J. London Math. Soc. (2) 73 (2006) 797–816 © 2006 London Mathematical Society
doi:10.1112/S0024610706022654

LARGE DEVIATIONS FOR HEAT KERNEL MEASURES ON LOOP SPACES VIA ROUGH PATHS

YUZURU INAHAMA AND HIROSHI KAWABI

Abstract

The main objective of this paper is to establish a large deviation principle for heat kernel measures on loop spaces. It gives an extension of Fang and Zhang's results on loop groups. For the proof, we use the continuity theorem of Lyons' rough path theory.

1. *Introduction*

In 2001, Fang and Zhang [**5**] proved a large deviation principle for heat kernel measures ν_t on the loop group over a compact Lie group as $t \searrow 0$. Roughly speaking, ν_t is the law of X_t, where $(X_t)_{t\geqslant 0}$ is the diffusion process on the loop group associated with the Gross Laplacian. The process $(X_t)_{t\geqslant 0}$ satisfies a certain infinite-dimensional stochastic differential equation (SDE) (see Malliavin [**15**] and (4.4) below). Their method is the projective limit technique (Dawson–Gärtner's theorem). They used the symmetry of the Lie group to prove the exponential tightness.

On the other hand, Ledoux, Qian and Zhang [**13**] gave a new proof for large deviations of Freidlin–Wentzell type by using Lyons' rough path theory. The basic idea in [**13**] is as follows: first they showed that the laws of Brownian rough paths satisfy the large deviation principle; then they used the contraction principle since the Itô map is continuous in the rough path theory.

In this paper we extend Fang and Zhang's result by using the rough path theory in a similar way to [**13**]. We slightly generalize the heat kernel measure ν_t so that the new definition includes not only heat kernel measures on loop groups but also heat kernel measures on loop spaces over compact Riemannian manifolds and many others. (See Examples 4.3–4.5 later.) In the main theorem (Theorem 4.9), we prove that a large deviation principle holds for the generalized heat kernel measures $\{\nu_t\}_{t\geqslant 0}$.

What seems novel in this paper is the following. (1) In Section 3 we prove a large deviation principle for the laws of Banach space-valued Brownian rough paths under the 'exactness' assumption. (2) We prove that the process $(X_t)_{t\geqslant 0}$ defined by the SDE (4.4) is a nice concrete example to which the Banach space-valued rough path theory can be applied. (3) Since Fang and Zhang used the symmetry of the Lie group in their proof for the exponential tightness, the case of loop spaces over compact Riemannian manifolds has not been solved. Our results include such a case in a natural way using the rough path theory. (Note that the result in this paper can be considered as a rigorous proof for a special case of Léandre's

Received 1 October 2004.

2000 *Mathematics Subject Classification* 60B12 (primary), 58J65, 60G15 (secondary).

This research was supported by JSPS Research Fellowships for Young Scientists.

https://doi.org/10.1112/S0024610706022654 Published online by Cambridge University Press

result in [**11**].) (4) We give a new proof for the existence of continuous modifications of heat processes (Theorem 4.8). In this paper we assume for simplicity that the starting loop is the constant loop at 0. However, it is easy to modify the proof for a general starting loop since the initial condition may be arbitrary in the rough path theory. A merit of this method is that Kolmogorov's criterion is not necessary.

The organization of this paper is as follows. In Section 2 we give a simple review of the rough path theory. In Section 3 we prove a large deviation principle for the laws of Banach space-valued Brownian rough paths in the same way as in [**13**]. In Section 4 we generalize heat kernel measures in Definition 4.2 and show in Lemma 4.8 that the above-defined heat process $(X_t)_{t\geqslant 0}$ can be regarded as the image of the Brownian rough paths via the Itô map. As a result our main theorem (Theorem 4.9) is easily proved by using the contraction principle. In Remark 4.10, we explain why we work on the space of continuous loops and not on the Sobolev–Slobodetski space $W^{\theta,p}(S^1)$.

2. *A simple review of the rough path theory*

In this section, we set notation and review some basic results of the rough path theory. This section contains no new results.

Let B be a real separable Banach space. The algebraic tensor product is denoted by $B\otimes_{\mathrm{a}} B$. There is a natural injection $\iota : B\otimes_{\mathrm{a}} B \hookrightarrow L(B^*,B)$ between two linear spaces. Here, $L(B^*,B)$ is the space of bounded linear maps form B^* to B. We consider a norm $|\cdot|$ on $B\otimes_{\mathrm{a}} B$ such that $|x\otimes y|\leqslant |x|_B\cdot|y|_B$ holds for all $x,y\in B$. We denote by $B\otimes B$ the completion of $B\otimes_{\mathrm{a}} B$ by this norm. When confusion may arise we write $|\cdot|_{B\otimes B}$ for $|\cdot|$.

Let $2<p<3$. (In this paper we only consider this case.) A continuous map $Y=(1,Y^1,Y^2)$ from the simplex $\Delta=\{(s,t)\mid 0\leqslant s\leqslant t\leqslant 1\}$ to the truncated tensor algebra $T^{(2)}(B)=\mathbb{R}\oplus B\oplus(B\otimes B)$ is said to be a rough path if, for every $s\leqslant u\leqslant t$, $Y_{s,t}=Y_{s,u}\otimes Y_{u,t}$ and

$$\left(\sup_D\sum_l |Y^j_{t_{l-1},t_l}|^{p/j}\right)^{j/p}<\infty \quad \text{for } j=1,2$$

hold, where $D=\{0=t_0<t_1<\ldots<t_N=1\}$ runs over all finite partitions of the interval $[0,1]$. For two rough paths X and Y, p-variation distance is defined as follows:

$$d_p(X,Y)=\max_{j=1,2}\left(\sup_D\sum_l |X^j_{t_{l-1},t_l}-Y^j_{t_{l-1},t_l}|^{p/j}\right)^{j/p}.$$

Let $y=(y_t)_{0\leqslant t\leqslant 1}$ be a B-valued continuous path of finite total variation and set $Y^1_{s,t}:=y_t-y_s, Y^2_{s,t}:=\int_s^t (y_u-y_s)\otimes dy_u$ for $0\leqslant s\leqslant t\leqslant 1$. A rough path obtained in this way is called a smooth rough path and we write $Y=F(y)$. A rough path obtained as the d_p-limit of a sequence of smooth rough paths is called a geometric rough path. The set of all the geometric rough paths is denoted by $G\Omega_p(B)$.

We introduce the dyadic polygonal approximation. Let $t^m_l=l/2^m$ for $m\in\mathbb{N}$ and $l=0,1,\ldots,2^m$. For a B-valued continuous path $y=(y_t)_{0\leqslant t\leqslant 1}$, we set

$$y(m)_t=y_{t^m_{l-1}}+2^m(t-t^m_{l-1})(y_{t^m_l}-y_{t^m_{l-1}}) \quad \text{if } t_{t^m_{l-1}}\leqslant t\leqslant t_{t^m_l}.$$

Clearly, $y(m)$ is of finite total variation. The smooth rough path associated to $y(m)$ is denoted by $Y(m)=(1,Y(m)^1,Y(m)^2)=F(y(m))$.

https://doi.org/10.1112/S0024610706022654 Published online by Cambridge University Press

The following equalities (2.1)–(2.5) are basic properties of $Y(m)$, which can easily be verified by straightforward computation. Later, these equalities will be used frequently. The first level path $Y(m)^1$ satisfies

$$Y(m)^1_{t^n_{l-1},t^n_l} = 2^{m-n}\Delta^m_{k(n,m,l)}y \quad \text{for } n \geqslant m \tag{2.1}$$

and

$$Y(m)^1_{t^n_{l-1},t^n_l} = \Delta^n_l y \quad \text{for } n \leqslant m. \tag{2.2}$$

Here $\Delta^n_l y := y_{t^n_l} - y_{t^n_{l-1}}$ and $k(n,m,l)$ is the unique integer such that $t^m_{k(n,m,l)-1} \leqslant t^n_{l-1} < t^n_l \leqslant t^m_{k(n,m,l)}$. The second level path $Y(m)^2$ satisfies

$$Y(m)^2_{t^n_{l-1},t^n_l} = 2^{2(m-n)-1}(\Delta^m_{k(n,m,l)}y \otimes \Delta^m_{k(n,m,l)}y) \quad \text{for } n \geqslant m \tag{2.3}$$

and

$$Y(m)^2_{t^n_{l-1},t^n_l} = \frac{1}{2}\Delta^n_l y \otimes \Delta^n_l y + \frac{1}{2}\sum(\Delta^m_r y \otimes \Delta^m_s y - \Delta^m_s y \otimes \Delta^m_r y) \quad \text{for } n \leqslant m, \tag{2.4}$$

where the sum is taken over $\{(r,s) \mid 2^{m-n}(l-1)+1 \leqslant r < s \leqslant 2^{m-n}l\}$.

By (2.4), we also have

$$\begin{aligned} &Y(m+1)^2_{t^n_{l-1},t^n_l} - Y(m)^2_{t^n_{l-1},t^n_l} \\ &\quad = \frac{1}{2}\Delta^n_l y \otimes \Delta^n_l y + \frac{1}{2}\sum_{r=2^{m-n}(l-1)+1}^{2^{m-n}l}(\Delta^{m+1}_{2r-1}y \otimes \Delta^{m+1}_{2r}y - \Delta^{m+1}_{2r}y \otimes \Delta^{m+1}_{2r-1}y) \\ &\qquad \text{for } n \leqslant m. \end{aligned} \tag{2.5}$$

The rest of this section is devoted to proving a property of the direct sum and the projective tensor products of Banach spaces. For basic results on the tensor product of Banach spaces, see, for example, Lyons and Qian [**14**, Section 5.6]. In the following, A, B, A_i, B_i, $i = 1, 2$, are real separable Banach spaces.

First we recall the definition of the projective tensor product $A \otimes_{\mathrm{pr}} B$. It is the completion of the algebraic tensor product $A \otimes_{\mathrm{a}} B$ with respect to the projective tensor norm defined by

$$|v|_{\mathrm{pr}} = \inf\left\{\sum_{k=1}^n |a^k|_A|b^k|_B \;\middle|\; \text{if } v = \sum_{k=1}^n a^k \otimes b^k \in A \otimes_{\mathrm{a}} B\right\}.$$

LEMMA 2.1. *Let the notation be as above. Then, the two Banach spaces* $(A_1 \oplus A_2) \otimes_{\mathrm{pr}} (B_1 \oplus B_2)$ *and* $\oplus_{i,j=1,2}(A_i \otimes_{\mathrm{pr}} B_j)$ *are isometric.*

Proof. In this proof we denote by $|\cdot|_{\mathrm{m}}$ the norm of the second Banach space. We denote the projective norm of the first Banach space and $A_i \otimes_{\mathrm{pr}} B_j$ by $|\cdot|_{\oplus,\mathrm{pr}}$ and $|\cdot|_{\mathrm{pr}}$, respectively. Since $|\cdot|_{\mathrm{pr}}$ is a cross norm on $A_i \otimes_{\mathrm{a}} B_j$, it is easy to see that $|\cdot|_{\mathrm{m}}$ is a cross norm on

$$(A_1 \oplus A_2) \otimes_{\mathrm{a}} (B_1 \oplus B_2) = \bigoplus_{i,j=1,2}(A_i \otimes_{\mathrm{a}} B_j)$$

in the sense that $|(a_1,a_2) \otimes (b_1,b_2)|_{\mathrm{m}} = (|a_1|+|a_2|)(|b_1|+|b_2|)$ holds for all $a_i \in A_i$ and $b_i \in B_i$. Since the projective norm is the largest among the cross norms, we have $|\cdot|_{\mathrm{m}} \leqslant |\cdot|_{\oplus,\mathrm{pr}}$ on $(A_1 \oplus A_2) \otimes_{\mathrm{a}} (B_1 \oplus B_2)$.

We prove the reverse inequality. Let $v_{11} \in A_1 \otimes_{\mathrm{a}} B_1$. Through the natural identification we may consider $v_{11} = (v_{11},0,0,0)$. If $(v_{11},0,0,0) = \sum_{k=1}^n (a^k_1, a^k_2) \otimes (b^k_1, b^k_2)$,

https://doi.org/10.1112/S0024610706022654 Published online by Cambridge University Press

then $v_{11} = \sum_{k=1}^{n} a_1^k \otimes b_1^k$ and $(v_{11}, 0, 0, 0) = \sum_{k=1}^{n} (a_1^k, 0) \otimes (b_1^k, 0)$. Moreover, it is obvious that

$$\sum_{k=1}^{n} |(a_1^k, a_2^k)| \cdot |(b_1^k, b_2^k)| \geqslant \sum_{k=1}^{n} |(a_1^k, 0)| \cdot |(b_1^k, 0)| = \sum_{k=1}^{n} |a_1^k| \cdot |b_1^k|.$$

By the definition of the projective norm we see that $|(v_{11}, 0, 0, 0)|_{\oplus,\mathrm{pr}} = |v_{11}|_{\mathrm{pr}}$. Other components can be done in the same way. By using the triangle inequality for $|\cdot|_{\oplus,\mathrm{pr}}$ and the above equality, we can easily see that $|\cdot|_{\mathrm{m}} \geqslant |\cdot|_{\oplus,\mathrm{pr}}$ on $(A_1 \oplus A_2) \otimes_{\mathrm{a}} (B_1 \oplus B_2)$. □

For a real separable Banach space B, we denote by $\mathrm{BV}(B)$ the set of B-valued continuous paths of finite total variation that start at the origin. Equipped with the total variation norm, $\mathrm{BV}(B)$ is a real Banach space.

LEMMA 2.2. *Let B_1 and B_2 be real separable Banach spaces. In this lemma, we consider the projective norm for Banach tensor products. Set $\iota(X, \lambda)$ by $\iota(X, \lambda)^1 (s, t) = (X^1(s, t), \Lambda^1(s, t))$ and*

$$\iota(X, \lambda)^2(s, t) = \left(X^2(s, t), \int_s^t X^1(s, u) \otimes d\lambda(u), \int_s^t \Lambda^1(s, u) \otimes dX^1(0, u), \Lambda^2(s, t) \right) \tag{2.6}$$

for $X \in G\Omega_p(B_1)$ and $\lambda \in \mathrm{BV}(B_2)$. Here, the second and the third components on the right-hand side should be understood as Young integrals and $\Lambda^1(s, t) = \lambda(t) - \lambda(s)$ and $\Lambda^2(s, t) = \int_s^t \Lambda^1(s, u) \otimes d\lambda^1(u)$. Then, $\iota : G\Omega_p(B_1) \times \mathrm{BV}(B_2) \to G\Omega_p(B_1 \oplus B_2)$ is continuous.

Proof. By Lemma 2.1, it is sufficient to estimate each component on the right-hand side of (2.6). Hence, we can prove the lemma in same way as in the finite-dimensional case. □

3. *Large deviation on the space of rough paths*

In this section, we prove a large deviation principle on the space of Banach-valued rough paths. Finite-dimensional cases are proved in Ledoux, Qian and Zhang [**13**]. However, the dimension of the vector space is (at least apparently) used in the proof in [**13**]. Hence, by using the materials from [**12**], we check that the proof in [**13**] also works for the infinite-dimensional case.

Let (B, H, μ) be an abstract Wiener space in the sense of Gross [**6**], that is, B is a real separable Banach space, H is a real separable Hilbert space which is continuously and densely embedded in B and μ is a Gaussian measure on B such that

$$\int_B \exp(\sqrt{-1}\langle \phi, x \rangle) \mu(dx) = \exp\left(-\frac{1}{2} |\phi|_{H^*}^2 \right) \quad \text{for all } \phi \in B^*.$$

Let $(w_t)_{0 \leqslant t \leqslant 1}$ be the B-valued Brownian motion on a completed probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with marginal distribution μ. We denote by $P(B)$ the Banach space of paths x from $[0, 1]$ into B with $x_0 = 0$ under the norm $\|x\|_{P(B)} := \sup\{|x_t|_B : 0 \leqslant t \leqslant 1\}$ and $L_0^{2,1}(H)$ the subspace of $P(B)$ consisting of absolutely continuous paths

https://doi.org/10.1112/S0024610706022654 Published online by Cambridge University Press

into H with square integrable derivatives, equipped with the norm $\|x\|^2_{L_0^{2,1}(H)} := \int_0^1 |\dot{x}_t|_H^2\, dt$. In the following we may use the convention that $|h|_H := \infty$ for $h \in B \backslash H$ and $\|x\|_{L_0^{2,1}(H)} := \infty$ for $x \in P(B) \setminus L_0^{2,1}(H)$.

We impose the following Assumption A on the tensor norm.

ASSUMPTION A (Exactness of $B \otimes B$ with respect to μ). There exist constants $C > 0$ and $\alpha \in [1/2, 1)$ such that, for any N and for any sequence $\{G_l, \tilde{G}_l\}_{l=1}^N$ of independent, B-valued random variables with common distribution μ,

$$\mathbb{E}\left[\left|\sum_{l=1}^N G_l \otimes \tilde{G}_l\right|_{B\otimes B}\right] \leqslant CN^\alpha$$

holds.

Under this assumption, Ledoux, Lyons and Qian proved in [**12**, Theorem 3] that the Brownian rough path $W = (1, W^1, W^2)$ exists as the almost sure limit of the dyadic polygonal approximations. In other words, $W = \lim_{m\to\infty} W(m)$, where $W(m) = F(w(m))$ is a smooth rough path associated to the dyadic polygonal approximation of w. (See the previous section.) Here, the limit is taken with respect to the p-variation metric d_p, $2 < p < 3$. We write $W = F(w)$. Now F is considered to be a measurable map from $P(B)$ to $G\Omega_p(B)$. Moreover, almost surely $W^1_{s,t} = w_t - w_s$ for $s \leqslant t$.

REMARK 3.1. By abusing the notation we denote by $\mathbb{P}$ the Wiener measure on $P(B)$ defined as above. For any $\varepsilon > 0$, we denote by $\mathbb{P}_\varepsilon$ the image measure of $\mathbb{P}$ by the mapping $w \in P(B) \mapsto \varepsilon w \in P(B)$. We have just stated that the set

$$A = \{w \in P(B) \mid \{F(w(m))\}_{m\in\mathbb{N}} \text{ is a Cauchy sequence in } G\Omega_p(B)\}$$

is of $\mathbb{P}$-measure one. Since A is invariant under the mapping $w \mapsto \varepsilon w$, A is of $\mathbb{P}_\varepsilon$-measure one, too. Therefore, how the measurable map F is defined on the complement set of A does not matter when we consider the family of measures $\{\mathbb{P}_\varepsilon\}_{\varepsilon>0}$. In particular, notation like $F(\varepsilon w) = (1, \varepsilon W^1, \varepsilon^2 W^2)$, etc., are well-defined even though $\mathbb{P}_\varepsilon$ and $\mathbb{P}$ are mutually singular unless $\varepsilon = 1$.

Our purpose in this section is to prove the following theorem.

THEOREM 3.2. *Let the notation be as above. We assume Assumption* A. *Then, a large deviation principle for the Brownian rough path $F(\varepsilon w) = (1, \varepsilon W^1, \varepsilon^2 W^2)$, $\varepsilon > 0$, on $G\Omega_p(B)$, $2 < p < 3$, holds with the following good rate function: for $X \in G\Omega_p(B)$,*

$$I(X) = \begin{cases} \frac{1}{2}\|x\|^2_{L_0^{2,1}(H)} & \text{if } X = F(x) \text{ for some } x \in L_0^{2,1}(H), \\ \infty & \text{otherwise.} \end{cases}$$

More precisely, for any measurable set $K \subset G\Omega_p(B)$, the following inequalities hold:

$$-\inf_{X\in K^\circ} I(X) \leqslant \liminf_{\varepsilon\searrow 0} \varepsilon^2 \log \mathbb{P}(F(\varepsilon w) \in K) \leqslant \limsup_{\varepsilon\searrow 0} \varepsilon^2 \log \mathbb{P}(F(\varepsilon w) \in K) \leqslant -\inf_{X\in\overline{K}} I(X).$$

We will prove this theorem at the end of this section. To do this we need some lemmas. The following is from Pisier [**16**, Corollary 2.5].

https://doi.org/10.1112/S0024610706022654 Published online by Cambridge University Press

LEMMA 3.3. *Let $0 < p < q < \infty$. There exists a positive constant $K(p,q)$ such that, for any Gaussian random variable X on a Banach space with the norm $\|\cdot\|$,*

$$\mathbb{E}[\|X\|^q]^{1/q} \leqslant K(p,q)\mathbb{E}[\|X\|^p]^{1/p}$$

holds. Moreover, for each fixed p, $K(p,q) = O(\sqrt{q})$ as $q \to \infty$.

LEMMA 3.4. *Let $B^N = B \oplus \ldots \oplus B$ be the direct sum of N and let $\mu^{\otimes N}$ be the product measure on B^N. Set, for $x = (x_1, \ldots, x_N) \in B^N$ and $q \geqslant 1$,*

$$\|x\|_q^q := \int_{B^N} \left| \sum_{l=1}^N x_l \otimes y_l \right|_{B\otimes B}^q \mu^{\otimes N}(dy).$$

Then, $\|\cdot\|_q$ defines a norm on B^N such that $\|\cdot\|_q \leqslant C\|\cdot\|_{B^N}$ holds for some constant $C > 0$. Here $\|\cdot\|_{B^N}$ is the usual direct sum norm. In particular, any B^N-valued Gaussian random variable can naturally be regarded as a $(\overline{B^N}, \|\cdot\|_q)$-valued Gaussian random variable.

Proof. The triangle inequality is trivial. We prove that $x = 0$ if $\|x\|_q = 0$. Suppose that $\|x\|_q = 0$. Then, $\sum_{l=1}^N x_l \otimes y_l = 0$ for almost all $y = (y_1, \ldots, y_N) \in B_N$. The mapping $y \in B_N \mapsto \sum_{l=1}^N x_l \otimes y_l$ is continuous since $B \otimes B$ is equipped with a tensor norm. Since $\mu^{\otimes N}$ is of full support, we see that the set of $\mu^{\otimes N}$-measure one is dense in B^N and that $\sum_{l=1}^N x_l \otimes y_l = 0$ for all $y = (y_1, \ldots, y_N) \in B_N$. From this we have $x = 0$. We can easily see from Fernique's theorem that $\|x\|_q \leqslant C\|x\|_{B^N}$ holds for all $x \in B^N$. Hence the last assertion is trivial. □

LEMMA 3.5. *Assume Assumption* A *and let $\{G_l, \tilde{G}_l\}_{l=1}^N$ be independent B-valued random variables with common distribution μ. Then there is a positive constant C independent of N and q such that*

$$\mathbb{E}\left[\left| \sum_{l=1}^N G_l \otimes \tilde{G}_l \right|_{B\otimes B}^q \right]^{1/q} \leqslant CqN^\alpha$$

holds for all $N \in \mathbb{N}$ and $q \geqslant 1$.

Proof. We denote by $\mathbb{E}_1$ and $\mathbb{E}_2$ the integrations with respect to $\{G_l\}_{l=1}^N$ and $\{\tilde{G}_l\}_{l=1}^N$, respectively. Note that $\mathbb{E} = \mathbb{E}_1\mathbb{E}_2$. Since, for each fixed sample $\{G_l\}_{l=1}^N$, $\sum_{l=1}^N G_l \otimes \tilde{G}_l$ is a $B \otimes B$-valued Gaussian random variable under $\mathbb{E}_2$, we obtain

$$\mathbb{E}_2\left[\left| \sum_{l=1}^N G_l \otimes \tilde{G}_l \right|_{B\otimes B}^q \right]^{1/q} \leqslant K(1,q)\mathbb{E}_2\left[\left| \sum_{l=1}^N G_l \otimes \tilde{G}_l \right|_{B\otimes B} \right],$$

where we have used Lemma 3.3. Then, by applying Lemmas 3.3 and 3.4 to the $(\overline{B^N}, \|\cdot\|_q)$-valued Gaussian random variables $\{G_l\}_{l=1}^N$, we obtain

$$\begin{aligned}
\mathbb{E}\left[\left| \sum_{l=1}^N G_l \otimes \tilde{G}_l \right|_{B\otimes B}^q \right]^{1/q} &= \mathbb{E}_1[\|(G_1, \ldots, G_N)\|_q^q]^{1/q} \\
&\leqslant K(1,q)\mathbb{E}_1[\|(G_1, \ldots, G_N)\|_q] \\
&\leqslant K(1,q)^2\mathbb{E}\left[\left| \sum_{l=1}^N G_l \otimes \tilde{G}_l \right|_{B\otimes B} \right].
\end{aligned}$$

https://doi.org/10.1112/S0024610706022654 Published online by Cambridge University Press

By Assumption A, we complete the proof. □

The following is well-known in the theory of large deviations. For details the readers should refer to Dembo and Zeitouni [**3**, Theorem 4.2.23] or Deuschel and Stroock [**4**, Lemma 2.1.4].

PROPOSITION 3.6. *Let I be a good rate function on a separable metric space $\mathcal{X}$ and f a measurable map from $\mathcal{X}$ into a second separable metric space $(\mathcal{Y}, d)$. Assume that there exists a sequence $\{f_m\}_{m=1}^{\infty} \subset C(\mathcal{X}, \mathcal{Y})$ such that, for each $0 < \beta < \infty$,*

$$\lim_{m\to\infty} \sup\{d(f_m(x), f(x)) \mid x \in \mathcal{X} \text{ with } I(x) \leqslant \beta\} = 0.$$

Then the map $I' : \mathcal{Y} \to [0, \infty]$ given by

$$I'(y) = \inf\{I(x) \mid x \in \mathcal{X} \text{ and } y = f(x)\}, \quad y \in \mathcal{Y},$$

is a good rate function. Moreover, if, in addition, $\{\mu_\varepsilon\}_{\varepsilon>0}$ has the property that

$$\lim_{m\to\infty} \limsup_{\varepsilon\to 0} \varepsilon^2 \log \mu_\varepsilon(\{x \in \mathcal{X} \mid d(f_m(x), f(x)) > \delta\}) = -\infty$$

for each $0 < \delta < \infty$, then I' governs the large deviation principle of $\{\mu_\varepsilon \circ f^{-1}\}_{\varepsilon>0}$ whenever I governs the large deviation principle of $\{\mu_\varepsilon\}_{\varepsilon>0}$.

Proof of Theorem 3.2. In this proof, we denote by C an unimportant positive constant which may vary from line to line. When we need to specify the dependence on some parameters, we use the subscripts C_p, $C_{\gamma,p}$, etc. The roughness $p \in (2, 3)$ is arbitrarily chosen and fixed throughout this proof.

Let $(w_t)_{0\leqslant t\leqslant 1}$ be a B-valued Brownian motion associated with μ. Then, the law of $(w_t)_{0\leqslant t\leqslant 1}$ is a Gaussian measure on $P(B)$ and its Cameron–Martin space is $L_0^{2,1}(H)$. Now we adopt Proposition 3.6. We set $\mathcal{X} = P(B)$, $\mathcal{Y} = G\Omega_p(B)$, μ_ε is the law of $(\varepsilon w_t)_{0\leqslant t\leqslant 1}$, $f_m(x) = F(x(m))$ and $f(x) = F(x)$. Note that f_m is continuous for every m and that $F|_{L_0^{2,1}(H)}$ is injective. We recall that Schilder's large deviation principle holds for $\{\mu_\varepsilon\}_{\varepsilon>0}$, that is, for any measurable set $K \subset P(B)$,

$$\begin{aligned} -\inf_{w\in K^\circ} I_0(w) &\leqslant \liminf_{\varepsilon\searrow 0} \varepsilon^2 \log \mu_\varepsilon(K) \\ &\leqslant \limsup_{\varepsilon\searrow 0} \varepsilon^2 \log \mu_\varepsilon(K) \leqslant -\inf_{w\in\overline{K}} I_0(w), \end{aligned}$$

where the good rate function $I_0 : P(B) \to [0, \infty]$ is given by $I_0(w) := \frac{1}{2}\|w\|^2_{L_0^{2,1}(H)}$.

Hence, in order to prove Theorem 3.2, we are left to show that, for any $\delta > 0$,

$$\lim_{m\to\infty} \limsup_{\varepsilon\to 0} \varepsilon^2 \log \mathbb{P}(d_p(F(\varepsilon w(m)), F(\varepsilon w)) > \delta) = -\infty \tag{3.1}$$

and, for any $\beta > 0$,

$$\lim_{m\to\infty} \sup\left\{ d_p(F(x(m)), F(x)) \;\middle|\; \int_0^1 |\dot{x}_t|_H^2 \, dt \leqslant \beta \right\} = 0. \tag{3.2}$$

To prove (3.1) and (3.2), we recall a useful bound on the p-variation distance. For B-valued rough paths X, Y of order 2 and $\gamma > p - 1$, we agree to denote, for $j = 1, 2$,

$$D_{j,p}(X, Y) = D_{j,p}(X^j, Y^j) = \left(\sum_{n=1}^{\infty} n^\gamma \sum_{l=1}^{2^n} |X^j_{t^n_{l-1}, t^n_l} - Y^j_{t^n_{l-1}, t^n_l}|^{p/j}_{B^{\otimes j}} \right)^{j/p},$$

https://doi.org/10.1112/S0024610706022654 Published online by Cambridge University Press

and we write $D_{j,p}(X)$ for $D_{j,p}(X,0)$. Here the values of $2<p<3$ and $\gamma>p-1$ are fixed. Then, by Lyons and Qian [**14**, Proposition 4.1.1 and Proposition 4.1.2], there exists a constant $C=C_{p,\gamma}>0$ such that

$$d_p(X,Y)\leqslant C\max\{D_{1,p}(X,Y),\, D_{1,p}(X,Y)(D_{1,p}(X)+D_{1,p}(Y)),\, D_{2,p}(X,Y)\} \tag{3.3}$$

holds.

First we prove (3.2). From (3.3) it is sufficient to show that

$$\lim_{m\to\infty}\sup\left\{D_{j,p}(F(x(m)),F(x))\;\middle|\;\int_0^1|\dot{x}_t|_H^2\,dt\leqslant\beta\right\}=0 \tag{3.4}$$

for $j=1,2$. (Note that we can show in a similar way to the case $j=1$ that $D_{1,p}(F(x))+\sup_{m\in\mathbf{N}}D_{1,p}(F(x(m)))<\infty$ holds.) Since H is continuously embedded in B, we see that if $\int_0^1|\dot{x}_t|_H^2\,dt\leqslant\beta$, then there is a constant $C>0$ such that

$$|x_t-x_s|_B\leqslant C|x_t-x_s|_H\leqslant C\beta^{1/2}|t-s|^{1/2}$$

for all $s,t\in[0,1]$. Hence, for $X^1_{s,t}=x_t-x_s$,

$$\begin{aligned}D_{1,p}(X(m)^1,X^1)^p&\leqslant\sum_{n=m+1}^{\infty}n^\gamma\sum_{l=1}^{2^n}|2^{m-n}\Delta^m_{k(n,m,l)}x-\Delta^n_l x|_B^p\\&\leqslant C_{\beta,p}\sum_{n=m+1}^{\infty}n^\gamma\sum_{l=1}^{2^n}(2^{(m-n)p}2^{-mp/2}+2^{-np/2})\\&\leqslant C_{\beta,p}\sum_{n=m+1}^{\infty}n^\gamma 2^{-(p-2)n/2},\end{aligned}$$

where $C_{\beta,p}>0$ is a constant that depends only on β and p. Thus we have obtained (3.4) for $j=1$.

We consider (3.4) for $j=2$. If $\int_0^1|\dot{x}_t|_H^2\,dt\leqslant\beta$, then by (2.4) we have for $n\leqslant m$ and $l=1,2,\dots,2^n$ that

$$\begin{aligned}|X(m+1)^2_{t^n_{l-1},t^n_l}-X(m)^2_{t^n_{l-1},t^n_l}|_{B\otimes B}&\leqslant\sum_{r=2^{m-n}(l-1)+1}^{2^{m-n}l}|\Delta^{m+1}_{2r-1}x|_B|\Delta^{m+1}_{2r}x|_B\\&\leqslant C\sqrt{\beta}2^{-(m+1)/2}\sum_{r=2^{m-n}(l-1)+1}^{2^{m-n}l}|\Delta^{m+1}_{2r-1}x|_B\\&\leqslant C^2\sqrt{\beta}2^{-(m+1)/2}\int_{t^n_{l-1}}^{t^n_l}|\dot{x}_t|_H\,dt\\&\leqslant C^2\beta 2^{-(m+n+1)/2}.\end{aligned}$$

For $n>m$ we obtain from (2.3) that

$$\begin{aligned}&|X(m+1)^2_{t^n_{l-1},t^n_l}-X(m)^2_{t^n_{l-1},t^n_l}|_{B\otimes B}\\&\quad\leqslant 2^{2(m+1-n)-1}|\Delta^{m+1}_{k(n,m+1,l)}x\otimes\Delta^{m+1}_{k(n,m+1,l)}x|_{B\otimes B}\\&\qquad+2^{2(m-n)-1}|\Delta^m_{k(n,m,l)}x\otimes\Delta^m_{k(n,m,l)}x|_{B\otimes B}\\&\quad\leqslant C^2\beta 2^{-n}.\end{aligned}$$

https://doi.org/10.1112/S0024610706022654 Published online by Cambridge University Press

By the same argument as in the finite-dimensional case, we obtain

$$\begin{aligned} &D_{2,p}(X(m+1)^2, X(m)^2)^{p/2} \\ &\quad = \sum_{n=1}^{\infty} n^{\gamma} \sum_{l=1}^{2^n} |X(m+1)^2_{t^n_{l-1},t^n_l} - X(m)^2_{t^n_{l-1},t^n_l}|_B^{p/2} \\ &\quad \leqslant C_p \beta^{p/2} \left(\sum_{n=1}^{m} n^{\gamma} \sum_{l=1}^{2^n} 2^{-p(m+n)/4} + \sum_{n=m+1}^{\infty} n^{\gamma} 2^{-(p-2)n/2} \right) \\ &\quad =: C_p \beta^{p/2} \varepsilon_m, \end{aligned}$$

where $C_p > 0$ is a constant depending only on p.

To show that $\{X(m)^2\}_{m\in\mathbb{N}}$ is a Cauchy sequence, it is sufficient to check that $\sum_{m=1}^{\infty} \varepsilon_m^{2/p} < \infty$. Since $2/p < 1$ and $(a+b)^{2/p} \leqslant a^{2/p} + b^{2/p}$ for all $a, b > 0$, it is enough to verify that $\sum_{m=1}^{\infty} \delta_m < \infty$, where δ_m is given by

$$\delta_m = \sum_{n=1}^{m} n^{2\gamma/p} 2^{2n/p} 2^{-(m+n)/2} + \sum_{n=m+1}^{\infty} n^{2\gamma/p} 2^{-(p-2)n/p} =: J_m^1 + J_m^2.$$

Note that, for any $k > 0$, there exists a constant $C > 0$ such that $n^{2\gamma/p} \leqslant C 2^{kn}$ holds for all n. Hence, choosing $k < (p-2)/p$, we see that

$$J_m^2 \leqslant C \sum_{n=m+1}^{\infty} 2^{n\{k-(p-2)/p\}} \leqslant C_{p,k} 2^{(m+1)\{k-(p-2)/p\}},$$

where $C_{p,k} > 0$ is a constant which depends only on p, k. This implies $\sum_{m=1}^{\infty} J_m^2 < \infty$. On the other hand, $J_m^1 \leqslant C 2^{-m/2} \sum_{n=1}^{m} 2^{n\{(2/p)-(1/2)+k\}}$. Taking k small enough we may assume that $r := (2/p)-(1/2)+k < 1/2$. Hence, $J_m^1 \leqslant C 2^{m\{r-(1/2)\}}$ for some constant $C > 0$. This implies $\sum_{m=1}^{\infty} J_m^1 < \infty$ and, hence, $\sum_{m=1}^{\infty} \delta_m < \infty$.

By the above argument we see that if we set X^2 to be the limit of $X(m)^2$ as $m \to \infty$ in the $p/2$-variational metric, then $(1, X^1, X^2)$ is the canonical rough path above x and is the limit of $(1, X(m)^1, X(m)^2)$, and (3.4) holds for $j = 2$. Thus, we have shown (3.2).

Next we prove (3.1). To do so, it is sufficient to show that there exists a sequence $c(m)$ which satisfies that $c(m) \to 0$ as $m \to \infty$ and that, for every q (or sufficiently large q),

$$\mathbb{E}[D_{1,p}(W(m), W)^q]^{1/q} \leqslant c(m)\sqrt{q} \tag{3.5}$$

and

$$\mathbb{E}[D_{2,p}(W(m), W)^q]^{1/q} \leqslant c(m) q. \tag{3.6}$$

This can be verified as follows. Suppose that (3.5) and (3.6) hold. Then, for $q = \varepsilon^{-2}$ and $j = 1, 2$, we have

$$\mathbb{P}(D_{j,p}(W(m), W) > \delta \varepsilon^{-j}) \leqslant (\delta^{-1}\varepsilon^j)^q c(m)^q q^{jq/2} \leqslant (\delta^{-1} c(m))^q$$

and, hence,

$$\limsup_{\varepsilon \to 0} \varepsilon^2 \log \mathbb{P}(D_{j,p}(W(m), W) > \delta \varepsilon^{-j}) \leqslant \log(\delta^{-1} c(m)).$$

Let $\tilde{P}(B) := \{x \in P(B) \mid D_{1,p}(X^1) < \infty \text{ with } X^1_{s,t} = x_t - x_s\}$. Then the law of the B-valued Brownian motion $(w_t)_{0\leqslant t\leqslant 1}$ is supported in $\tilde{P}(B)$ (see [**12**, pp. 560–563]) and induces a Gaussian measure on $\tilde{P}(B)$. Since $W^1_{s,t} = w_t - w_s$, we may apply

https://doi.org/10.1112/S0024610706022654 Published online by Cambridge University Press

Lemma 3.3 to obtain that $\mathbb{E}[D_{1,p}(W)^q]^{1/q} \leqslant C\sqrt{q}$ for some constant C depending only on p and γ.

Hence, we have for every q that

$$\mathbb{E}[\{D_{1,p}(W(m),W)(D_{1,p}(W(m))+D_{1,p}(W))\}^q]^{1/q} \leqslant 2c(m)(2C+c(m))q,$$

where we have used the triangle inequality for $D_{1,p}$. In a similar way as above, we have

$$\begin{aligned}\limsup_{\varepsilon\to 0}\varepsilon^2 \log \mathbb{P}(D_{1,p}(W(m),W)(D_{1,p}(W(m))+D_{1,p}(W)) > \delta\varepsilon^{-2}) \\ \leqslant \log\{2\delta^{-1}c(m)(2C+c(m))\}.\end{aligned}$$

Thus, we have shown that (3.5) and (3.6) are sufficient for (3.1).

We prove (3.5). In the same way as in the finite-dimensional case (see [**13**, p. 274]), we have for $q > p$ that

$$\mathbb{E}[D_{1,p}(W(m),W)^q] \leqslant A(m,q)\sum_{n=m+1}^{\infty} a_n^{q/p}\sum_{l=1}^{2^n}\mathbb{E}[|2^{m-n}\Delta^m_{k(n,m,l)}w - \Delta^n_l w|_B^q].$$

Here, $\{a_n\}_{n\in\mathbb{N}}$ is a sequence such that, for some $\varepsilon \in (0,(1/2)-(1/p))$, $a_n^{q/p} = 2^{-\varepsilon qn}2^{(q-2)n/2}$ and $A(m,q)^{1/q}$ is a bounded double-sequence in m and q. It is easy to see that

$$\begin{aligned}\mathbb{E}[|2^{m-n}\Delta^m_{k(n,m,l)}w - \Delta^n_l w|_B^q] &\leqslant 2^{q-1}\mathbb{E}[|2^{m-n}\Delta^m_{k(n,m,l)}w|_B^q + |\Delta^n_l w|_B^q] \\ &= 2^{q-1}(2^{(m-2n)q/2}+2^{-nq/2})\int_B |x|_B^q\mu(dx) \\ &\leqslant 2^q 2^{-nq/2}\int_B |x|_B^q\mu(dx).\end{aligned}$$

Hence, we have

$$\begin{aligned}&\mathbb{E}[D_{1,p}(W(m),W)^q]^{1/q} \\ &\quad\leqslant 2A(m,q)^{1/q}\left(\int_B |x|_B^q\mu(dx)\right)^{1/q}\left(\sum_{n=m+1}^{\infty} a_n^{q/p}2^{-(q-2)n/2}\right)^{1/q} \\ &\quad\leqslant C\left(\int_B |x|_B^q\mu(dx)\right)^{1/q}2^{-\varepsilon(m+1)},\end{aligned}$$

where $C > 0$ is a constant independent of m and q. From Lemma 3.3, inequality (3.5) follows.

Now we prove (3.6). In the same way as in the finite-dimensional case ([**13**, pp. 274–275]), we have, for every m and $q > p$,

$$\mathbb{E}[D_{2,p}(W(m),W)^q] \leqslant A(q)\sum_{n=1}^{\infty} a_n^{2q/p}\sum_{l=1}^{2^n}\mathbb{E}[|W(m)^2_{t^n_{l-1},t^n_l} - W^2_{t^n_{l-1},t^n_l}|^q], \tag{3.7}$$

where

$$A(q)^{p/(2q-p)} = \sum_{n=1}^{\infty} 2^n\left(\frac{n^\gamma}{a_n}\right)^{2q/(2q-p)}.$$

Noting (2.5) and that the distribution of $2^{(m+1)/2}\Delta^{m+1}_r w$ is μ, we have from Lemma 3.5 that

$$\mathbb{E}[|W(m+1)^2_{t^n_{l-1},t^n_l} - W(m)^2_{t^n_{l-1},t^n_l}|^q]^{1/q} \leqslant 2^{-(m+1)}Cq2^{(m-n)\alpha} = Cq2^{-1}2^{-(1-\alpha)m-\alpha n}$$

https://doi.org/10.1112/S0024610706022654 Published online by Cambridge University Press

for $q \geqslant 1$, $n \leqslant m$ and $1 \leqslant l \leqslant 2^n$. Note that without loss of generality we may assume $\alpha \in (2/p, 1)$. By using the triangle inequality and Fatou's lemma, we have

$$\begin{aligned}\mathbb{E}[|W(m)^2_{t^n_{l-1},t^n_l} - W^2_{t^n_{l-1},t^n_l}|^q]^{1/q} &\leqslant \liminf_{k\to\infty} \mathbb{E}[|W(m)^2_{t^n_{l-1},t^n_l} - W(m+k)^2_{t^n_{l-1},t^n_l}|^q]^{1/q} \\ &\leqslant \sum_{k=0}^{\infty} Cq2^{-(1-\alpha)(m+k)-\alpha n} \leqslant Cq2^{-(1-\alpha)m-\alpha n}.\end{aligned} \tag{3.8}$$

Note that the right-hand side of (3.8) is dominated by $Cq2^{-n}$ since $n \leqslant m$.

For $n \geqslant m$, we have from (2.3) that

$$\mathbb{E}[|W(m)^2_{t^n_{l-1},t^n_l}|^q]^{1/q} \leqslant Cq2^{-n}.$$

Combining this with (3.8) we have that

$$\mathbb{E}[|W^2_{t^n_{l-1},t^n_l}|^q]^{1/q} + \sup_{m\in\mathbb{N}} \mathbb{E}[|W(m)^2_{t^n_{l-1},t^n_l}|^q]^{1/q} \leqslant Cq2^{-n}$$

for $q \geqslant 1$ and $1 \leqslant l \leqslant 2^n$ and that

$$\mathbb{E}[|W(m)^2_{t^n_{l-1},t^n_l} - W^2_{t^n_{l-1},t^n_l}|^q]^{1/q} \leqslant Cq2^{-n} \tag{3.9}$$

for $n \geqslant m$, $q \geqslant 1$ and $1 \leqslant l \leqslant 2^n$.

By (3.7)–(3.9), we see that

$$\begin{aligned}&\mathbb{E}[D_{2,p}(W(m), W)^q] \\ &\quad\leqslant (Cq)^q A(q)\bigg(\sum_{n=1}^{m} a_n^{2q/p} 2^{n-(1-\alpha)mq-\alpha nq} + \sum_{n=m+1}^{\infty} a_n^{2q/p} 2^{-(q-1)n}\bigg).\end{aligned} \tag{3.10}$$

For sufficiently small $\varepsilon > 0$, we set $a_n^{2q/p} 2^{(1-\alpha q)n} = 2^{-2\varepsilon qn}$. Then,

$$\sum_{n=1}^{m} a_n^{2q/p} 2^{n-(1-\alpha)mq-\alpha nq} \leqslant 2^{-(1-\alpha)mq} \sum_{n=1}^{\infty} 2^{-2\varepsilon qn} \leqslant C2^{-(1-\alpha)mq},$$

where C is a constant independent of m and q. Similarly,

$$\begin{aligned}\sum_{n=m+1}^{\infty} a_n^{2q/p} 2^{-(q-1)n} &= \sum_{n=m+1}^{\infty} 2^{-2\varepsilon qn-(1-\alpha q)n-(q-1)n} \\ &= \sum_{n=m+1}^{\infty} 2^{-nq(2\varepsilon+(1-\alpha))} \leqslant C2^{-mq(2\varepsilon+(1-\alpha))},\end{aligned}$$

where C is a constant independent of m and q.

We complete the proof of (3.6) if we show that $A(q)^{1/q}$ is bounded uniformly in q. First note that, for any small $\varepsilon' > 0$, there is a constant $C = C_{\varepsilon'}$ such that, for all n, $n^{2\gamma q/(2q-p)} \leqslant C2^{\varepsilon' n}$ holds. Then,

$$\begin{aligned}A(q)^{p/(2q-p)} &= \sum_{n=1}^{\infty} 2^n n^{2\gamma q/(2q-p)} 2^{-[np/(2q-p)](\alpha q-1-2\varepsilon q)} \\ &\leqslant C\sum_{n=1}^{\infty} 2^{[nq/(2q-p)](2+2\varepsilon'-p\alpha+2\varepsilon p)} 2^{-np\varepsilon'/(2q-p)} \\ &\leqslant C\sum_{n=1}^{\infty} 2^{[nq/(2q-p)](2+2\varepsilon'-p\alpha+2\varepsilon p)}.\end{aligned}$$

https://doi.org/10.1112/S0024610706022654 Published online by Cambridge University Press

Take $\varepsilon \leqslant (p\alpha - 2)/8p$ and $\varepsilon' \leqslant (p\alpha - 2)/8$. Then,

$$A(q)^{p/(2q-p)} \leqslant C \sum_{n=1}^{\infty} 2^{-[(p\alpha-2)q/2(2q-p)]n} \leqslant C \sum_{n=1}^{\infty} 2^{-cn},$$

where c and C are positive constants independent of q. This shows that $A(q)^{p/(2q-p)}$ is bounded in q. Since $(2q-p)/pq$ is bounded in q, $A(q)^{1/q}$ is also bounded in q. Thus, we have shown (3.6) and have completed the proof of Theorem 3.2. □

In the following we extend Theorem 3.2 so that we can treat differential equations with a drift term.

THEOREM 3.7. *Let us keep the same assumption as in Theorem* 3.2. *Furthermore, we set for* $\varepsilon \geqslant 0$, $\lambda_t^\varepsilon = (\varepsilon^2 t, t)$. *Note that* $\lambda^\varepsilon \in \mathrm{BV}(\mathbb{R}^2)$ *for every* $\varepsilon \geqslant 0$. *We denote* $(\varepsilon W, \Lambda^{(\varepsilon)}) := \iota(\varepsilon W, \lambda^\varepsilon)$, *where* $\varepsilon W = (1, \varepsilon W^1, \varepsilon^2 W^2)$ *and* $\iota : G\Omega_p(B) \times \mathrm{BV}(\mathbb{R}^2) \to G\Omega_p(B \oplus \mathbb{R}^2)$ *is the map defined in Lemma* 2.2. *Then, a large deviation principle for* $(\varepsilon W, \Lambda^{(\varepsilon)})$, $\varepsilon > 0$, *on* $G\Omega_p(B \oplus \mathbb{R}^2)$ *holds with the good rate function* I, *where*

$$I(X) = \begin{cases} \frac{1}{2}\|x\|^2_{L_0^{2,1}(H)} & \text{if } X = F((x, \lambda^0)) \text{ for some } x \in L_0^{2,1}(H), \\ \infty & \text{otherwise.} \end{cases}$$

Proof. This is a direct consequence of Lemma 2.2, Theorem 3.2 and Lemma 3.9 below. □

REMARK 3.8. Strictly speaking, we should replace $\mathrm{BV}(\mathbb{R}^2)$ with the closure of $L_0^{2,1}(\mathbb{R}^2)$ with respect to the total variation norm. Obviously, this subset is complete and separable.

LEMMA 3.9. *Let* $\mathcal{X}$ *and* $\mathcal{Y}$ *be complete separable metric spaces. Let* $\{\nu_\varepsilon\}_{\varepsilon>0}$ *be a family of probability measures on* $\mathcal{X}$ *which satisfies a large deviation principle with the good rate function* I. *Let* $\varepsilon \in [0, \infty) \mapsto a(\varepsilon) \in \mathcal{Y}$ *be a continuous curve in* $\mathcal{Y}$. *We write* $a(0) = a$ *for simplicity.*

Then, $\{\nu_\varepsilon \otimes \delta_{a(\varepsilon)}\}_{\varepsilon>0}$ *satisfies a large deviation principle on the product space* $\mathcal{X} \times \mathcal{Y}$ *with the good rate function* $\tilde{I}(x, y) = I(x) + J(y)$. *Here,* δ_y *denotes the point mass at* $y \in \mathcal{Y}$ *and*

$$J(y) := \begin{cases} 0 & \text{if } y = a, \\ \infty & \text{otherwise.} \end{cases}$$

Proof. It is obvious that $\{\delta_{a(\varepsilon)}\}_{\varepsilon>0}$ satisfies a large deviation principle with the good rate function J.

First we take an arbitrary open set $O \subset \mathcal{X} \times \mathcal{Y}$ and set $c := -\inf\{\tilde{I}(x, y) \mid (x, y) \in O\}$. We may assume $c \neq -\infty$. For any $n \in \mathbb{N}$, there exist $(x_n, y_n) \in O$ such that $-\tilde{I}(x_n, y_n) \geqslant c - 1/n$. By the definition of the product topology, there exist open

https://doi.org/10.1112/S0024610706022654 Published online by Cambridge University Press

sets $U_n^1 \subset \mathcal{X}$ and $U_n^2 \subset \mathcal{Y}$ such that $(x_n, y_n) \in U_n^1 \times U_n^2 \subset O$. Then, we have

$$\begin{aligned}\liminf_{\varepsilon \searrow 0} \varepsilon^2 \log(\nu_\varepsilon \otimes \delta_{a(\varepsilon)})(O) &\geqslant \liminf_{\varepsilon \searrow 0} \varepsilon^2 \log(\nu_\varepsilon \otimes \delta_{a(\varepsilon)})(U_n^1 \times U_n^2) \\ &\geqslant \liminf_{\varepsilon \searrow 0} \varepsilon^2 \log \nu_\varepsilon(U_n^1) + \liminf_{\varepsilon \searrow 0} \varepsilon^2 \log \delta_{a(\varepsilon)}(U_n^2) \\ &\geqslant -\inf\{I(x) \mid x \in U_n^1\} - \inf\{J(y) \mid y \in U_n^2\} \\ &\geqslant -I(x_n) - J(y_n) \geqslant c - 1/n.\end{aligned}$$

By letting $n \to \infty$, we have $\liminf_{\varepsilon \searrow 0} \varepsilon^2 \log(\nu_\varepsilon \otimes \delta_{a(\varepsilon)})(O) \geqslant -\inf\{\tilde{I}(x, y) \mid (x, y) \in O\}$.

Next we take an arbitrary closed set $F \subset \mathcal{X} \times \mathcal{Y}$ and set $A_r := \{a(\varepsilon) \mid 0 \leqslant \varepsilon \leqslant r\}$ and $F_r := (\mathcal{X} \times A_r) \cap F$ for $r > 0$. Obviously F_r is closed. Since A_r is compact, $\tilde{F}_r := \pi_1(F_r)$ is also closed, where $\pi_1 : \mathcal{X} \times \mathcal{Y} \to \mathcal{X}$ is the canonical projection. Then, for any fixed $r > 0$, we have

$$\begin{aligned}\limsup_{\varepsilon \searrow 0} \varepsilon^2 \log(\nu_\varepsilon \otimes \delta_{a(\varepsilon)})(F) &= \limsup_{\varepsilon \searrow 0} \varepsilon^2 \log(\nu_\varepsilon \otimes \delta_{a(\varepsilon)})(F_r) \\ &\leqslant \limsup_{\varepsilon \searrow 0} \varepsilon^2 \log(\nu_\varepsilon \otimes \delta_{a(\varepsilon)})(\tilde{F}_r \times A_r) \\ &= \limsup_{\varepsilon \searrow 0} \varepsilon^2 \log \nu_\varepsilon(\tilde{F}_r) \leqslant -\inf\{I(x) \mid x \in \tilde{F}_r\}.\end{aligned}$$

Now we set $c(r) := \inf\{I(x) \mid x \in \tilde{F}_r\}$. It is sufficient to show that

$$c_0 := \limsup_{r \searrow 0} c(r) \geqslant \inf\{I(x) \mid x \in \tilde{F}_0\}.$$

We may assume $c_0 < \infty$. There exists a sequence $r_n \searrow 0$ such that $c(r_n) \to c_0$. For every $n \in \mathbb{N}$, there exists $(x_n, a(s_n)) \in F_{r_n}$ such that $I(x_n) \leqslant c(r_n) + 1/n$. Since x_n is an element of a compact set $\{x \in \mathcal{X} \mid I(x) \leqslant c_0 + 1\}$, there exists a subsequence $\{x_{n(k)}\}_{k \in \mathbb{N}}$ which converges to an element $x_\infty \in \mathcal{X}$. Hence, $(x_{n(k)}, a(s_{n(k)})) \to (x_\infty, a)$ as $n \to \infty$. Since F is closed, we see that $(x_\infty, a) \in F_0$ and $x_\infty \in \tilde{F}_0$. Then,

$$\begin{aligned}\inf\{I(x) \mid x \in \tilde{F}_0\} \leqslant I(x_\infty) &\leqslant \liminf_{k \to \infty} I(x_{n(k)}) \\ &\leqslant \liminf_{k \to \infty} (c(r_{n(k)}) + 1/n(k)) \\ &= \lim_{k \to \infty} (c(r_{n(k)}) + 1/n(k)) = c_0.\end{aligned}$$

Thus, by noting $\inf\{I(x) \mid x \in \tilde{F}_0\} = \inf\{\tilde{I}(x, y) \mid (x, y) \in F\}$, we have obtained

$$\limsup_{\varepsilon \searrow 0} \varepsilon^2 \log(\nu_\varepsilon \otimes \delta_{a(\varepsilon)})(F) \leqslant -\inf\{\tilde{I}(x, y) \mid (x, y) \in F\}.$$

This completes the proof. □

4. *Large deviation for heat kernel measures*

In this section we introduce heat kernel measures and show that the rough path theory is applicable to them. Let $s > 1/2$ and $\mathcal{L}_0(\mathbb{R}^r) := \{x \in C([0, 1], \mathbb{R}^r) \mid x(0) = x(1) = 0\}$. For $h \in \mathcal{L}_0(\mathbb{R}^r)$ of the form

$$h(\tau) = \sum_{n \neq 0} \hat{h}(n)(e^{2\pi\sqrt{-1}n\tau} - 1),$$

https://doi.org/10.1112/S0024610706022654 Published online by Cambridge University Press

we set

$$\|h\|^2_{H_0^s(\mathbb{R}^r)} := \sum_{n\neq 0} |2\pi n|^{2s} |\hat{h}(n)|^2$$

and $H_0^s(\mathbb{R}^r) := \{h \in \mathcal{L}_0(\mathbb{R}^r) \mid \|h\|_{H_0^s(\mathbb{R}^r)} < \infty\}$. It is well-known that $H_0^s(\mathbb{R}^r)$ is a Hilbert space embedded in $\mathcal{L}_0(\mathbb{R}^r)$ and that there exists a Gaussian measure μ^s such that the triplet $(\mathcal{L}_0(\mathbb{R}^r), H_0^s(\mathbb{R}^r), \mu^s)$ becomes an abstract Wiener space. (For a proof, see Inahama [**8**, Section 4] and see also [**10**] for the pinned case. Inequalities such as (4.1) and (4.2) below were essentially obtained in, [**9**, Section 3]. It easily follows from those that $H_0^s(\mathbb{R}^r)$-Gaussian measure μ^s is supported in $\mathcal{L}_0(\mathbb{R}^r)$.) When $s = 1$, μ^1 is the usual r-dimensional pinned Wiener measure and is of particular importance. For $\tau \in [0,1]$ and $j = 1, 2, \dots, r$, we denote by δ_τ^j the element in $\mathcal{L}_0(\mathbb{R}^r)^*$ defined by $\langle \delta_\tau^j, x\rangle = x^j(\tau)$ and set $x(\tau) := (x^1(\tau), \dots, x^r(\tau))$. Let $(w_t)_{t\geqslant 0}$ be a $\mathcal{L}_0(\mathbb{R}^r)$-valued Brownian motion associated with μ^s. We set $w_t^j(\tau) := \langle \delta_\tau^j, w_t\rangle$ and $w_t(\tau) := (w_t^1(\tau), \dots, w_t^r(\tau))$.

LEMMA 4.1. *Let the notation be as above and $s > 1/2$. The projective tensor product $\mathcal{L}_0(\mathbb{R}^r) \otimes_{\mathrm{pr}} \mathcal{L}_0(\mathbb{R}^r)$ is exact with respect to μ^s, that is, it satisfies Assumption* A.

Proof. It is easy to see that $\delta_\tau^j \in \mathcal{L}_0(\mathbb{R}^r)^* \subset H_0^s(\mathbb{R}^r)^* \simeq H_0^s(\mathbb{R}^r)$ and that, as an element in $H_0^s(\mathbb{R}^r)$, δ_τ^j is written as follows:

$$\delta_\tau^j(\eta) = \sum_{n\neq 0} |2\pi n|^{-2s} (e^{-2\pi\sqrt{-1}n\tau} - 1)(e^{2\pi\sqrt{-1}n\eta} - 1)e_j, \quad \eta \in [0,1],$$

where $\{e_j\}_{j=1}^r$ is the natural orthonormal basis of $\mathbb{R}^r$. By a straightforward computation, we see that

$$\begin{aligned}
\|\delta_\tau^j - \delta_\sigma^j\|^2_{H_0^s(\mathbb{R}^r)} &= \sum_{n\neq 0} |2\pi n|^{-2s} |e^{-2\pi\sqrt{-1}n\tau} - e^{-2\pi\sqrt{-1}n\sigma}|^2 \\
&\leqslant 2^{2-\varepsilon} \sum_{n\neq 0} |2\pi n|^{-2s} |e^{-2\pi\sqrt{-1}n\tau} - e^{-2\pi\sqrt{-1}n\sigma}|^\varepsilon \\
&\leqslant C \sum_{n\neq 0} |2\pi n|^{-2s+\varepsilon} |\tau - \sigma|^\varepsilon \leqslant C|\tau - \sigma|^\varepsilon, \qquad (4.1)
\end{aligned}$$

where $\varepsilon > 0$ is taken so small that $2s - \varepsilon > 1$ and $\varepsilon \leqslant 1$.

Since the law of $x^j(\tau) - x^j(\sigma)$ is one-dimensional Gaussian measure with variance $\|\delta_\tau^j - \delta_\sigma^j\|^2_{H_0^s(\mathbb{R}^r)}$, there exists a constant $C_{p,r} > 0$ such that

$$\begin{aligned}
\mathbb{E}^{\mu^s}[|x(\tau) - x(\sigma)|^p_{\mathbb{R}^r}] &\leqslant C_{p,r} \mathbb{E}^{\mu^s}[|x^j(\tau) - x^j(\sigma)|^p] \\
&\leqslant C_{p,r} |\tau - \sigma|^{p\varepsilon/2} \qquad (4.2)
\end{aligned}$$

holds for all $\tau, \sigma \in [0,1]$. Here, the positive constant p is chosen so large that $p\varepsilon/2 > 1$. Then, by [**14**, Theorem 4.6.8], (4.2) implies that any Banach tensor product on $\mathcal{L}_0(\mathbb{R}^r) \otimes \mathcal{L}_0(\mathbb{R}^r)$ is exact with respect to μ^s. □

Now we are in a position to define heat kernel measures in a slightly generalized form. Let

$$A_j(x) = \sum_{i=1}^{r'} a_{ij}(x) \frac{\partial}{\partial x_i}, \quad A_0(x) = \sum_{i=1}^{r'} b_i(x) \frac{\partial}{\partial x_i}, \quad \tilde{A}_0(x) = \sum_{i=1}^{r'} \beta_i(x) \frac{\partial}{\partial x_i}$$

https://doi.org/10.1112/S0024610706022654 Published online by Cambridge University Press

be vector fields on $\mathbb{R}^{r'}$, $j = 1, \dots, r$. In this paper, we impose the following regularity condition:

$$a_{ij}, b_i, \beta_i \in C_b^4(\mathbb{R}^{r'}, \mathbb{R}) \quad \text{for } 1 \leqslant i \leqslant r', 1 \leqslant j \leqslant r. \tag{4.3}$$

We write σ for the $(r' \times r)$-matrix $\{a_{ij}\}_{1 \leqslant i \leqslant r', 1 \leqslant j \leqslant r}$ and write b and β for the column vectors $(b_1, \dots, b_{r'})^{\mathrm{T}}$ and $(\beta_1, \dots, \beta_{r'})^{\mathrm{T}}$, respectively. Let $\varepsilon > 0$. For each $\tau \in [0, 1]$, we consider the following SDE:

$$\begin{aligned} dX_t^\varepsilon(\tau) &= \sum_{j=1}^{r} A_j(X_t^\varepsilon(\tau)) \circ \varepsilon \, dw_t^j(\tau) + A_0(X_t^\varepsilon(\tau))\varepsilon^2 \, dt + \tilde{A}_0(X_t^\varepsilon(\tau)) \, dt \\ &= \sigma(X_t^\varepsilon(\tau)) \circ \varepsilon \, dw_t(\tau) + b(X_t^\varepsilon(\tau))\varepsilon^2 \, dt + \beta(X_t^\varepsilon(\tau)) \, dt. \end{aligned} \tag{4.4}$$

In this paper, we assume $X_0^\varepsilon(\tau) = 0$ for simplicity. We will often write $X^\varepsilon(t, \tau) := X_t^\varepsilon(\tau)$. In Lemma 4.8 below, we will prove that $X^\varepsilon(t, \tau)$ has a bi-continuous modification. Note that if $\tilde{A}_0 = 0$, then $(X^\varepsilon(t, \cdot))_{t \geqslant 0}$ and $(X^1(\varepsilon^2 t, \cdot))_{t \geqslant 0}$ have the same law.

DEFINITION 4.2. Let $r, r' \in \mathbb{N}$, $s > 1/2$, $\varepsilon > 0$ and vector fields $\{A_j\}_{j=0}^r \cup \{\tilde{A}_0\}$ be as above.

(i) We denote by $\mathcal{V}_\varepsilon$ the law of the heat process X^ε. This measure is supported in $P(\mathcal{L}_0(\mathbb{R}^{r'}))$.

(ii) Consider the case $\tilde{A}_0 = 0$. We define the heat kernel measure ν_t to be the law of $X^1(t, \cdot)$. This measure ν_t is supported in $\mathcal{L}_0(\mathbb{R}^{r'})$.

EXAMPLE 4.3. Let G be a compact connected Lie group of dimension r. We assume that its Lie algebra $\mathfrak{g}$ is equipped with adjoint invariant inner product. By the left translation we define a Riemannian structure on G. Let $\{A_j\}_{j=1}^r$ be an orthonormal basis of $\mathfrak{g}$. ($A_0 = \tilde{A}_0 = 0$.) By the left translation, elements of $\mathfrak{g}$ are identified with left invariant vector fields on G. Embed G into a Euclidean space of sufficiently high dimension so that unit element e is mapped to 0 and then extend A_j so that the coefficients satisfy our assumption. In this case ν_t is clearly supported in the loop group $\mathcal{L}_e(G)$. This measure was introduced by Malliavin [**15**] and is called the heat kernel measure on $\mathcal{L}_e(G)$. (In particular, the case $s = 1$ has been extensively studied. For example, Fang and Zhang's result on the large deviation principle in [**5**] is for the case $s = 1$.)

EXAMPLE 4.4. Let M be a compact connected Riemannian manifold of dimension r with an initial point m_0. We denote by $\pi : O(M) \to M$ its orthonormal frame bundle with initial point u_0. The point u_0 is an element in the fiber of m_0. We consider the Levi-Civita connection on M. Then, the canonical horizontal vector fields $\{A_j\}_{1 \leqslant j \leqslant r}$ on $O(M)$ are defined. (We set $A_0 = \tilde{A}_0 = 0$.) By embedding $O(M)$ in a Euclidean space, we can consider the heat process $X(t, \tau)$ with the initial condition $X(0, \tau) = u_0$ in the same way as in Example 4.3. We set $\tilde{\nu}_t$ to be the law of $\pi(X(t, \cdot))$ and call it the heat kernel measure on M. The measure $\tilde{\nu}_t$ is supported in the loop space $\mathcal{L}_{m_0}(M)$. It is easy to see that, when M is a Lie group as in Example 4.3, this definition of heat kernel measure $\tilde{\nu}_t$ coincides with the one in Example 4.3.

https://doi.org/10.1112/S0024610706022654 Published online by Cambridge University Press

EXAMPLE 4.5. In [**5**], Fang and Zhang give another definition of heat kernel measures on a compact connected Riemannian manifold, by using Nash's embedding and Elworthy's gradient systems (see [**5**, pp. 478–482]). Our Definition 4.2 covers this case, too. However, at this moment, the authors do not know whether the definitions in Examples 4.4 and 4.5 are the same or not.

EXAMPLE 4.6. If $\tilde{A}_0 = 0$, then X^1 is the process constructed in Brzeźniak and Elworthy [**2**, Section 6]. More precisely, they also treated the case that the coefficients are unbounded, while we must assume boundedness in order to use the rough path theory. The process in [**2**] takes its values in a Sobolev–Slobodetski space, which is embedded in $\mathcal{L}_0(\mathbb{R}^{r'})$.

In the following we show that heat process X^ε defined by (4.4) is a typical example to which the Banach space-valued rough path theory can be applied.

Before doing so we set some notation. In the following let V_i denote a Banach space ($i \in \mathbb{N}$). We denote by $L(V_1, V_2)$ the set of bounded linear maps from V_1 to V_2 and by $L^k(V_1, \dots, V_k; V_{k+1})$ the set of bounded k-multi-linear maps from $V_1 \times \cdots \times V_k$ to V_{k+1}. As usual we identify $L(V_1, L^k(V_1, V_2))$ with $L^{k+1}(V_1, V_2)$, etc. Let $\mathcal{O}$ be an open set in V_1 and $f : \mathcal{O} \to V_2$ be a map. We say f is Fréchet differentiable at $y \in \mathcal{O}$ if there exists $Df(y) \in L(V_1, V_2)$ such that

$$\lim_{x \to 0} \frac{|f(y+x) - f(y) - Df(y)x|_{V_2}}{|x|_{V_1}} = 0.$$

If f is Fréchet differentiable at all $y \in \mathcal{O}$, then Df is a map from $\mathcal{O}$ to $L(V_1, V_2)$. If $f : \mathcal{O} \to V_2$ is sufficiently smooth in the sense of Fréchet, then $D^k f$ is a map from $\mathcal{O}$ to $L^k(V_1, \dots, V_1; V_2)$.

In the following, we set $V_1 = \mathcal{L}_0(\mathbb{R}^r) \oplus \mathbb{R}^2$ and $V_2 = \mathcal{L}_0(\mathbb{R}^{r'})$ and consider a map $f : V_2 \to L(V_1, V_2)$ defined by

$$(f(y)(x,u))(\tau) := \sigma(y(\tau))x(\tau) + b(y(\tau))u_1 + \beta(y(\tau))u_2$$

for $(x, u) \in V_1$ with $u = (u_1, u_2)$, $y \in V_2$ and $\tau \in [0, 1]$. Then we have the following.

LEMMA 4.7. *Let $f : V_2 \to L(V_1, V_2)$ be as above. Then, f is three times Fréchet differentiable. Moreover, there exists a constant $M > 0$ such that*

$$|D^k f(y)| \leqslant M \tag{4.5}$$

and

$$|D^k f(y) - D^k f(y')| \leqslant M|y - y'| \tag{4.6}$$

hold for all $y, y' \in V_2$ and $k = 0, 1, 2, 3$.

Proof. Note that $Df : V_2 \to L(V_2, L(V_1, V_2))$, where $L(V_2, L(V_1, V_2))$ can naturally be identified with $L^2(V_2, V_1; V_2)$ (the set of bounded bilinear maps from $V_2 \times V_1$ to V_2). By straightforward computation, $Df(y)$ can be expressed for each $y \in V_2$ as follows:

$$\begin{aligned}(Df(y)\langle p, (x,u)\rangle_{V_2 \times V_1})(\tau) \\ = \sum_{i=1}^{r'} \sum_{l=1}^{r'} \left(\sum_{j=1}^{r} \partial_l a_{ij}(y(\tau))x^j(\tau) + \partial_l b_i(y(\tau))u_1 + \partial_l \beta_i(y(\tau))u_2 \right) p^l(\tau)\tilde{e}_i,\end{aligned}$$

https://doi.org/10.1112/S0024610706022654 Published online by Cambridge University Press

where $p = (p^1, \dots, p^{r'}) \in V_2$, $(x, u) = (x^1, \dots, x^r; u_1, u_2) \in V_1$ and $\{\tilde{e}_i\}_{i=1}^{r'}$ is the natural orthonormal basis of $\mathbb{R}^{r'}$. It is obvious from the boundedness of the first derivatives of a_{ij}, b_i and β_i that (4.5) holds for $k = 1$. It is also easy to see from the boundedness of the second derivatives of a_{ij}, b_i and β_i that (4.6) holds for $k = 1$.

We omit a proof for the cases $k = 0, 2, 3$, since we can show them in a similar way. Note that we use the condition (4.3) on the fourth derivative of a_{ij} , b_i and β_i to prove (4.6) for $k = 3$. □

Now we are in a position to apply the rough path theory in Lyons and Qian [**14**, Chapter 6] to our case. By Lemmas 4.1 and 4.7, we may use the theory of the Itô maps in the sense of the rough path theory. Consider the following differential equation in the sense of the rough path theory:

$$dY_t = f(Y_t)\, dX_t, \quad Y_0 = 0. \tag{4.7}$$

Then, by [**14**, Theorem 6.2.1 and Corollary 6.2.2], for any $X \in G\Omega_p(V_1)$ there exists $Y \in G\Omega_p(V_2)$ which uniquely solves the equation (4.7). By [**14**, Theorem 6.2.2], the Itô map $\phi : G\Omega_p(V_1) \to G\Omega_p(V_2)$, which is defined by $Y = \phi(X)$, is continuous. (Note that we consider the projective tensor products of V_1, V_2 and $\mathcal{L}_0(\mathbb{R}^r)$.)

LEMMA 4.8. *Let $\varepsilon > 0$ and W be the Brownian rough path in $G\Omega_p(\mathcal{L}_0(\mathbb{R}^r))$. Then, $Y^{(\varepsilon)} = \phi(\varepsilon W, \Lambda^{(\varepsilon)}) = \phi(\iota(\varepsilon W, \lambda^\varepsilon))$ is a random element in $G\Omega_p(V_2)$ and $(t, \tau) \mapsto Y_{0,t}^{(\varepsilon),1}(\tau)$ is a bi-continuous modification of the two-parameter process $(X^\varepsilon(t, \tau))_{0 \leqslant t \leqslant 1, 0 \leqslant \tau \leqslant 1}$ defined in* (4.4).

Proof. It is sufficient to show that, for each fixed $\varepsilon > 0$ and $\tau \in [0, 1]$,

$$\mathbb{P}(Y_{0,t}^{(\varepsilon),1}(\tau) = X^\varepsilon(t, \tau) \text{ for } 0 \leqslant t \leqslant 1) = 1$$

holds. For $m \in \mathbb{N}$, we denote by $(z_t^{\varepsilon,m})_{0 \leqslant t \leqslant 1}$ a V_2-valued continuous path which solves the following integral equation (in the usual sense):

$$z_t^{\varepsilon,m} = \int_0^t f(z_u^{\varepsilon,m}) d(\varepsilon w(m)_u, \lambda_u^\varepsilon), \tag{4.8}$$

where the integral on the right-hand side is Riemannian integral. Then, by [**14**, Proposition 6.2.2], $z_t^{\varepsilon,m} = \phi(\iota(\varepsilon W(m), \lambda^\varepsilon))_{0,t}^1$ holds, where ϕ is the Itô map. Let $\delta_\tau : V_2 \to \mathbb{R}^{r'}$ be the evaluation map at τ, that is, $\langle \delta_\tau, x \rangle = x(\tau)$ for $x \in V_2$. Since δ_τ is bounded and linear, we see from (4.8) that

$$\begin{aligned} z_t^{\varepsilon,m}(\tau) &= \langle \delta_\tau, z_t^{\varepsilon,m} \rangle \\ &= \int_0^t \langle \delta_\tau, f(z_u^{\varepsilon,m}) d(\varepsilon w(m)_u, \lambda_u^\varepsilon) \rangle \\ &= \int_0^t \sigma(z_u^{\varepsilon,m}(\tau)) \varepsilon \, dw(m)_u(\tau) + \int_0^t b(z_u^{\varepsilon,m}(\tau)) \varepsilon^2 \, du + \int_0^t \beta(z_u^{\varepsilon,m}(\tau)) \, du. \end{aligned}$$

Here, $w_u(\tau) = (w_u^1(\tau), \dots, w_u^{r'}(\tau))$ and we have used the definition of f. By the Wong–Zakai approximation theorem (for example, see Ikeda and Watanabe

https://doi.org/10.1112/S0024610706022654 Published online by Cambridge University Press

[**7**, Theorem VI-7.2]), a subsequence of $\{(z_t^{\varepsilon,m}(\tau))_{0\leqslant t\leqslant 1}\}_{m\in\mathbb{N}}$ converges to $(X^\varepsilon(t,\tau))_{0\leqslant t\leqslant 1}$ in the sup-norm almost surely. On the other hand, on the set $\{\lim_{m\to\infty} W(m) = W \text{ in } G\Omega_p(V_1)\}$, we have that, for all $t\in[0,1]$,

$$\lim_{m\to\infty} z_t^{\varepsilon,m}(\tau) = \lim_{m\to\infty} \langle \delta_\tau, \phi(\iota(\varepsilon W(m), \lambda^\varepsilon))_{0,t}^1\rangle = \langle \delta_\tau, \phi(\iota(\varepsilon W, \lambda^\varepsilon))_{0,t}^1\rangle = Y_{0,t}^{(\varepsilon),1}(\tau).$$

Here, we have used the continuity of ϕ. □

The following theorem is the main result is this paper. It reproves and generalizes the large deviation principle for heat processes (and heat kernel measures) on loop groups in Fang and Zhang [**5**].

THEOREM 4.9. *Let $r, r' \in \mathbb{N}$, $s > 1/2$ and $\varepsilon > 0$, and let the vector fields $\{A_j\}_{j=0}^r \cup \{\tilde{A}_0\}$ be as above. In this theorem we simply write H for $H_0^s(\mathbb{R}^r)$.*

(i) *Let $\mathcal{V}_\varepsilon$ be the law of the process X^ε as in Definition* 4.2. *Then, $\mathcal{V}_\varepsilon$ satisfies the large deviation principle as $\varepsilon \searrow 0$ with the good rate function I_1, where*

$$I_1(\xi) = \begin{cases} \frac{1}{2}\inf\{\|x\|^2_{L_0^{2,1}(H)} \mid \phi(\iota(F(x),\lambda^0))^1 = \xi\} & \text{if } \xi_\cdot = \phi(\iota(F(x),\lambda^0))^1_{0,\cdot} \\ & \text{for some } x \in L_0^{2,1}(H), \\ \infty & \text{otherwise.} \end{cases}$$

More precisely, for any measurable set $K \subset P(\mathcal{L}_0(\mathbb{R}^{r'}))$, the following inequalities hold:

$$-\inf_{\xi\in K^\circ} I_1(\xi) \leqslant \liminf_{\varepsilon\searrow 0} \varepsilon^2 \log \mathcal{V}_\varepsilon(K) \leqslant \limsup_{\varepsilon\searrow 0} \varepsilon^2 \log \mathcal{V}_\varepsilon(K) \leqslant -\inf_{\xi\in\overline{K}} I_1(\xi).$$

(ii) *We consider the case $\tilde{A}_0 = 0$. Let ν_t be the heat kernel measure as in Definition* 4.2. *Then, ν_t satisfies the large deviation principle as $t \searrow 0$ with the good rate function I_2, where*

$$I_2(y) = \begin{cases} \frac{1}{2}\inf\{\|x\|^2_{L_0^{2,1}(H)} \mid \phi(\iota(F(x),\lambda^0))^1_{0,1} = y\} & \text{if } y = \phi(\iota(F(x),\lambda^0))^1_{0,1} \\ & \text{for some } x \in L_0^{2,1}(H), \\ \infty & \text{otherwise.} \end{cases}$$

More precisely, for any measurable set $K \subset \mathcal{L}_0(\mathbb{R}^{r'})$, the following inequalities hold:

$$-\inf_{y\in K^\circ} I_2(y) \leqslant \liminf_{t\searrow 0} t\log \nu_t(K) \leqslant \limsup_{t\searrow 0} t\log \nu_t(K) \leqslant -\inf_{y\in\overline{K}} I_2(y).$$

Proof. By (4.4) and Lemma 4.8, $\mathcal{V}_\varepsilon$ is equal to the law of $(\phi(\iota(F(\varepsilon w), \lambda^\varepsilon))_{0,t}^1)_{0\leqslant t\leqslant 1}$. Since ϕ is continuous, we may use the contraction principle of large deviations. From Theorem 3.7, the first item follows.

The second item follows directly from the first one and the fact that $(X^\varepsilon(t,\cdot))_{t\geqslant 0}$ and $(X^1(\varepsilon^2 t,\cdot))_{t\geqslant 0}$ have the same law since the evaluation at time 1 is a continuous map. □

REMARK 4.10. In this remark we explain why we do not work on the Sobolev–Slobodetski space $W^{\theta,p}(S^1)$. (See Brzeźniak and Elworthy [**2**] for the precise definition.) Note that $W^{\theta,p}(S^1)$ is an M-type 2 Banach space and is continuously embedded in the space of continuous loops if $p > 1$ and $1/p < \theta < 1$ (see [**2**, Section 5]).

https://doi.org/10.1112/S0024610706022654 Published online by Cambridge University Press

In [**2**, Section 6] and [**1**, Section 4] it is proved that $(X_t)_{t\geqslant 0}$ defined by the SDE (4.4) can be considered as the solution of a $W^{\theta,p}(S^1)$-valued SDE if $p > 2$ and $1/p < \theta < 1/2$ (at least for the case $s = 1$).

For general $s > 1/2$, under a suitable condition on p and θ, we see by straightforward computation that the Gaussian measure μ_s is supported on $W^{\theta,p}(S^1)$ and see from inequality (6.8) in [**12**, p. 575] that $W^{\theta,p}(S^1) \otimes_{\mathrm{pr}} W^{\theta,p}(S^1)$ is exact with respect to μ_s. Moreover, Brzeźniak and Caroll proved the Wong–Zakai theorem for SDEs on M-type 2 Banach spaces ([**1**, Theorem 3]).

Hence, one may wonder why we do not work on the Sobolev–Slobodetski space $W^{\theta,p}(S^1)$. The main reason why we avoided $W^{\theta,p}(S^1)$ is that the map f in Lemma 4.7, which is called the Nemytski operator, is not bounded with respect to the topology of $W^{\theta,p}(S^1)$ (see [**2**, Section 5]). Since we would not like to treat the Itô maps or ordinary differential equations with unbounded coefficients, we choose to work on the space of continuous loops.

References

1. Z. Brzeźniak and A. Carroll, 'Approximations of the Wong–Zakai type for stochastic differential equations in M-type 2 Banach spaces with applications to loop spaces', *Séminaire de Probabilités XXXVII*, Lecture Notes in Mathematics 1832 (Springer, Berlin, 2003) 251–289.

2. Z. Brzeźniak and K. D. Elworthy, 'Stochastic differential equations on Banach manifolds', *Methods Funct. Anal. Topology* 6 (2000) 43–84.

3. A. Dembo and O. Zeitouni, *Large deviations techniques and applications*, 2nd edn, Applications of Mathematics (New York) 38 (Springer, New York, 1998).

4. J. D. Deuschel and D. W. Stroock, *Large deviations* (Academic Press, Boston, 1989).

5. S. Fang and T. Zhang, 'Large deviations for the Brownian motion on loop groups', *J. Theoret. Probab.* 14 (2001) 463–483.

6. L. Gross, 'Abstract Wiener spaces', Proc. Fifth Berkeley Sympos. Math. Statist. and Probability (Berkeley, CA, 1965/66), vol. II: *Contributions to Probability Theory*, part 1 (University of California, Berkeley, CA, 1997) 31–42.

7. N. Ikeda and S. Watanabe, *Stochastic differential equations and diffusion processes*, 2nd edn, North-Holland Mathematical Library 24 (North-Holland, Amsterdam and Kodansha, Tokyo, 1989).

8. Y. Inahama, 'Logarithmic Sobolev inequality on free loop groups for heat kernel measures associated with the general Sobolev spaces', *J. Funct. Anal.* 179 (2001) 170–213.

9. Y. Inahama, 'Convergence of finite dimensional distributions of heat kernel measures on loop groups', *J. Funct. Anal.* 198 (2003) 311–340.

10. Y. Inahama, 'Logarithmic Sobolev inequality for H_0^s-metric on pinned loop groups', *Infin. Dimens. Anal. Quantum Probab. Relat. Top.* 7 (2004) 1–26.

11. R. Léandre, 'Large deviation for non-linear random fields', *Nonlinear Phenom. Complex Syst.* 5 (2002) 314–317.

12. M. Ledoux, T. Lyons and Z. Qian, 'Lévy area of Wiener processes in Banach spaces', *Ann. Probab.* 30 (2002) 546–578.

13. M. Ledoux, Z. Qian and T. Zhang, 'Large deviations and support theorem for diffusion processes via rough paths', *Stochastic Process. Appl.* 102 (2002) 265–283.

14. T. Lyons and Z. Qian, *System control and rough paths* (Oxford University Press, Oxford, 2002).

15. P. Malliavin, 'Hypoellipticity in infinite dimensions', *Diffusion processes and related problems in analysis*, vol. I (Evanston, IL, 1989), Progress in Probability 22 (Birkhäuser, Boston, MA, 1990) 17–31.

16. G. Pisier, 'Probabilistic methods in the geometry of Banach spaces', *Probability and analysis* (Varenna, 1985), Lecture Notes in Mathematics 1206 (Springer, Berlin, 1986) 167–241.

https://doi.org/10.1112/S0024610706022654 Published online by Cambridge University Press

Yuzuru Inahama and Hiroshi Kawabi
Department of Mathematical Science
Graduate School of Engineering Science
Osaka University
Toyonaka
Osaka, 560-8531
Japan

Current addresses:
Yuzuru Inahama
Department of Mathematics
Graduate School of Science and Engineering
Tokyo Institute of Technology
2-12-1, Oh-okayama, Meguro-ku
Tokyo, 152-8551
Japan

inahama@kusm.kyoto-u.ac.jp

Hiroshi Kawabi
Graduate School of Mathematics
Kyushu University
6-10-1, Hakozaki, Higashi-ku
Fukuoka, 812-8581
Japan

kawabi@sigmath.es.osaka-u.ac.jp

https://doi.org/10.1112/S0024610706022654 Published online by Cambridge University Press