} - \frac{F(x, y, X_1^\varepsilon)}{\varepsilon^2}\right) \\
&\quad = \mathbb{E}\left[\exp\left(\frac{\mathrm{i}\zeta \cdot (Z_1^\varepsilon - y)}{\varepsilon} - \frac{F(x, y, Z_1^\varepsilon)}{\varepsilon^2}\right)\right. \\
&\qquad \left. \times \exp\left(-\frac{1}{\varepsilon}\int_0^1 \left((K_H^*)^{-1}(K_H^{-1}\dot{\gamma})\right)_s \mathrm{d}B_s - \frac{\|\gamma\|_{\mathscr{H}_H}^2}{2\varepsilon^2}\right)\right] \\
&\quad = \exp\left[-\frac{a}{\varepsilon^2}\right]\mathbb{E}_x\left[\exp\Big(\mathrm{i}\zeta \cdot g_1^1(x, y)\Big)\exp\Big(\mathrm{i}\zeta \cdot V(\varepsilon, x, y) - U(\varepsilon, x, y)\Big)\right].
\end{aligned}$$

In the above

$$a(x, y) = F(x, y, y) + \frac{d^2(x, y)}{2} = 0,$$

and

$$V(\varepsilon, x, y) = \frac{Z_1^\varepsilon(x, y) - y - \varepsilon g_1^1(x, y)}{\varepsilon} = \varepsilon R_1^2(\varepsilon, x, y).$$

Like in the argument in Section 2, we need to estimate

$$\mathbb{E}_x\left[\exp\Big(\mathrm{i}\zeta \cdot g_1^1(x, y)\Big)\exp\Big(\mathrm{i}\zeta \cdot V(\varepsilon, x, y) - U(\varepsilon, x, y)\Big)\right].$$

For this purpose, we need:

Lemma 4.8. *There exist $C > 0$ and $\varepsilon_0 > 0$ such that*

$$\sup_{(x,y)\in V\times V}\ \sup_{\varepsilon<\varepsilon_0} \mathbb{E}\mathrm{e}^{-(1+C)U(\varepsilon,x,y)} < \infty.$$

Proof. We only sketch the proof. Details can be found in [8] (with minor modifications) and will not be repeated here.

Fix any $1/2 < \lambda < H$. One can show that for $\rho > 0$ there exist constants $C > 0$, $b > 0$ and $\varepsilon_0 > 0$ such that for all $\varepsilon < \varepsilon_0$ and all $(x, y) \in V \times V$ we have

$$\mathbb{E}_x\left\{\mathrm{e}^{-(1+C)U(\varepsilon,x,y)};\ \left\|Z_t^\varepsilon - \phi_t(x, y)\right\|_{\lambda,1} \geq \rho\right\} \leq \mathrm{e}^{\frac{-b}{\varepsilon^2}}. \tag{4.11}$$

Here $\|\cdot\|_{\lambda,t}$ is the λ-Hölder norm up to time t. The above estimate is a consequence of the following application of the large deviation principle to X_1^ε:

$$\limsup_{\varepsilon\to 0} \varepsilon^2 \log \mathbb{E}_x \left\{ e^{-\frac{F(x,y,X_1^\varepsilon)}{\varepsilon^2}}; \|X^\varepsilon - \phi(x,y)\|_{\lambda,1} \geq \rho \right\} < -a(x,y) = 0.$$

On the other hand, applying Lemma 3.9 we have that for each $(x, y) \in V \times V$ there exists $C > 0$ and $\varepsilon_0 > 0$ such that

$$\sup_{\varepsilon<\varepsilon_0} \mathbb{E}_x \left\{ e^{-(1+C)U(\varepsilon,x,y)}; \|Z^\varepsilon - \phi(x,y)\|_{\lambda,1} \leq \rho \right\} < \infty.$$

Since we have smoothness of $Z^\varepsilon(x, y)$ (in x and y) and $V \times V$ is contained in a compact subset of $M \times M$, the above estimate leads to

$$\sup_{\varepsilon<\varepsilon_0} \sup_{(x,y)\in V\times V} \mathbb{E}_x \left\{ e^{-(1+C)U(\varepsilon,x,y)}; \|Z^\varepsilon - \phi(x,y)\|_{\lambda,1} \leq \rho \right\} < \infty.$$

Combining this with (4.11), the proof is completed. □

Set

$$\Upsilon(\varepsilon, x, y) = e^{i\zeta\cdot V(\varepsilon,x,y)-U(\varepsilon,x,y)}$$

and consider its stochastic Taylor expansion:

$$\Upsilon(\varepsilon, x, y, \zeta) = \sum_{m=0}^{N} \partial_\varepsilon^m \Upsilon(0, x, y, \zeta) \frac{\varepsilon^m}{m!} + S_{N+1}(\varepsilon, x, y, \zeta)\varepsilon^{N+1}, \tag{4.12}$$

where

$$S_{N+1}(\varepsilon, x, y, \zeta) = \int_0^1 \partial_\varepsilon^{N+1} \Upsilon(\varepsilon v, x, y, \zeta) \frac{(1-v)^N}{N!} dv.$$

Lemma 4.9. *For any non-negative integers k, l, m and n, and multi-indices $\alpha \in \{1, 2, \ldots, d\}^k$ and $\beta \in \{1, 2, \ldots, d\}^l$, we have:*

(1) *For all $p \geq 2$, there exists $\varepsilon_0 > 0$ such that*

$$\sup_{\varepsilon\leq\varepsilon_0} \sup_{x,y\in V\times V} \mathbb{E}\left(\sum_{j=0}^{n} \sup_{t\in[0,1]} \|D^j(\partial_x^\alpha \partial_y^\beta \partial_\varepsilon^m i\zeta \cdot V(\varepsilon,x,y)) - U(\varepsilon,x,y)\|_{HS}^p \right) < \infty.$$

(2) *There exist $C > 0$, $K > 0$ and $\varepsilon_0 > 0$ such that*

$$\sup_{\varepsilon\leq\varepsilon_0} \sup_{x,y\in V\times V} \mathbb{E}\left(\sum_{j=0}^{n} \sup_{t\in[0,1]} \|D^j(\partial_x^\alpha \partial_y^\beta \partial_\varepsilon^m \Upsilon(\varepsilon,x,y,\zeta))\|_{HS}^{1+C} \right) < K\left(\|\zeta\|+1\right)^{m+k+l}.$$

Moreover, we have

$$\sup_{\varepsilon\leq\varepsilon_0} \sup_{x,y\in V\times V} \mathbb{E}\left(\sum_{j=0}^{n} \sup_{t\in[0,1]} \left\| D^j\left(\partial_x^\alpha \partial_y^\beta \partial_\varepsilon^m \left(e^{i\zeta\cdot g_1^1(x,y)} \Upsilon(\varepsilon,x,y,\zeta)\right)\right) \right\|_{HS}^{1+C} \right)$$
$$< K\left(\|\zeta\|+1\right)^{m+k+l}.$$

Proof. We follow the argument in [8]. Note that

$$\mathrm{i}\zeta \cdot V(\varepsilon, x, y) - U(\varepsilon, x, y) = \mathrm{i}\zeta \int_0^1 \partial_\varepsilon^2 Z_1^{\varepsilon v}(x, y)(1-v)\mathrm{d}v - \int_0^1 \partial_\varepsilon^2 \theta(\varepsilon v, x, y)(1-v)\mathrm{d}v.$$

The estimate in (1) follows directly from Lemma 4.7.

For the second statement, first note that

$$\mathrm{e}^{-U} \in \mathrm{Dom}(D).$$

This is seen by an approximating argument and because D is a closed operator. Moreover, we have

$$D(\mathrm{e}^{-U}) = -(DU)\mathrm{e}^{-U}.$$

Hence Υ is also in the domain of D.

It is clear that $\partial_x^\alpha \partial_y^\beta \partial_\varepsilon^m \Upsilon$ is of the form $W\Upsilon$, where W is a polynomial in ζ of degree $m + |\alpha| + |\beta|$ with, as coefficients, derivatives (w.r.t. x, y and ε) of $U(\varepsilon, x, y)$ and $V(\varepsilon, x, y)$. Moreover,

$$D(\partial_x^\alpha \partial_y^\beta \partial_\varepsilon^m \Upsilon) = (DW + \mathrm{i}\zeta \cdot DV - DU)\Upsilon.$$

The first estimate in (2) now follows immediately from (1) and Lemma 4.8. The last estimate in (2) then follows from the first one in (2) and Lemma 4.7. This completes the proof. □

With the above lemma, we are now able to obtain an asymptotic expansion for

$$\mathbb{E}_x \left[\exp\left(\mathrm{i}\zeta \cdot g_1^1(x, y) \right) \exp\left(\mathrm{i}\zeta \cdot V(\varepsilon, x, y) - U(\varepsilon, x, y) \right) \right].$$

Define

$$\alpha_m(x, y, \zeta) = \mathbb{E}_x \left[\exp\big(\mathrm{i}\zeta \cdot g_1^1(x, y)\big) \partial_\varepsilon^m \Upsilon(0, x, y, \zeta) \right],$$

and

$$T_{N+1}(\varepsilon, x, y, \zeta) = \mathbb{E}_x \left[\exp\big(\mathrm{i}\zeta \cdot g_1^1(x, y)\big) S_{N+1}(\varepsilon, x, y, \zeta) \right].$$

Recall (4.12); we obtain

$$\begin{aligned} &\mathbb{E}_x \left[\exp\left(\mathrm{i}\zeta \cdot g_1^1(x, y) \right) \exp\left(\mathrm{i}\zeta \cdot V(\varepsilon, x, y) - U(\varepsilon, x, y) \right) \right] \\ &\quad = \mathbb{E}_x \left[\exp\left(\mathrm{i}\zeta \cdot g_1^1(x, y) \right) \Upsilon(\varepsilon, x, y, \zeta) \right] \\ &\quad = \sum_{m=0}^{N} \alpha_m(x, y, \zeta)\varepsilon^m + T_{N+1}(\varepsilon, x, y, \zeta)\varepsilon^{N+1}. \end{aligned}$$

Remark 4.10. Lemma 4.9 in fact provides the smoothness and boundedness of α_m and T_{N+1}.

So far, we have obtained that for all $\zeta \in \mathbb{R}^d$

$$\mathbb{E}_x \exp\left(\frac{\mathrm{i}\zeta \cdot (X_1^\varepsilon - y)}{\varepsilon} - \frac{F(x, y, X_1^\varepsilon)}{\varepsilon^2} \right)$$

$$= \mathrm{e}^{-\frac{a(x,y)}{\varepsilon^2}} \left(\sum_{m=0}^{N} \alpha_m(x,y,\zeta)\varepsilon^m + T_{N+1}(\varepsilon,x,y,\zeta)\varepsilon^{N+1} \right)$$

$$= \sum_{m=0}^{N} \alpha_m(x,y,\zeta)\varepsilon^m + T_{N+1}(\varepsilon,x,y,\zeta)\varepsilon^{N+1}.$$

To apply the inversion of Fourier transformation, we need integrability of α_m and T_{N+1} in ζ, which is provided by the following lemma.

Lemma 4.11. *For any non-negative integers p, k and l, and multi-indices $\alpha \in \{1,2,\dots,d\}^k$ and $\beta \in \{1,2,\dots,d\}^l$, we have:*

(1) *There exists $K = K_p(\alpha,\beta) > 0$ such that*

$$\sup_{(x,y)\in V\times V} \left| \partial_x^\alpha \partial_y^\beta \alpha_m(x,y,\zeta) \right| \le \frac{K}{\|\zeta\|^{2p}} (\|\zeta\|+1)^{m+k+l}.$$

(2) *There exist $\varepsilon_0 > 0$ and $K = K(p,N,\alpha,\beta,m) > 0$ such that*

$$\sup_{\varepsilon<\varepsilon_0} \sup_{(x,y)\in V\times V} \left| \partial_x^\alpha \partial_y^\beta \partial_\varepsilon^m T_{N+1}(\varepsilon,x,y,\zeta) \right| \le \frac{K}{\|\zeta\|^{2p}} (\|\zeta\|+1)^{(N+1)+k+l}.$$

Proof. The lemma follows from integration by parts in Malliavin calculus. Indeed, first note that by Eq. (A.7), the Malliavin matrix of g^1 is deterministic, non-degenerate and uniform in x and y. By Propositions 5.7 and 5.8 in [22] and Lemma 4.7, for any proper test function ψ, $G \in \mathbb{D}^{|\alpha|,q}$, there exist $l_\alpha G$ and $r < q$ such that

$$\mathbb{E}\big(\partial^\alpha \psi(g_1^1)G\big) = E\big(\psi(g_1^1) l_\alpha(G)\big)$$

and

$$\big(\mathbb{E}|l_\alpha(G)|^r\big)^{\frac{1}{r}} \le K \left(\sum_{j=0}^{|\alpha|} \mathbb{E}\|D^j G\|_{\mathrm{HS}}^q \right)^{\frac{1}{q}}.$$

Here K depends on$|\alpha|$, g_1^1 and its Malliavin matrix and K is uniform in x and y.

We apply the above integration by parts formula with

$$\psi(u) = \mathrm{e}^{\mathrm{i}\zeta\cdot u} \quad \text{and} \quad \partial^\alpha = \left(\sum_{i=1}^{d} \partial_{u_i}^2 \right)^p.$$

We have

$$\left| \mathbb{E}\big(\mathrm{e}^{\mathrm{i}\zeta\cdot g_1^1} G\big) \right| \le \frac{K}{\|\zeta\|^{2p}} \left(\sum_{j=0}^{2p} \mathbb{E}\big(\|D^j G\|_{\mathrm{HS}}^q\big) \right)^{\frac{1}{q}}.$$

Now the lemma follows by Lemma 4.9 and replacing G in the above by

$$G_1 = \partial_x^\alpha \partial_y^\beta \partial_\varepsilon^m \Upsilon(0,x,y,\zeta),$$

and

$$G_2 = \partial_x^\alpha \partial_y^\beta \partial_\varepsilon^m \big(S_{N+1}(\varepsilon,x,y,\zeta)\mathrm{e}^{\mathrm{i}\zeta\cdot g_1^1}\big)\mathrm{e}^{-\mathrm{i}\zeta\cdot g_1^1}. \quad \square$$

Now we only need to chose $2p > d + (N + 1) + k + l$ in the previous lemma and obtain

$$p_\varepsilon(x, y)\mathrm{e}^{-\frac{F(x,y,y)}{\varepsilon^2}} = \frac{\mathrm{e}^{-\frac{a(x,y)}{\varepsilon^2}}}{\varepsilon^d}\left(\sum_{m=0}^{N} \beta_m(x, y)\varepsilon^m + t_{N+1}(\varepsilon, x, y)\varepsilon^{N+1}\right).$$

Here

$$\beta_m(x, y) = \frac{1}{(2\pi)^d}\int \alpha_m(x, y, \zeta)\mathrm{d}\zeta,$$

and

$$t_{N+1}(\varepsilon, x, y) = \frac{1}{(2\pi)^d}\int T_{N+1}(\varepsilon, x, y, \zeta)\mathrm{d}\zeta.$$

Notice that $\beta_m(x, y, \zeta)$ is an odd function in ζ when m is odd (cf. [8]). Now by the self-similarity of the fractional Brownian motion and $\varepsilon = t^H$ we obtain the desired asymptotic formula for the density function.

4.3. *The on-diagonal asymptotics*

As a straightforward corollary of Theorem 4.1, we have the following on-diagonal asymptotics:

$$p(t; x, x) = \frac{1}{t^{Hd}}\left(a_0(x) + a_1(x)t^{2H} + \cdots + a_n(x)t^{2nH} + o(t^{2nH})\right).$$

In this subsection, we analyze the coefficients $a_n(x)$ and show how they are related to some functionals of the underlying fractional Brownian motion.

We first introduce some notation and recall some results that may be found in [3,5,19,11].

If $I = (i_1, \ldots, i_k) \in \{1, \ldots, d\}^k$ is a word, we denote by V_I the Lie commutator defined by

$$V_I = [V_{i_1}, [V_{i_2}, \ldots, [V_{i_{k-1}}, V_{i_k}]] \ldots].$$

The group of permutations of the set $\{1, \ldots, k\}$ is denoted as $\mathfrak{S}_k$. If $\sigma \in \mathfrak{S}_k$, we denote by $e(\sigma)$ the cardinality of the set

$$\{j \in \{1, \ldots, k - 1\}, \sigma(j) > \sigma(j + 1)\}.$$

Finally, for the iterated integrals, defined in Young's sense, we use the following notation:

(1)

$$\Delta^k[0, t] = \{(t_1, \ldots, t_k) \in [0, t]^k, t_1 \leq \cdots \leq t_k\}.$$

(2) If $I = (i_1, \ldots, i_k) \in \{1, \ldots, d\}^k$ is a word with length k,

$$\int_{\Delta^k[0,t]} \mathrm{d}B^I = \int_{0\leq t_1\leq\cdots\leq t_k\leq t} \mathrm{d}B^{i_1}_{t_1} \cdots \mathrm{d}B^{i_k}_{t_k}.$$

(3) If $I = (i_1, \ldots, i_k) \in \{1, \ldots, d\}^k$ is a word with length k,

$$\Lambda_I(B)_t = \sum_{\sigma\in\mathcal{S}_k} \frac{(-1)^{e(\sigma)}}{k^2\binom{k-1}{e(\sigma)}}\int_{0\leq t_1\leq\cdots\leq t_k\leq t} \mathrm{d}B^{\sigma^{-1}(i_1)}_{t_1} \cdots \mathrm{d}B^{\sigma^{-1}(i_k)}_{t_k}, \quad t \geq 0.$$

Theorem 4.12. *For $f \in C_b^\infty(\mathbb{R}^d, \mathbb{R})$, $x \in \mathbb{R}^d$, and $N \geq 0$, when $t \to 0$,*

$$\begin{aligned} f(X_t^x) &= f(x) + \sum_{k=1}^N t^{2kH} \sum_{I=(i_1,\ldots,i_{2k})} (V_{i_1} \ldots V_{i_{2k}} f)(x) \int_{\Delta^{2k}[0,1]} \mathrm{d}B^I + o(t^{(2N+1)H}) \\ &= f\left(\exp\left(\sum_{I,|I|\leq N} \Lambda_I(B)_t V_I\right) x\right) + o(t^{NH}) \end{aligned}$$

and

$$\begin{aligned} \mathbb{E}(f(X_t^x)) &= f(x) + \sum_{k=1}^N t^{2kH} \sum_{I=(i_1,\ldots,i_{2k})} (V_{i_1} \ldots V_{i_{2k}} f)(x) \mathbb{E}\left(\int_{\Delta^{2k}[0,1]} \mathrm{d}B^I\right) \\ &\quad + o(t^{(2N+1)H}) \\ &= \mathbb{E}\left(f\left(\exp\left(\sum_{I,|I|\leq N} \Lambda_I(B)_t V_I\right) x\right)\right) + o(t^{NH}). \end{aligned}$$

As a consequence, we obtain the following proposition which may be proved as in [4] (or [14]).

Proposition 4.13. *For $N \geq 1$, when $t \to 0$,*

$$p(t; x_0, x_0) = d_t^N(x_0) + O\left(t^{H(N+1-d)}\right),$$

where $d_t^N(x_0)$ is the density at 0 of the random variable $\sum_{I,|I|\leq N} \Lambda_I(B)_t V_I(x_0)$.

This proposition may be used to understand the geometric meaning of the coefficients $a_k(x_0)$ of the small-time asymptotics

$$p(t; x, x) = \frac{1}{t^{Hd}} \left(a_0(x) + a_1(x) t^{2H} + \cdots + a_n(x) t^{2nH} + o(t^{2nH})\right).$$

For instance, by applying the previous proposition with $N = 1$, we get

$$a_0(x_0) = \frac{1}{(2\pi)^{\frac{d}{2}}} \frac{1}{|\det(V_1(x_0), \ldots, V_d(x_0))|}.$$

The computation of $a_1(x)$ is technically more involved. We wish to apply the previous proposition with $N = 2$. For that, we need to understand the law of the random variable

$$\Theta_t = \sum_{i=1}^d B_t^i V_i(x_0) + \frac{1}{2} \sum_{1\leq i<j\leq d} \int_0^t B_s^i \mathrm{d}B_s^j - B_s^j \mathrm{d}B_i^s [V_i, V_j](x_0).$$

From the structure equations, we have

$$\Theta_t = \sum_{k=1}^d \left(B_t^k + \frac{1}{2} \sum_{1\leq i<j\leq d} \omega_{ij}^k \int_0^t B_s^i \mathrm{d}B_s^j - B_s^j \mathrm{d}B_s^i\right) V_k(x_0).$$

By a simple linear transformation, this is reduced to the problem of the computation of the law of the $\mathbb{R}^d$-valued random variable

$$\theta_t = \left(B_t^k + \frac{1}{2} \sum_{1 \le i < j \le d} \omega_{ij}^k \int_0^t B_s^i \mathrm{d}B_s^j - B_s^j \mathrm{d}B_s^i \right)_{1 \le k \le d}.$$

At this time, to the knowledge of the authors, there is no explicit formula for this distribution. However, the scaling property of fractional Brownian motion and the inverse Fourier transform formula lead easily to the following expression:

$$p_t(x_0, x_0) = \frac{1}{|\det(V_1(x_0), \dots, V_d(x_0))|} \frac{1}{(2\pi t^{2H})^{d/2}} \left(1 - q_H(\omega) t^{2H} + o(t^{2H})\right),$$

where $q_H(\omega)$ is the quadratic form given by

$$q_H(\omega) = \frac{1}{8(2\pi)^{\frac{d}{2}}} \int_{\mathbb{R}^d} \mathbb{E} \left(\mathrm{e}^{\mathrm{i}\langle \lambda, B_1 \rangle} \left(\sum_{1 \le i < j \le d} \langle \omega_{ij}, \lambda \rangle \int_0^1 B_s^i \mathrm{d}B_s^j - B_s^j \mathrm{d}B_s^i \right)^2 \right) \mathrm{d}\lambda.$$

Acknowledgements

The first author of this research was supported in part by NSF Grant DMS 0907326. The authors would like to thank the anonymous referee for a careful reading of the paper and for many helpful comments and suggestions as regards improving the paper. All errors are our responsibility.

Appendix

In this last section, we give proofs for the technical lemmas that we used before.

Fix $1/2 < \lambda < H$. Let $B(\phi, \rho) \in C^\lambda(0, T; \mathbb{R}^d)$ be the ball centered at ϕ with radius ρ under the λ-Hölder topology

$$\|f\|_{\lambda,T} := \|f\|_\infty + \sup_{0 \le s < t \le T} \frac{|f(t) - f(s)|}{(t-s)^\lambda}, \quad \text{for all } f \in C^\lambda(0, T; \mathbb{R}^d).$$

Note that the λ-Hölder topology is a stronger topology than the usual supreme topology.

Recall the two expressions for Z^ε:

$$\mathrm{d}Z_t^\varepsilon = \sigma(Z_t^\varepsilon)(\varepsilon \mathrm{d}B_t + \mathrm{d}\gamma_t) + b(\varepsilon, Z_t^\varepsilon)\mathrm{d}t \tag{A.1}$$

and

$$Z^\varepsilon = \phi + \sum_{j=0}^N \frac{g_j \varepsilon^j}{j!} + \varepsilon^{N+1} R_{N+1}^\varepsilon. \tag{A.2}$$

Here $\gamma \in \mathscr{H}_H$, hence $\gamma \in I_0^{H+1/2}(L^2) \subset C^H(0, T; \mathbb{R}^d)$.

A.1. Proof of Lemma 3.7

We show, for all $t \in [0, T]$, that there exists a constant C such that for r large enough we have

$$\mathbb{P}\{\|g_1\|_{\lambda,t} \geq r\} \leq \exp\left\{-\frac{Cr^2}{t^{2H}}\right\} \tag{A.3}$$
$$\mathbb{P}\{\|g_2\|_{\lambda,t} \geq r\} \leq \exp\left\{-\frac{Cr}{t^{2H}}\right\},$$

and on $\{t \leq T^\varepsilon\}$, where T^ε is the first existence time of Z^ε from $B(\phi, \rho)$, we have

$$\|\varepsilon R_1^\varepsilon\|_{\lambda,t} \leq \rho \tag{A.4}$$
$$\mathbb{P}\{\|\varepsilon R_2^\varepsilon\|_{\lambda,t} \geq r; t \leq T^\varepsilon\} \leq \exp\left\{-\frac{Cr^2}{\rho^2 t^{2H}}\right\}$$
$$\mathbb{P}\{\|\varepsilon R_3^\varepsilon\|_{\lambda,t} \geq r; t \leq T^\varepsilon\} \leq \exp\left\{-\frac{Cr}{\rho t^{2H}}\right\}.$$

We first prove the estimates for the g_i's. Write

$$\sigma(Z^\varepsilon) = \sigma(\phi) + \sigma_x(\phi)(Z^\varepsilon - \phi) + \frac{1}{2}\sigma_{xx}(\phi)(Z^\varepsilon - \phi)^2 + O(\varepsilon^3) \tag{A.5}$$

and

$$\begin{aligned} b(\varepsilon, z^\varepsilon) &= b(0, \phi) + b_x(0, \phi)(Z^\varepsilon - \phi) + \frac{1}{2}b_{xx}(0, \phi)(Z^\varepsilon - \phi)^2 + O(\varepsilon^3) \\ &\quad + b_\varepsilon(0, \phi)\varepsilon + b_{\varepsilon x}(0, \phi)(Z^\varepsilon - \phi)\varepsilon + O(\varepsilon^3) + \frac{1}{2}b_{\varepsilon\varepsilon}(0, \phi)\varepsilon^2 + O(\varepsilon^3). \end{aligned} \tag{A.6}$$

Substituting into the two expressions for Z^ε gives us

$$\mathrm{d}g_1(s) = \sigma(\phi_s)\mathrm{d}B_s + \sigma_x(\phi_s)g_1(s)\mathrm{d}\gamma_s + b_x(0, \phi_s)g_1(s)\mathrm{d}s + b_\varepsilon(0, \phi_s)\mathrm{d}s \tag{A.7}$$

and

$$\begin{aligned} \mathrm{d}g_2(s) &= 2\sigma_x(\phi_s)g_1(s)\mathrm{d}B_s + \sigma_{xx}(\phi_s)g_1(s)^2\mathrm{d}\gamma_s + \sigma_x(\phi_s)g_2(s)\mathrm{d}\gamma_s + b_{xx}(0, \phi_s)g_1(s)^2\mathrm{d}s \\ &\quad + b_x(0, \phi_s)g_2(s)\mathrm{d}s + b_{\varepsilon\varepsilon}(0, \phi_s)\mathrm{d}s + 2b_{\varepsilon x}(0, \phi_s)g_1(s)\mathrm{d}s. \end{aligned} \tag{A.8}$$

By (A.7) and Remark 2.1, it is clear that

$$\|g_1\|_{\lambda,t} \leq C\|B\|_{\lambda,t}, \quad t \in [0, T],$$

where C is a constant depending only on $\|\phi\|_{\lambda,T}$, $\|\gamma\|_{\lambda,T}$ and T. This gives us the first estimate in (A.3).

Similarly, by (A.8) and Remark 2.1 together with the estimate that we just obtained for g_1, we have

$$\begin{aligned} \|g_2\|_{\lambda,t} &\leq C(1 + \|g_1\|_{\lambda,t} + \|g_1\|_{\lambda,t}^2 + \|g_1\|_{\lambda,t}\|B\|_{\lambda,t}) \\ &\leq C\|B\|_{\lambda,t}^2. \end{aligned}$$

Here C is also a constant, depending only on $\|\phi\|_{\lambda,T}$, $\|\gamma\|_{\lambda,T}$ and T. Hence we have proved (A.3).

In what follows, we prove (A.4). To lighten our notation, in the discussion that follows, we suppress the superscript ε in R_i^ε whenever there is no confusion.

Since we work in $B(\phi, \rho)$, the first inequality in (A.4) is apparent. We therefore only need to concentrate on the last two inequalities.

First we use a similar idea to deduce the equations satisfied by $R_i,\ i = 1, 2, 3$. For this purpose, define μ_1, μ_2 and ν_1, ν_2 by

$$\begin{aligned}\sigma(Z^\varepsilon) &= \sigma(\phi) + \mu_1\varepsilon \\ &= \sigma(\phi) + \sigma_x(\phi)(Z^\varepsilon - \phi) + \mu_2\varepsilon^2\end{aligned} \tag{A.9}$$

and

$$\begin{aligned}b(\varepsilon, Z^\varepsilon) &= b(0, \phi) + \nu_1\varepsilon \\ &= b(0, \phi) + b_x(0, \phi)(Z^\varepsilon - \phi) + b_\varepsilon(0, \phi)\varepsilon + \nu_2\varepsilon^2.\end{aligned} \tag{A.10}$$

It is clear that the μ_i's (resp. ν_i's), $i = 1, 2$ are of the form $\psi_i^\mu(\varepsilon R_1)(R_1)^i$ (resp. $\psi_i^\nu(\varepsilon R_1)(R_1)^i$), where the ψ_i's are some functions of bounded derivatives determined by σ and b. Hence in $B(\phi, \rho)$, μ_1, ν_1 are functions of R_1 with bounded derivatives, and there exists a constant C, depending only on the derivatives of σ and b, such that

$$\|\mu_1\|_{\lambda,t}, \|\nu_1\|_{\lambda,t} \le C(1 + \|R_1\|_{\lambda,t}) \quad \text{and} \quad \|\mu_2\|_{\lambda,t}, \|\nu_2\|_{\lambda,t} \le C(1 + \|R_1\|_{\lambda,t})^2. \tag{A.11}$$

Eqs. (A.2), (A.1), (A.9) and (A.10) give us

$$\begin{aligned}&\mathrm{d}R_1(s) = \sigma(Z_s^\varepsilon)\mathrm{d}B_s + \mu_1\mathrm{d}\gamma_s + \nu_1\mathrm{d}s \\ &\mathrm{d}R_2(s) = 2\mu_1\mathrm{d}B_s + 2\mu_2\mathrm{d}\gamma_s + \sigma_x(\phi)R_2\mathrm{d}\gamma_t + b_x(0, \phi_s)R_2\mathrm{d}s + 2\nu_2\mathrm{d}s.\end{aligned} \tag{A.12}$$

Since we work with in $B(\phi, \rho)$, we have

$$\|Z^\varepsilon\|_{\lambda,t} \le \|\phi\|_{\lambda,t} + \rho$$

and hence

$$\left\| \int_0^t \sigma(Z_s^\varepsilon)\mathrm{d}B_s \right\|_{\lambda,t} < C\|B\|_{\lambda,t}$$

for some constant C depending only on ρ, ϕ and the derivatives of σ. By standard Picard iteration, we conclude that

$$\|R_1\|_{\lambda,t} < C\|B\|_{\lambda,t}(1 + \|\gamma\|_{\lambda,t}), \quad \text{in } B(\phi, \rho) \tag{A.13}$$

for some constant C uniformly bounded in ε.

The equation for R_2 is

$$\mathrm{d}R_2(s) - \sigma_x(\phi)R_2\mathrm{d}\gamma_s - b_x(0, \phi_s)R_2\mathrm{d}s = 2\mu_1\mathrm{d}B_s + 2\mu_2\mathrm{d}\gamma_s + 2\nu_2\mathrm{d}s. \tag{A.14}$$

Recall that μ_1 is of the form $\psi_1(\varepsilon R_1)R_1$, and in $B(\phi, \rho)$,

$$\|\varepsilon R_1\|_{\lambda,t} < \rho.$$

We obtain

$$\left\| \varepsilon \int_0^t \mu_1\mathrm{d}B_s \right\|_{\lambda,t} = \left\| \int_0^t \psi_1^\mu(\varepsilon R_1)(\varepsilon R_1)\mathrm{d}B_s \right\|_{\lambda,t} < \rho^2 C\|B\|_{\lambda,t}$$

for some constant C uniformly bounded in ε. Similarly,

$$\begin{aligned}\left\|\varepsilon\int_0^t \mu_2 \mathrm{d}\gamma_s + \varepsilon\int_0^t \nu_2 \mathrm{d}s\right\|_\alpha &= \left\|\int_0^t \psi_2^\mu(\varepsilon R_1)(\varepsilon R_1^2)\mathrm{d}\gamma_s + \int_0^t \psi_2^\nu(\varepsilon R_1)(\varepsilon R_1^2)\mathrm{d}s\right\|_{\lambda,t} \\ &\le \rho^2 C\|R_1\|_{\lambda,t}(1+\|\gamma\|_{\lambda,t}) \\ &\le \rho^2 C\|B\|_{\lambda,t}(1+\|\gamma\|_{\lambda,t})\end{aligned}$$

for some constant C uniformly bounded in ε. Hence by multiplying by a factor

$$\exp\left\{-\int \sigma_x(\phi)\mathrm{d}\gamma - \int b_x(0,\phi)\mathrm{d}u\right\}$$

on both sides of (A.14) and integrating from 0 to t, we conclude that

$$\mathbb{P}\{\|\varepsilon R_2\|_{\lambda,t} \ge r; t \le T^\varepsilon\} \le \exp\left\{-\frac{Cr^2}{\rho^2 t^{2H}}\right\}.$$

This gives us the desired estimate for εR_2. A similar argument also gives us

$$\mathbb{P}\{\|R_2\|_{\lambda,t} \ge r; t \le T^\varepsilon\} \le \exp\left\{-\frac{Cr}{\rho t^{2H}}\right\}.$$

Continuing this type of argument, the equation for R_3 is given by

$$\begin{aligned}&\mathrm{d}R_3(s) - \sigma_x(\phi)R_3\mathrm{d}\gamma_s - b_\varepsilon(0,\phi)R_3\mathrm{d}s \\ &\quad = \sigma(\phi)R_2\mathrm{d}B_s + \mu_2\mathrm{d}B_s + \mu_3\mathrm{d}\gamma_s + \frac{1}{2}\sigma_{xx}(\phi)R_1R_2\mathrm{d}\gamma_s \\ &\qquad + \frac{1}{4}b_{xx}(0,\phi)R_1R_2\mathrm{d}s + \nu_3\mathrm{d}s + b_{\varepsilon,x}(0,\phi)R_2\mathrm{d}s.\end{aligned}$$

By (A.11) and (A.13) we conclude that in $B(\phi,\rho)$ we have for all $0 < \varepsilon \le \rho$

$$\left\|\varepsilon\int_0^t \sigma(\phi)R_2 + \mu_2\mathrm{d}B_s\right\|_{\lambda,t} < \rho C\|B\|_{\lambda,t}^2,$$

and similar estimates for the rest of the terms on the right hand side of the equation for R_3. Hence

$$\mathbb{P}\{\|\varepsilon R_3\|_{\lambda,t} \ge r; t \le T^\varepsilon\} \le \exp\left\{-\frac{Cr}{\rho t^{2H}}\right\}.$$

Therefore, we have proved (A.4).

A.2. Proof of Lemma 3.9

For the convenience of quick reference, we re-state the lemma here.

Lemma A.1. *Let*

$$\theta(\varepsilon) = F(Z^\varepsilon) = \theta(0) + \varepsilon\theta'(0) + \varepsilon^2 U(\varepsilon)$$

where $U(\varepsilon) = \int_0^1 (1-v)\theta''(\varepsilon v)\mathrm{d}v$. *There exist* $\beta > 0$ *and* $\varepsilon_0 > 0$ *such that*

$$\sup_{0\le\varepsilon\le\varepsilon_0} \mathbb{E}\left(\mathrm{e}^{-(1+\beta)U(\varepsilon)}; Z^\varepsilon \in B(\phi,\rho)\right) < \infty.$$

Observe that

$$(Z^\varepsilon - \phi)^2 = \varepsilon^2 g_1^2 + \frac{1}{2}\varepsilon^3 g_1 R_2 + \frac{1}{2}\varepsilon^3 R_1 R_2.$$

Thus, if we write

$$U(\varepsilon) = \frac{1}{2}\theta''(0) + \varepsilon R(\varepsilon)$$

then

$$|R(\varepsilon)| \le C(|R_3| + |g_1|\,|R_2| + |R_2|\,|R_1| + |R_1^3|). \tag{A.15}$$

Together with the fact that $|\varepsilon R_1| \le \rho$, this gives

$$|\varepsilon R(\varepsilon)| \le C(|\varepsilon R_3| + |g_1|\,|\varepsilon R_2| + \rho|R_2| + \rho|R_1^2|).$$

Hence, from the estimates in Lemma 3.7, we conclude that for each $\alpha > 0$, there exists $\rho(\alpha)$ such that for all $\varepsilon \le \rho \le \rho(\alpha)$, we have

$$\sup_{0\le\varepsilon\le\rho} \mathbb{E}\left(\mathrm{e}^{(1+\alpha)|\varepsilon R(\varepsilon)|}; t \le T^\varepsilon\right) < \infty.$$

Therefore, proving Lemma 3.9 is reduced to proving the following:

Lemma A.2. *There exists a $\beta > 0$ such that*

$$\mathbb{E}\exp\left\{-(1+\beta)\left[\frac{1}{2}\theta''(0)\right]\right\} < \infty.$$

Proof. We follow the proof in [7]. Since

$$U(0) = \frac{1}{2}\theta''(0) = \frac{1}{2}\left[\mathrm{d}F(\theta)g_2 + \mathrm{d}^2F(\theta)g_1^2\right],$$

it is clear that to prove the above lemma, it suffices to prove that for sufficiently large r we have

$$\mathbb{P}\left\{-\frac{1}{2}\left[\mathrm{d}F(\phi)g_2 + \mathrm{d}^2F(\phi)g_1^2\right] \ge r\right\} \le \mathrm{e}^{-Cr}, \quad \text{with } C > 1. \tag{A.16}$$

Set

$$Y^\varepsilon = (\varepsilon g_1, \varepsilon^2 g_2)$$

with

$$\mathrm{d}Y_s^\varepsilon = \varepsilon\bar{\sigma}(s, Y^\varepsilon)\mathrm{d}B_s + \bar{b}(\varepsilon, s, Y^\varepsilon)\mathrm{d}s, \quad Y_0^\varepsilon = 0.$$

Here $\bar{\sigma}$ and $\bar{b}$ are determined by (A.7) and (A.8). Define $A \subset C([0,T], \mathbb{R}^{2d})$ by

$$A = \{\psi = (\psi_1, \psi_2) \in C([0,T], \mathbb{R}^d \times \mathbb{R}^d) : \ \mathrm{d}F(\phi)\psi_2 + \mathrm{d}^2F(\phi)\psi_1^2 \le -2\}.$$

We have

$$\mathbb{P}\{Y^\varepsilon \in A\} = \mathbb{P}\left\{-\frac{1}{2}\left[\mathrm{d}F(\phi)g_2 + \mathrm{d}^2F(\phi)g_1^2\right] \ge \frac{1}{\varepsilon^2}\right\}$$

and by the large deviation principle for Y^ε,

$$\limsup_{\varepsilon\to 0} \varepsilon^2 \log \mathbb{P}\{Y^\varepsilon \in A\} \le -\Lambda^*(A).$$

Here Λ^* is a good rate function of Y^ε. It is clear that to prove inequality (A.16) it suffices to prove that $\Lambda^*(A) > 1$.

Recall that

$$\Lambda^*(A) = \inf\left\{\frac{1}{2}|k|^2_{\mathscr{H}_H};\ \Phi^*(k) \in A\right\}$$

where $u = \Phi^*(k)$ is the solution to the ordinary differential equation

$$\mathrm{d}u_s = \bar{\sigma}(s, u_s)\mathrm{d}k_s + \bar{b}(0, s, u_s)\mathrm{d}s, \quad \text{with } u_0 = 0.$$

It is easy to see from (3.4), (3.5), (A.7) and (A.8) that we have explicitly

$$u = (\mathrm{d}\Phi(\gamma)k, \mathrm{d}^2\Phi(\gamma)k^2).$$

By our assumption H2 and the explanation after it, there exists $\nu \in (0, 1)$ such that for all $k \in \mathscr{H}_H - \{0\}$ we have

$$\mathrm{d}^2 F \circ \Phi(\gamma)k^2 > (-1 + \nu)|k|^2_{\mathscr{H}_H},$$

or

$$|k|^2_{\mathscr{H}_H} > -\frac{1}{1-\nu}(\mathrm{d}^2 F \circ \Phi(\gamma)k^2) = -\frac{1}{1-\nu}(\mathrm{d}^2 F(\phi)(\mathrm{d}\Phi(\gamma)k) + \mathrm{d}F(\phi)(\mathrm{d}^2\Phi(\gamma)k^2)).$$

Therefore, if $\Phi^*(k) \in A$, we have

$$\frac{1}{2}|k|_{\mathscr{H}_H} > \frac{1}{1-\nu} > 1,$$

which implies $\Lambda^*(A) > 1$ and completes the proof. □

References

[1] R. Azencott, Formule de Taylor stochastique et développements asymptotiques d'intégrales de Feynman, in: Yor Azema (Ed.), Séminaire de Probabilités XVI, in: LNM, vol. 921, Springer, 1982, pp. 237–284.

[2] R. Azencott, Densité des diffusions en temps petit: développements asymptotiques (part I), Sem. Probab. 18 (1984) 402–498.

[3] F. Baudoin, An Introduction to the Geometry of Stochastic Flows, Imperial College Press, 2004.

[4] F. Baudoin, Stochastic Taylor expansions and heat kernel asymptotics, ESAIM Probab. Stat. (2010) (in press).

[5] F. Baudoin, L. Coutin, Operators associated with a stochastic differential equation driven by fractional Brownian motions, Stochastic Process. Appl. 117 (5) (2007) 550–574.

[6] F. Baudoin, M. Hairer, A version of Hörmander's theorem for the fractional Brownian motion, Probab. Theory Related Fields 139 (3–4) (2007) 373–395.

[7] G. Ben Arous, Méthodes de Laplace et de la phase stationnaire sur l'espace de Wiener [The Laplace and stationary phase methods on Wiener space], Stochastics 25 (3) (1988) 125–153 (in French).

[8] G. Ben Arous, Développement asymptotique du noyau de la chaleur hypoelliptique hors du cut-locus [Asymptotic expansion of the hypoelliptic heat kernel outside of the cut-locus], Ann. Sci. Éc. Norm. Supér (4) 21 (3) (1988) 307–331 (in French).

[9] G. Ben Arous, Flots et séries de Taylor stochastiques, J. Probab. Theory Related Fields 81 (1989) 29–77.

[10] L. Decreusefond, A.S. Üstünel, Stochastic analysis of the fractional Brownian motion, Potential Anal. 10 (1998) 177–214.

[11] P. Friz, N. Victoir, Multidimensional Stochastic Processes as Rough Paths. Theory and Applications, in: Cambridge Studies in Advanced Mathematics, CUP, 2009.

[12] Y. Hu, D. Nualart, Differential equations driven by Holder continuous functions of order greater than 1/2, in: Abel Symp., vol. 2, Springer, Berlin, 2007, pp. 399–413.
[13] Y. Inahama, Laplace approximation for rough differential equation driven by fractional Brownian motion, ArXiv Preprint, 2010. arXiv:1004.1478.
[14] H. Kunita, Asymptotic self-similarity and short time asymptotics of stochastic flows, J. Math. Sci. Univ. Tokyo 4 (1997) 595–619.
[15] R. Léandre, Développement asymptotique de la densité d'une diffusion dégénérée, Forum Math. 4 (1) (1992) 45–75.
[16] T. Lyons, M. Caruana, T. Lévy, Differential Equations Driven by Rough Paths, Lectures from the 34th Summer School on Probability Theory Held in Saint-Flour, July 6–24, 2004. With an Introduction Concerning the Summer School by Jean Picard, in: Lecture Notes in Mathematics, vol. 1908, Springer, Berlin, 2007, xviii+109 pp.
[17] P. Malliavin, Stochastic Analysis, in: Grundlehren der Mathematischen Wissenschaften (Fundamental Principles of Mathematical Sciences), vol. 313, Springer-Verlag, Berlin, 1997, xii+343 pp.
[18] A. Millet, M. Sanz, Large deviations for rough paths of the fractional Brownian motion, Ann. Inst. H. Poincaré Probab. Statist. 42 (2) (2006) 245–271 (English, French summary).
[19] A. Neuenkirch, I. Nourdin, A. Rössler, S. Tindel, Trees and asymptotic expansions for fractional stochastic differential equations, Ann. Inst. H. Poincaré Probab. Statist. 45 (1) (2009) 157–174. (English, French summary).
[20] D. Nualart, A. Răşcanu, Differential equations driven by fractional Brownian motion, Collect. Math. 53 (1) (2002) 55–81.
[21] D. Nualart, B. Saussereau, Malliavin calculus for stochastic differential equations driven by a fractional Brownian motion, Stochastic Process. Appl. 119 (2) (2009) 391–409.
[22] I. Shigekawa, Stochastic Analysis, in: Translations of Mathematical Monographs, vol. 224, AMS, 2004.
[23] L.C. Young, An inequality of the Hölder type connected with Stieltjes integration, Acta Math. 67 (1936) 251–282.
[24] M. Zähle, Integration with respect to fractal functions and stochastic calculus I, Probab. Theory Related Fields 111 (1998) 333–374.