Quasi-sure Existence of Gaussian Rough Paths and Large Deviation Principles for Capacities

H. Boedihardjo*, X. Geng† and Z. Qian‡

Abstract

We construct a quasi-sure version (in the sense of Malliavin) of geometric rough paths associated with a Gaussian process with long-time memory. As an application we establish a large deviation principle (LDP) for capacities for such Gaussian rough paths. Together with Lyons' universal limit theorem, our results yield immediately the corresponding results for pathwise solutions to stochastic differential equations driven by such Gaussian process in the sense of rough paths. Moreover, our LDP result implies the result of Yoshida on the LDP for capacities over the abstract Wiener space associated with such Gaussian process.

1 Introduction

The theory of rough paths, established by Lyons in his groundbreaking paper [13], gives us a fundamental way of understanding path integrals along one forms and pathwise solutions to differential equations driven by rough signals. After his work, the study of the (geometric) rough path nature of stochastic processes (e.g. Brownian motion, Markov processes, martingales, Gaussian processes, etc.) becomes rather important, since it will then immediately lead to a pathwise theory of stochastic differential equations driven by such processes, which is one of the central problems in stochastic analysis. The rough path regularity of Brownian motion was first studied in the unpublished Ph.D. thesis of Sipiläinen [22]. Later on Coutin and Qian [3] proved that the sample paths of fractional Brownian motion with Hurst parameter $H > 1/4$ can be lifted as geometric rough paths in a canonical way, and such canonical lifting does not exist when $H \leqslant 1/4$. Of course their result covers the Brownian motion case. The systematic study of stochastic processes as rough paths then appeared in

*Department of Mathematics and Statistics, University of Reading, Oxford RG6 6AX, England. Email: h.s.boedihardjo@reading.ac.uk

†Mathematical Institute, University of Oxford, Oxford OX2 6GG and the Oxford-Man Institute, Eagle House, Walton Well Road, Oxford OX2 6ED, England. Email: xi.geng@maths.ox.ac.uk

‡Exeter College, University of Oxford, Oxford OX1 3DP, England. Email: qianz@maths.ox.ac.uk.

the monographs on rough path theory by Lyons and Qian [15] and by Friz and Victoir [6].

The continuity of the solution map for rough differential equations, which was also proved by Lyons [13] and usually known as the universal limit theorem, is a fundamental result in rough path theory. To some extent it gives us a way of understanding the right topology under which differential equations are stable on rough path space. An easy but important application of the universal limit theorem is large deviation principles (or simply LDPs) for pathwise solutions to stochastic differential equations according to the contraction principle, once the LDP for the law of the driving process as rough paths is established under the rough path topology. This is also the main motivation of strengthening the classical LDPs for probability measures on path space under the uniform topology to the rough path setting. Since the rough path topology is stronger than the uniform topology, a direct corollary is the classical Freidlin-Wentzell theory on path space, which does not follow immediately from the contraction principle and is in fact highly nontrivial as the solution map is not continuous in this case. In the case of Brownian motion, Ledoux, Qian and Zhang first established the LDP for the law of Brownian rough paths. Their result was then extended to the case of fractional Brownian motion by Millet and Sanz-Solé [19]. The general study of LDPs for different stochastic processes in particular for Gaussian processes as rough paths can be found in [6].

We first recall some basic notions from rough path theory which we use throughout the rest of this article. We refer the readers to [6], [14], [15] for a detailed presentation.

For $n \geqslant 1$, let

$$T^{(n)}\left(\mathbb{R}^d\right) = \oplus_{i=0}^n \left(\mathbb{R}^d\right)^{\otimes i}$$

be the truncated tensor algebra over $\mathbb{R}^d$ degree n, where $\left(\mathbb{R}^d\right)^{\otimes 0} := 0$. We use Δ to denote the standard 2-simplex $\{(s,t): \ 0 \leqslant s \leqslant t \leqslant 1\}$.

We call an $\mathbb{R}^d$-valued continuous paths over $[0,1]$ *smooth* if it has bounded total variation. Given a smooth path w, for $k \in \mathbb{N}$ define

$$w_{s,t}^k = \int\limits_{s<t_1<\cdots<t_k<t} dw_{t_1} \otimes \cdots \otimes dw_{t_k}, \ (s,t) \in \Delta. \tag{1.1}$$

From classical integration theory we know that (1.1) is well-defined as the limit of Riemann-Stieltjes sums. Let $\boldsymbol{w} : \Delta \to T^{(n)}\left(\mathbb{R}^d\right)$ be the functional given by

$$\boldsymbol{w}_{s,t} = \left(1, w_{s,t}^1, \cdots, w_{s,t}^n\right), \quad (s,t) \in \Delta.$$

This is usually called the *lifting* of w up to degree n. The additivity property of integration over disjoint intervals is then summarized as the following so-called *Chen's identity*:

$$\boldsymbol{w}_{s,u} \otimes \boldsymbol{w}_{u,t} = \boldsymbol{w}_{s,t}, \quad \text{for all } 0 \leqslant s \leqslant u \leqslant t \leqslant 1. \tag{1.2}$$

We use $\Omega_n^\infty\left(\mathbb{R}^d\right)$ to denote the space of all such functionals which are liftings of smooth paths w. In the definition of Ω_n^∞, the starting point of the path is irrelevant, and we always assume that paths start at the origin.

Let $p \geqslant 1$ be fixed and $[p]$ denote the integer part of p (not greater than p). The *p-variation metric* d_p on $\Omega_{[p]}^\infty$ is defined by

$$d_p(\boldsymbol{u},\boldsymbol{w}) = \max_{1\leqslant i\leqslant [p]} \sup_D \left(\sum_l \left|u^i_{t_{l-1},t_l} - w^i_{t_{l-1},t_l}\right|^{\frac{p}{i}}\right)^{\frac{i}{p}},$$

where the supremum $\sup_D$ is taken over all possible finite partitions of $[0,1]$. The completion of $\Omega_{[p]}^\infty$ under d_p is called the space of *geometric p-rough paths* over $\mathbb{R}^d$, and it is denoted by $G\Omega_p\left(\mathbb{R}^d\right)$. If $\boldsymbol{w} = \left(1, w^1, \cdots, w^{[p]}\right) \in G\Omega_p\left(\mathbb{R}^d\right)$, then $\boldsymbol{w}$ also satisfies Chen's identity (1.2) in $T^{[p]}\left(\mathbb{R}^d\right)$, and $\boldsymbol{w}$ has finite p-variation in the sense that $\sup_D \sum_l \left|w_{t_{l-1},t_l}\right|^{\frac{p}{i}} < \infty$ for all $1 \leqslant i \leqslant [p]$.

The fundamental result in rough path theory is the following so-called *Lyons' universal limit theorem* (see [13], and also [6], [15]) for differential equations driven by geometric rough paths.

Theorem 1.1. *Let $\{V_1, \cdots, V_d\}$ be a family of γ-Lipschitz vector fields on $\mathbb{R}^N$ for some $\gamma > p$. For any given $x_0 \in \mathbb{R}^N$, define the map*

$$F(x_0,\cdot):\ \Omega_{[p]}^\infty\left(\mathbb{R}^d\right) \to G\Omega_p\left(\mathbb{R}^N\right)$$

in the following way. For any $\boldsymbol{w} \in \Omega_{[p]}^\infty\left(\mathbb{R}^d\right)$ which is the lifting of some smooth path w, let x be the unique smooth path which is the solution in $\mathbb{R}^N$ of the ODE

$$dx_t = \sum_{\alpha=1}^d V_\alpha(x_t)dw_t^\alpha,\ \ t \in [0,1],$$

with initial value x_0. $F(x_0, \boldsymbol{w})$ is then defined to be the lifting of x in $\Omega_p^\infty(\mathbb{R}^N)$. Then the map $F(x_0,\cdot)$ is uniformly continuous on bounded sets with respect to the p-variation metric.

Remark 1.1. Theorem 1.1 is not the original version of Lyons' result in [13] but an equivalent form. The original result of Lyons is formulated in terms of rough path integrals and does not restrict to geometric rough paths only. Here we state the result in a more elementary form to avoid the machinery of rough path integrals.

The theory of rough paths can be applied to quasi-sure analysis for Gaussian measures on path space. The notion of quasi-sure analysis was originally introduced by Malliavin [16] (see also [17]) to the study of non-degenerate conditioning and disintegration of Gaussian measures on abstract Wiener spaces. The fundamental concept in quasi-sure analysis is capacity, which specifies more precise scales for "negligible" subsets of an abstract Wiener space. In particular, a set of capacity zero is always a null set, while in general a null set may have positive capacity. According to Malliavin, the theory of quasi-sure analysis can be regarded as an infinite dimensional version of non-linear potential theory. It enables us to disintegrate a Gaussian measure continuously in the infinite dimensional setting, which for instance applies to the study of bridge processes and pinned diffusions. Moreover, it also leads to sharper estimates than classical methods.

The main goal of the present article is to initiate the study of Gaussian rough paths in the setting of quasi-sure analysis. Due to powerful tools in rough path theory, our results lead to the verification of many classical results for the quasi-sure analysis on Wiener space.

The first aim of this article is to study the quasi-sure existence of canonical lifting for sample paths of Gaussian processes as geometric rough paths. The Brownian motion case was studied by Inahama [10] under the p-variation metric, and Aida [1], Higuchi [9], Inahama [11] and Watanabe [23] independently under the Besov norm, by exploiting methods from the Malliavin calculus. More precisely, it was proved that for quasi-surely, Brownian sample paths can be lifted as geometric p-rough paths for $2 < p < 3$. In the next section, we extend this result to a class of Gaussian processes with long-time memory which includes fractional Brownian motion with Hurst parameter $H > 1/4$, by applying techniques both from rough path theory and the Malliavin calculus. Combining our result with Lyons' universal limit theorem, we obtain immediately a quasi-sure limit theorem for pathwise solutions to stochastic differential equations driven by Gaussian processes, which improves the Wong-Zakai type limit theorem and its quasi-sure version (see for example Ren [21], Malliavin-Nualart [18] and the references therein).

The technique we use in the next section enables us to establish a large deviation principle for capacities for Gaussian rough paths with long-time memory, which is the second aim of this article. LDPs for capacities for transformations on an abstract Wiener space was first studied by Yoshida [24]. The general definition and the basic properties of LDPs for induced capacities on a Polish space first appeared in Gao and Ren [7], in which the case of stochastic flows driven by Brownian motion was also investigated. Before establishing our LDP result, we first prove two fundamental results on transformations of LDPs for capacities: the contraction principle and exponential good approximations, which are both easy adaptations from the classical results for probability measures. Our LDP result is then based on the result and method developed in the next section and finite dimensional approximations. It turns out that the general result of Yoshida in the case of Gaussian processes is a direct corollary of our result due to the continuity of the projection map from a geometric rough path onto its first level path. The original proof of Yoshida relies crucially on the infinite dimensional structure of abstract Wiener space, and in particular deep properties of capacity and analytic properties of the Ornstein-Uhlenbeck semigroup. However, our technique here replies only on basic properties of capacity and finite dimensional Gaussian spaces. Moreover, again from Lyons' universal limit theorem, our LDP result immediately yields the LDPs for capacities for pathwise solutions to stochastic differential equations driven by Gaussian processes. In this respect our result is stronger than the result of Yoshida since we are working in a stronger topology (the p-variation topology) instead of the uniform topology, which is too weak to support the continuity of the solution map for differential equations. It is also interesting to note that Inahama [11] was already able to applied techniques from quasi-sure analysis to establish LDPs for pinned diffusion measures.

2 Quasi-sure Existence of Gaussian Rough Paths

In the present article, we consider the following class of Gaussian processes with long-time memory in the sense of Coutin-Qian [3].

Definition 2.1. A d-dimensional centered, continuous Gaussian process $\{B_t\}_{t\geqslant 0}$ starting at the origin with independent components is said to have *h-long-time memory* for some $0<h<1$ and if there is a constant C_h such that

$$\mathbb{E}\left[|B_t-B_s|^2\right]\leqslant C_h\,|t-s|^{2h}$$

for $s,t\geqslant 0$ and

$$\left|\mathbb{E}\left[\left(B_t^i-B_s^i\right)\left(B_{t+\tau}^i-B_{s+\tau}^i\right)\right]\right|\leqslant C_h\tau^{2h}\left|\frac{t-s}{\tau}\right|^2$$

for $1\leqslant i\leqslant d$, $s,t\geqslant 0$, $\tau>0$ with $(t-s)/\tau\leqslant 1$.

A fundamental example of Gaussian processes with long-time memory is fractional Brownian motion with h being the Hurst parameter (see [15]).

From now on, we always assume that such Gaussian process is realized on the path space over the finite time period $[0,1]$. This is of course equivalent to the consideration of the process over any $[0,T]$. Let W be the space of all $\mathbb{R}^d$-valued continuous paths w over $[0,1]$ with $w_0=0$, and equip W with the Borel σ-algebra $\mathcal{B}(W)$. Let $\mathbb{P}$ be the law on $(W,\mathcal{B}(W))$ of some Gaussian process with h-long-time memory in the sense of Definition 2.1.

It is a fundamental result of Coutin and Qian [3] that if $h>1/4,\ 2<p<4$ with $hp>1$, then outside a $\mathbb{P}$-null set each sample path $w\in W$ can be lifted as geometric p-rough paths in a canonical way. More precisely, for $m\geqslant 1$, let $t_m^k=k/2^m$ $(k=0,1,\cdots,2^m)$ be the m-th dyadic partition of $[0,1]$. Given $w\in W$, define $w^{(m)}$ to be the dyadic piecewise linear interpolation of w by

$$w_t^{(m)}=w_{t_m^{k-1}}+2^m\left(t-t_m^{k-1}\right)\left(w_{t_m^k}-w_{t_m^{k-1}}\right),\ t\in\left[t_m^{k-1},t_m^k\right],$$

and let

$$\boldsymbol{w}_{s,t}^{(m)}=\left(1,w_{s,t}^{(m),1},w_{s,t}^{(m),2},w_{s,t}^{(m),3}\right),\ (s,t)\in\Delta,$$

be the geometric rough path associated with $w^{(m)}$ up to level 3. Let $\mathcal{A}_p$ be the totality of all $w\in W$ such that $\{\boldsymbol{w}^{(m)}\}_{m\geqslant 1}$ is a Cauchy sequence under the p-variation metric d_p. Then $\mathcal{A}_p^c$ is a $\mathbb{P}$-null set and hence $\boldsymbol{w}^{(m)}$ converges to a unique geometric p-rough path $\boldsymbol{w}$ for $\mathbb{P}$-almost-surely. The convergence holds in $L^1(W,\mathbb{P})$ as well.

Remark 2.1. Although a geometric p-rough path is defined up to level $[p]$, by Lyons' extension theorem (see [13]) it does not make a difference if we always consider up to level 3 under d_p since $2 < p < 4$.

Remark 2.2. Coutin and Qian [3] also showed that if $h \leqslant 1/4$, no subsequence of $\boldsymbol{w}_{s,t}^{(m)}$ converges in probability or in L^1, and hence such canonical lifting of sample paths as geometric rough paths does not exist.

The goal of this section is to strengthen the result of Coutin-Qian to the quasi-sure setting in the sense of Malliavin. The main result and technique developed in this section are essential to establish a large deviation principle for capacities as we will see later on.

Throughout the rest of this article, we fix $h \in (1/4, 1/2]$, $p \in (2, 4)$ with $hp > 1$ (the case of $h > 1/2$ is trivial from the rough path point of view), and consider a d-dimensional Gaussian process with h-long-time memory.

We first recall some basic notions about the Malliavin calculus and quasi-sure analysis. We refer the readers to [17], [20] for a systematic discussion.

Let $\mathcal{H}$ be the Cameron-Martin space associated with the corresponding Gaussian measure $\mathbb{P}$ on W. $\mathcal{H}$ is canonically defined to be the space of all paths in W of the form

$$h_t = \mathbb{E}[Zw_t],\ t \in [0, 1],$$

where Z is an element of the L^2 space generated by the process (i.e. the L^2-closure of $\text{Span}\{w_t :\ t \in [0, 1]\}$), and the inner product is given by $\langle h_1, h_2 \rangle = \mathbb{E}[Z_1 Z_2]$. It follows that the identity map ι defines a continuous and dense embedding from $\mathcal{H}$ into W which makes $(W, \mathcal{H}, \mathbb{P})$ into an abstract Wiener space in the sense of Gross. Let $\iota^* :\ W^* \to \mathcal{H}^* \cong \mathcal{H}$ be the dual of ι. Then the identity map $\mathcal{I} :\ W^* \hookrightarrow L^2(W, \mathbb{P})$ uniquely extends to an isometric embedding from $\mathcal{H}$ into $L^2(W, \mathbb{P})$ via ι^*.

If f is a smooth Schwarz function on $\mathbb{R}^n$, and $\varphi_1, \cdots, \varphi_n \in W^*$, then $F = f(\varphi_1, \cdots, \varphi_n)$ is called a *smooth (Wiener) functional* on W. The collection of all smooth functionals on W is denoted by $\mathcal{S}$. The *Malliavin derivative* of F is defined to be the $\mathcal{H}$-valued functional

$$DF = \sum_{i=1}^{n} \frac{\partial f}{\partial x^i}(\varphi_1, \cdots, \varphi_n)\iota^*\varphi_i,$$

Such definition can be generalized to smooth functionals taking values in a separable Hilbert space E. Let $\mathcal{S}(E)$ be the space of E-valued functionals of the form $F = \sum_{i=1}^{k} F_i e_i$, where $F_i \in \mathcal{S}$, $e_i \in E$. The Malliavin derivative of F is defined to be the $\mathcal{H} \otimes E$-valued functional $DF = \sum_{i=1}^{k} DF_i \otimes e_i$. Such definition is independent of the form of F, and by induction we can define higher order derivatives $D^N F$ for $N \in \mathbb{N}$, which is then an $\mathcal{H}^{\otimes N} \otimes E$-valued functional. Given $q \geqslant 1,\ N \in \mathbb{N}$, the Sobolev norm $\|\cdot\|_{q,N;E}$ on $\mathcal{S}(E)$ is defined by

$$\|F\|_{q,N;E} = \left(\sum_{i=0}^{N} \mathbb{E}\left[\|D^i F\|_{\mathcal{H}^{\otimes i} \otimes E}^q \right] \right)^{\frac{1}{q}}.$$

We use $\|\cdot\|_{q;E}$ to denote the norm corresponding to the case $N = 0$ (the L^q-norm). The completion of $(\mathcal{S}(E), \|\cdot\|_{q,N;E})$ is called the (q,N)-*Sobolev space* for E-valued functionals over W, and it is denoted by $\mathbb{D}_N^q(E)$.

For any $q > 1, N \in \mathbb{N}$, the (q,N)-*capacity* $\mathrm{Cap}_{q,N}$ is a functional defined on the collection of all subsets of W. If O is an open subset of W, then

$$\mathrm{Cap}_{q,N}(O) := \inf\{\|u\|_{q,N} :\ u \in \mathbb{D}_N^q,\ u \geqslant 1 \text{ on } O,\ u \geqslant 0 \text{ on } W, \mathbb{P}\text{-a.s.}\}$$

and for any arbitrary subset A of W,

$$\mathrm{Cap}_{q,N}(A) := \inf\left\{\mathrm{Cap}_{q,N}(O) :\ O \text{ open}, A \subset O\right\}.$$

A subset $A \subset W$ is called *slim* if $\mathrm{Cap}_{q,N}(A) = 0$ for all $q > 1$ and $N \in \mathbb{N}$. A property for paths in W is said to hold for *quasi-surely if* it holds outside a slim set.

The (q,N)-capacity has the following basic properties:

(1) if $A \subset B$, then

$$0 \leqslant \mathrm{Cap}_{q,N}(A) \leqslant \mathrm{Cap}_{q,N}(B);$$

(2) $\mathrm{Cap}_{q,N}$ is increasing in N, and in q up to a constant depending on N;

(3) $\mathrm{Cap}_{q,N}$ is sub-additive, i.e.,

$$\mathrm{Cap}_{q,N}\left(\bigcup_{i=1}^{\infty} A_i\right) \leqslant \sum_{i=1}^{\infty} \mathrm{Cap}_{q,N}(A_i).$$

The following quasi-sure version of Tchebycheff's inequality and Borel-Cantelli's lemma play an essential role in the study of quasi-sure convergence in our approach. We refer the readers to [17] for the proof.

Proposition 2.1. *(1) For any $\lambda > 0$ and any $u \in \mathbb{D}_N^q$ which is lower semi-continuous, we have*

$$\mathrm{Cap}_{q,N}\{w \in W : u(w) > \lambda\} \leqslant \frac{C_{q,N}}{\lambda}\|u\|_{q,N},$$

where $C_{q,N}$ is a constant depending only on q and N.

(2) For any sequence $\{A_n\}_{n=1}^{\infty}$ of subsets of W, if $\sum_{n=1}^{\infty} \mathrm{Cap}_{q,N}(A_n) < \infty$, then

$$\mathrm{Cap}_{q,N}\left(\limsup_{n\to\infty} A_n\right) = 0.$$

Now we are in a position to state our main result of this section.

Theorem 2.1. *Suppose that $\mathbb{P}$ is the Gaussian measure on $(W, \mathcal{B}(W))$ associated with a d-dimensional Gaussian process with h-long-time memory for some $h \in (1/4, 1/2]$, $p \in (2,4)$ with $hp > 1$. Then $\mathcal{A}_p^c$ is a slim set. In particular, sample paths of the Gaussian processes can be lifted as geometric p-rough paths in a canonical way quasi-surely, as the limit of the lifting of dyadic piecewise linear interpolation under d_p.*

By applying Lyons' universal limit theorem (Theorem 1.1) for rough differential equations driven by geometric rough paths, an immediate consequence of Theorem 2.1 is the quasi-sure existence and uniqueness for pathwise solutions to stochastic differential equations driven by Gaussian processes with h-long-time memory in the sense of geometric rough paths, under certain regularity conditions on the generating vector fields.

The main idea of proving Theorem 2.1 is to use a crucial control on the p-variation metric which is defined over dyadic partitions only, and to apply basic results for Gaussian polynomials in the Malliavin calculus.

If $\boldsymbol{w} = (1, w^1, w^2, w^3)$ and $\tilde{\boldsymbol{w}} = (1, \tilde{w}^1, \tilde{w}^2, \tilde{w}^3)$ are two functionals on Δ taking values in $T^3(\mathbb{R}^d)$, define

$$\rho_i(\boldsymbol{w}, \tilde{\boldsymbol{w}}) = \left(\sum_{n=1}^{\infty} n^{\gamma} \sum_{k=1}^{2^n} \left| w^i_{t_n^{k-1}, t_n^k} - \tilde{w}^i_{t_n^{k-1}, t_n^k} \right|^{\frac{p}{i}} \right)^{\frac{i}{p}}, \tag{2.1}$$

where $i = 1, 2, 3$ and $\gamma > p - 1$ is a fixed constant . We use $\rho_j(\boldsymbol{w})$ to denote $\rho_j(\boldsymbol{w}, \tilde{\boldsymbol{w}})$ with $\tilde{\boldsymbol{w}} = (1, 0, 0, 0)$. These functionals were originally introduced by Hambly and Lyons [8] for constructing the stochastic area processes associated with Brownian motions on the Sierpinski gasket. They were then used by Ledoux, Qian and Zhang [12] to establish a large deviation principle for Brownian rough paths under the p-variation topology. We also use these functionals to prove a large deviation principle for capacity in the next section.

The following estimate is contained implicitly in [8] and made explicit in [15].

Lemma 2.1. *There exists a positive constant $C_{d,p,\gamma}$ depending only on d, p, γ, such that for any $\boldsymbol{w}, \widetilde{\boldsymbol{w}}$,*

$$\begin{aligned} d_p(\boldsymbol{w}, \widetilde{\boldsymbol{w}}) \leqslant\ & C_{d,p,\gamma} \max \{ \rho_1(\boldsymbol{w}, \widetilde{\boldsymbol{w}}), \rho_2(\boldsymbol{w}, \widetilde{\boldsymbol{w}}), \rho_1(\boldsymbol{w}, \widetilde{\boldsymbol{w}}) (\rho_1(\boldsymbol{w}) + \rho_1(\widetilde{\boldsymbol{w}})), \\ & \rho_3(\boldsymbol{w}, \widetilde{\boldsymbol{w}}), \rho_2(\boldsymbol{w}, \widetilde{\boldsymbol{w}}) (\rho_1(\boldsymbol{w}) + \rho_1(\widetilde{\boldsymbol{w}})), \\ & \rho_1(\boldsymbol{w}, \widetilde{\boldsymbol{w}}) \left(\rho_2(\boldsymbol{w}) + \rho_2(\widetilde{\boldsymbol{w}}) + (\rho_1(\boldsymbol{w}) + \rho_1(\widetilde{\boldsymbol{w}}))^2 \right) \}. \end{aligned} \tag{2.2}$$

The main difficulty of proving Theorem 2.1 is that it is unknown if the p-variation metric is differentiable in the sense of Malliavin. We get around this difficulty first by controlling the p-variation metric using Lemma 2.1 and then by observing that the capacity of $\left\{ \rho_i \left(\boldsymbol{w}^{(m+1)}, \boldsymbol{w}^{(m)} \right) > \lambda \right\}$ is "evenly distributed" over the dyadic sub-intervals (see (2.7) in the following). Our task is then reduced to the estimation of the Sobolev norms of certain Gaussian polynomials, which is contained in the following basic result in the Malliavin calculus (see [20]).

Lemma 2.2. *Fix $N \in \mathbb{N}$. Let $\mathcal{P}^N(E)$ be the space of E-valued polynomial functionals of degree less than or equal to N. Then for any $q > 2$ and any $F \in \mathcal{P}^N(E)$, we have*

$$\|F\|_{q;E} \leqslant (N+1)(q-1)^{\frac{N}{2}} \|F\|_{2;E}. \tag{2.3}$$

Moreover, for any $F \in \mathcal{P}^N(E)$ and for any $i \leqslant N$ we have

$$\|D^i F\|_{2;\mathcal{H}^{\otimes i}\otimes E} \leqslant N^{\frac{i+1}{2}} \|F\|_{2;E}. \tag{2.4}$$

The following L^2-estimates for the dyadic piecewise linear interpolation, which are contained in a series of calculations in [15], are crucial for us.

Lemma 2.3. *Let $m, n \geqslant 1$ and $k = 1, \cdots, 2^n$.*
1) For $i = 1, 2, 3$, we have

$$\left\| w^{(m),i}_{t_n^{k-1},t_n^k} \right\|_{2;(\mathbb{R}^d)^{\otimes i}} \leqslant \begin{cases} C_{d,h} 2^{-inh}, & n \leqslant m, \\ C_{d,h} 2^{im(1-h)-in}, & n > m. \end{cases}$$

2) We also have

$$\left\| w^{(m+1),1}_{t_n^{k-1},t_n^k} - w^{(m),1}_{t_n^{k-1},t_n^k} \right\|_{2;\mathbb{R}^d} \leqslant \begin{cases} 0, & n \leqslant m, \\ C_{d,h} 2^{m(1-h)-n}, & n > m; \end{cases}$$

$$\left\| w^{(m+1),2}_{t_n^{k-1},t_n^k} - w^{(m),2}_{t_n^{k-1},t_n^k} \right\|_{2;(\mathbb{R}^d)^{\otimes 2}} \leqslant \begin{cases} C_{d,h} 2^{\frac{1}{2}(1-4h)m - \frac{1}{2}n}, & n \leqslant m, \\ C_{d,h} 2^{2m(1-h)-2n}, & n > m; \end{cases}$$

$$\left\| w^{(m+1),3}_{t_n^{k-1},t_n^k} - w^{(m),3}_{t_n^{k-1},t_n^k} \right\|_{2;(\mathbb{R}^d)^{\otimes 3}} \leqslant \begin{cases} C_{d,h} 2^{\frac{1}{2}(1-4h)m - \frac{1+2h}{2}n}, & n \leqslant m, \\ C_{d,h} 2^{3m(1-h)-3n}, & n > m. \end{cases}$$

Here $C_{d,h}$ is a constant depending only on d and h.

Now we can proceed to the proof of Theorem 2.1. The key step is to establish estimates for the capacities of the tail events $\left\{w : \ \rho_i\left(\boldsymbol{w}^{(m+1)}, \boldsymbol{w}^{(m)}\right) > \lambda\right\}$ and $\left\{w : \ \rho_i\left(\boldsymbol{w}^{(m)}\right) > \lambda\right\}$ $(i = 1, 2, 3)$. This is contained in the following lemma.

Lemma 2.4. *Let $\theta \in \left(\left(\frac{p(2h+1)}{6} - 1\right)^+, hp - 1\right)$, $\tilde{N} > \frac{N}{2} \vee \left(2\left(h - \frac{\theta+1}{p}\right)\right)^{-1}$. Then we have*
(1)

$$\mathrm{Cap}_{q,N}\left\{w : \rho_i\left(\boldsymbol{w}^{(m+1)}, \boldsymbol{w}^{(m)}\right) > \lambda\right\} \leqslant C_i \lambda^{-2\tilde{N}} \left(\frac{1}{2^m}\right)^{2i\tilde{N}\left(h - \frac{\theta+1}{p}\right)-1}, \tag{2.5}$$

(2)

$$\mathrm{Cap}_{q,N}\left\{w : \rho_i\left(\boldsymbol{w}^{(m)}\right) > \lambda\right\} \leqslant C_i \lambda^{-2\tilde{N}}. \tag{2.6}$$

Here C_i is a positive constant of the form $C_i = C_1 C_2^{\tilde{N}} g\left(\tilde{N}; N\right) \tilde{N}^{i\tilde{N}}$, where C_1 depends only on q and N, C_2 depends only on $d, p, h, \gamma, \theta, q$ and $g\left(\tilde{N}; N\right)$ is a polynomial in $\tilde{N}$ with degree depending only on N and universal constant coefficients.

Proof. For $i = 1, 2, 3$, set

$$\begin{aligned} I_i(m;\lambda) &= \mathrm{Cap}_{q,N}\left\{w:\ \rho_i\left(\boldsymbol{w}^{(m+1)}, \boldsymbol{w}^{(m)}\right) > \lambda\right\} \\ &= \mathrm{Cap}_{q,N}\left\{w:\ \rho_i\left(\boldsymbol{w}^{(m+1)}, \boldsymbol{w}^{(m)}\right)^{\frac{p}{i}} > \lambda^{\frac{p}{i}}\right\}. \end{aligned}$$

By the definition of ρ_i, for every $\theta > 0$ we have

$$\begin{aligned} &\left\{w:\ \rho_i\left(\boldsymbol{w}^{(m+1)}, \boldsymbol{w}^{(m)}\right)^{\frac{p}{i}} > \lambda^{\frac{p}{i}}\right\} \\ &\subset \bigcup_{n=1}^{\infty}\left\{w:\ \sum_{k=1}^{2^n}\left|w^{(m+1),i}_{t_n^{k-1},t_n^k} - w^{(m),i}_{t_n^{k-1},t_n^k}\right|^{\frac{p}{i}} > C_{\gamma,\theta}\lambda^{\frac{p}{i}}\left(\frac{1}{2^n}\right)^{\theta}\right\}, \end{aligned}$$

where $C_{\gamma,\theta} = \left(\sum_{n=1}^{\infty} n^{\gamma} 2^{-n\theta}\right)^{-1}$. Therefore,

$$\begin{aligned} &I_i\left(m;\lambda\right) \\ &\leqslant \sum_{n=1}^{\infty}\mathrm{Cap}_{q,N}\left\{w:\ \sum_{k=1}^{2^n}\left|w^{(m+1),i}_{t_n^{k-1},t_n^k} - w^{(m),i}_{t_n^{k-1},t_n^k}\right|^{\frac{p}{i}} > \lambda^{\frac{p}{i}}C_{\gamma,\theta}\left(\frac{1}{2^n}\right)^{\theta}\right\} \\ &\leqslant \sum_{n=1}^{\infty}\sum_{k=1}^{2^n}\mathrm{Cap}_{q,N}\left\{w:\ \left|w^{(m+1),i}_{t_n^{k-1},t_n^k} - w^{(m),i}_{t_n^{k-1},t_n^k}\right|^{\frac{p}{i}} > \lambda^{\frac{p}{i}}C_{\gamma,\theta}\left(\frac{1}{2^n}\right)^{\theta+1}\right\}. \qquad (2.7) \end{aligned}$$

On the other hand, for any $\tilde{N} > 0$ we have

$$\begin{aligned} &\mathrm{Cap}_{q,N}\left\{\left|w^{(m+1),i}_{t_n^{k-1},t_n^k} - w^{(m),i}_{t_n^{k-1},t_n^k}\right|^{\frac{p}{i}} > \lambda^{\frac{p}{i}}C_{\gamma,\theta}\left(\frac{1}{2^n}\right)^{\theta+1}\right\} \\ &= \mathrm{Cap}_{q,N}\left\{f^i_{m,n,k} > \left[\lambda C_{\gamma,\theta}^{\frac{i}{p}}\left(\frac{1}{2^n}\right)^{\frac{i}{p}(\theta+1)}\right]^{2\tilde{N}}\right\}, \end{aligned}$$

where

$$f^i_{m,n,k}(w) = \left|w^{(m+1),i}_{t_n^{k-1},t_n^k} - w^{(m),i}_{t_n^{k-1},t_n^k}\right|^{2\tilde{N}}, \text{ for } w \in W.$$

Since $\tilde{N}$ is a natural number, $f^i_{m,n,k}$ are polynomial functionals of degree $2i\tilde{N}$, and hence they are N times differentiable in the sense of Malliavin provided $\tilde{N} \geqslant \frac{N}{2}$. Consequently, we can apply Tchebycheff's inequality (the first part of Proposition 2.1) to obtain

$$I_i(m;\lambda) \leqslant C_{q,N}\sum_{n=1}^{\infty}\sum_{k=1}^{2^n}\left(C_{\gamma,\theta}^{\frac{i}{p}}\lambda\left(\frac{1}{2^n}\right)^{\frac{i}{p}(\theta+1)}\right)^{-2\tilde{N}}\left\|f^i_{m,n,k}\right\|_{q,N}.$$

If $q > 2$, by (2.3) of Lemma 2.2, we have

$$\begin{aligned}\left\|f^i_{m,n,k}\right\|_{q,N} &\leqslant \sum_{l=0}^{N} \|D^l f^i_{m,n,k}\|_{q;\mathcal{H}^{\otimes l}} \\ &\leqslant \left(2i\tilde{N}+1\right)(q-1)^{i\tilde{N}} \sum_{l=0}^{N} \left\|D^l f^i_{m,n,k}\right\|_{2;\mathcal{H}^{\otimes l}}.\end{aligned}$$

By (2.4) of Lemma 2.2, we have

$$\left\|D^l f^i_{m,n,k}\right\|_{2;\mathcal{H}^{\otimes l}} \leqslant \left(2i\tilde{N}\right)^{\frac{N+1}{2}} \left\|f^i_{m,n,k}\right\|_2.$$

Therefore,

$$\left\|f^i_{m,n,k}\right\|_{q,N} \leqslant (N+1)\left(2i\tilde{N}+1\right)(q-1)^{i\tilde{N}}\left(2i\tilde{N}\right)^{\frac{N+1}{2}} \left\|f^i_{m,n,k}\right\|_2.$$

Moreover, since $w^{(m+1),i}_{t_n^{k-1},t_n^k} - w^{(m),i}_{t_n^{k-1},t_n^k}$ is an $(\mathbb{R}^d)^{\otimes i}$-valued polynomial functional of degree i, we know again from (2.3) that

$$\begin{aligned}\left\|f^i_{m,n,k}\right\|_2 &= \left\|w^{(m+1),i}_{t_n^{k-1},t_n^k} - w^{(m),i}_{t_n^{k-1},t_n^k}\right\|^{2\tilde{N}}_{4\tilde{N};(\mathbb{R}^d)^{\otimes i}} \\ &\leqslant (i+1)^{2\tilde{N}}\left(4\tilde{N}-1\right)^{i\tilde{N}}\left\|w^{(m+1),i}_{t_n^{k-1},t_n^k} - w^{(m),i}_{t_n^{k-1},t_n^k}\right\|^{2\tilde{N}}_{2;(\mathbb{R}^d)^{\otimes i}}.\end{aligned}$$

Therefore,

$$\begin{aligned}&\left\|f^i_{m,n,k}\right\|_{q,N} \\ \leqslant&(N+1)\left((q-1)^i(i+1)^2\right)^{\tilde{N}}\left(2i\tilde{N}+1\right)\left(2i\tilde{N}\right)^{\frac{N+1}{2}} \\ &\cdot\left(4\tilde{N}-1\right)^{i\tilde{N}}\left\|w^{(m+1),i}_{t_n^{k-1},t_n^k} - w^{(m),i}_{t_n^{k-1},t_n^k}\right\|^{2\tilde{N}}_{2;(\mathbb{R}^d)^{\otimes i}} \\ \leqslant&(N+1)\left(1024(q-1)^3\right)^{\tilde{N}}\left(6\tilde{N}+1\right)\left(6\tilde{N}\right)^{N}\tilde{N}^{i\tilde{N}}\left\|w^{(m+1),i}_{t_n^{k-1},t_n^k} - w^{(m),i}_{t_n^{k-1},t_n^k}\right\|^{2\tilde{N}}_{2;(\mathbb{R}^d)^{\otimes i}}.\end{aligned} \tag{2.8}$$

Let C_i be the constant before $\left\|w^{(m+1),i}_{t_n^{k-1},t_n^k} - w^{(m),i}_{t_n^{k-1},t_n^k}\right\|^{2\tilde{N}}_{2;(\mathbb{R}^d)^{\otimes i}}$ on the right hand side of (2.8).

By absorbing the constant in Tchebycheff's inequality into C_i, we arrive at

$$\begin{aligned}&I_i(m;\lambda) \\ &\leqslant C_i \sum_{n=1}^{\infty}\sum_{k=1}^{2^n}\left(C_\theta^{\frac{i}{p}}\lambda\left(\frac{1}{2^n}\right)^{\frac{i}{p}(\theta+1)}\right)^{-2\tilde{N}}\left\|w^{(m+1),i}_{t_n^{k-1},t_n^k} - w^{(m),i}_{t_n^{k-1},t_n^k}\right\|^{2\tilde{N}}_{2;(\mathbb{R}^d)^{\otimes i}}.\end{aligned} \tag{2.9}$$

Exactly the same computation yields

$$\begin{aligned}&\mathrm{Cap}_{q,N}\left\{w:\ \rho_i\left(\boldsymbol{w}^{(m)}\right)>\lambda\right\}\\ \leqslant\ & C_i\sum_{n=1}^{\infty}\sum_{k=1}^{2^n}\left(C_{\gamma,\theta}^{\frac{i}{p}}\lambda\left(\frac{1}{2^n}\right)^{\frac{i}{p}(\theta+1)}\right)^{-2\tilde{N}}\left\|w^{(m),i}_{t_n^{k-1},t_n^k}\right\|^{2\tilde{N}}_{2;(\mathbb{R}^d)^{\otimes i}}.\end{aligned} \tag{2.10}$$

We now substitute the estimates in Lemma 2.3 into (2.9) and (2.10). In what follows, we assume that $\theta\in\left(\left(\frac{p(2h+1)}{6}-1\right)^{+},hp-1\right)$, $\widetilde{N}>\frac{N}{2}\vee\left(2\left(h-\frac{\theta+1}{p}\right)\right)^{-1}$ for summability reason. We also absorb the constant $C_{d,h}$ in Lemma 2.3 and $C_{\gamma,\theta}$.

For $i=1$, this gives

$$\begin{aligned}I_1(m;\lambda)&\leqslant C_1\lambda^{-2\tilde{N}}2^{2\tilde{N}m(1-h)}\sum_{n=m+1}^{\infty}\sum_{k=1}^{2^n}2^{-2n\tilde{N}\left(1-\frac{\theta+1}{p}\right)}\\&\leqslant C_1\lambda^{-2\tilde{N}}2^{-m\left(2\tilde{N}\left(h-\frac{\theta+1}{p}\right)-1\right)}.\end{aligned}$$

For $i=2$, this gives

$$\begin{aligned}I_2(m;\lambda)\quad&\leqslant\quad C_2\lambda^{-2\tilde{N}}\left(\sum_{n=1}^{m}\sum_{k=1}^{2^n}2^{-n\tilde{N}\left(1-\frac{4(\theta+1)}{p}\right)-m\tilde{N}(4h-1)}\right.\\&\qquad\left.+\sum_{n=m+1}^{\infty}\sum_{k=1}^{2^n}2^{-4n\tilde{N}\left(1-\frac{\theta+1}{p}\right)+4m\tilde{N}(1-h)}\right)\\&\leqslant\quad C_2\lambda^{-2\tilde{N}}2^{-m\left(4\tilde{N}\left(h-\frac{\theta+1}{p}\right)-1\right)}.\end{aligned}$$

For $i=3$, this gives

$$\begin{aligned}I_3(m;\lambda)\quad&\leqslant\quad C_3\lambda^{-2\tilde{N}}\left(\sum_{n=1}^{m}\sum_{k=1}^{2^n}2^{-n\tilde{N}\left(1+2h-\frac{6(\theta+1)}{p}\right)-m\tilde{N}(4h-1)}\right.\\&\qquad\left.+\sum_{n=m+1}^{\infty}\sum_{k=1}^{2^n}2^{-6n\tilde{N}\left(1-\frac{\theta+1}{p}\right)+6m\tilde{N}(1-h)}\right)\\&\leqslant\quad C_3\lambda^{-2\tilde{N}}2^{-m\left(6\tilde{N}\left(h-\frac{\theta+1}{p}\right)-1\right)}.\end{aligned}$$

Therefore, for $i=1,2,3$, we have

$$I_i(m;\lambda)\leqslant C_i\lambda^{-2\tilde{N}}2^{-m\left(2i\tilde{N}\left(h-\frac{\theta+1}{p}\right)-1\right)},$$

which gives (2.5). From the computation it is easy to see that the constants C_i here are of the form stated in the lemma.

Similar computation yields that for $i = 1, 2, 3$,

$$\begin{aligned}
&\mathrm{Cap}_{q,N}\left\{w:\ \rho_i\left(\boldsymbol{w}^{(m)}\right) > \lambda\right\}\\
\leqslant\ & C_i\left(\lambda^{-2\tilde{N}}\sum_{n=1}^{m}\sum_{k=1}^{2^n}2^{-2n\tilde{N}i\left(h-\frac{\theta+1}{p}\right)}\right.\\
&\left.+\lambda^{-2\tilde{N}}\sum_{n=m+1}^{\infty}\sum_{k=1}^{2^n}2^{-2\tilde{N}i\left(n\left(1-\frac{\theta+1}{p}\right)-m(1-h)\right)}\right)\\
\leqslant\ & C_i\lambda^{-2\tilde{N}},
\end{aligned}$$

with C_i of the form stated in the lemma. This gives (2.6). □

Remark 2.3. The explicit form of the constants in Lemma 2.4 is used in the next section when proving a large deviation principle for capacities.

Now we are in a position to complete the proof of Theorem 2.1.

Proof of Theorem 2.1. By rewriting (2.2) as

$$\begin{aligned}
&d_p(\boldsymbol{w}, \widetilde{\boldsymbol{w}})\\
&\leqslant C_{d,p,\gamma}\max\left\{\rho_i(\boldsymbol{w}, \widetilde{\boldsymbol{w}})\left(\rho_j(\boldsymbol{w}) + \rho_j(\widetilde{\boldsymbol{w}})\right)^k : (i,j,k) \in \mathbb{N}\times\mathbb{N}\times\mathbb{Z}_+, i + jk \leqslant 3\right\},
\end{aligned} \tag{2.11}$$

we only need to show that there exists a positive constant β, such that for any $(i, j, k) \in \mathbb{N}\times\mathbb{N}\times\mathbb{Z}_+$ satisfying $i + jk \leqslant 3$, we have

$$\begin{aligned}
&\sum_{m=1}^{\infty}\mathrm{Cap}_{q,N}\left\{w:\ \rho_i\left(\boldsymbol{w}^{(m+1)}, \boldsymbol{w}^{(m)}\right)\left(\rho_j\left(\boldsymbol{w}^{(m)}\right) + \rho_j\left(\boldsymbol{w}^{(m+1)}\right)\right)^k > \frac{1}{2^{m\beta}}\right\}\\
&< \infty.
\end{aligned} \tag{2.12}$$

Indeed, if the above result holds, then by Lemma 2.1, we have

$$\sum_{m=1}^{\infty}\mathrm{Cap}_{q,N}\left\{w:\ d_p\left(\boldsymbol{w}^{(m)}, \boldsymbol{w}^{(m+1)}\right) > C'_{d,p,\gamma}\frac{1}{2^{m\beta}}\right\} < \infty,$$

where $C'_{d,p,\gamma}$ is some constant depending only on d, p, γ. It follows from the quasi-sure version of Borel-Catelli's lemma (the second part of Proposition 2.1) that

$$\mathrm{Cap}_{q,N}\left(\limsup_{m\to\infty}\left\{w:\ d_p\left(\boldsymbol{w}^{(m)}, \boldsymbol{w}^{(m+1)}\right) > C'_{d,p,\gamma}\frac{1}{2^{m\beta}}\right\}\right) = 0.$$

Since

$$\begin{aligned}\mathcal{A}_p^c &= \Big\{w:\ \boldsymbol{w}^{(m)} \text{ is not a Cauchy sequence in under } d_p\Big\}\\ &\subset \left\{w:\ \sum_{m=1}^{\infty} d_p\left(\boldsymbol{w}^{(m)}, \boldsymbol{w}^{(m+1)}\right) = \infty\right\}\\ &\subset \limsup_{m\to\infty}\left\{w:\ d_p\left(\boldsymbol{w}^{(m)}, \boldsymbol{w}^{(m+1)}\right) > C'_{d,p,\gamma}\frac{1}{2^{m\beta}}\right\},\end{aligned}$$

it follows that $\mathrm{Cap}_{q,N}(\mathcal{A}_p^c) = 0$ which completes the proof.

Now we prove (2.12).

First consider the case $k > 0$. For any $\alpha, \beta > 0$, we have

$$\begin{aligned}&\mathrm{Cap}_{q,N}\left\{w:\ \rho_i\left(\boldsymbol{w}^{(m+1)}, \boldsymbol{w}^{(m)}\right)\left(\rho_j\left(\boldsymbol{w}^{(m)}\right) + \rho_j\left(\boldsymbol{w}^{(m+1)}\right)\right)^k > \frac{1}{2^{m\beta}}\right\}\\ \leqslant\ &\mathrm{Cap}_{q,N}\left\{w:\ \rho_i\left(\boldsymbol{w}^{(m+1)}, \boldsymbol{w}^{(m)}\right) > \frac{1}{2^{m(\beta+\alpha)}}\right\}\\ &+\mathrm{Cap}_{q,N}\left\{w:\ \left(\rho_j\left(\boldsymbol{w}^{(m)}\right) + \rho_j\left(\boldsymbol{w}^{(m+1)}\right)\right)^k > 2^{m\alpha}\right\}\\ \leqslant\ &\mathrm{Cap}_{q,N}\left\{w:\ \rho_i\left(\boldsymbol{w}^{(m+1)}, \boldsymbol{w}^{(m)}\right) > \frac{1}{2^{m(\beta+\alpha)}}\right\}\\ &+\mathrm{Cap}_{q,N}\left\{w:\ \rho_j\left(\boldsymbol{w}^{(m)}\right) > 2^{\frac{m\alpha}{k}-1}\right\}\\ &+\mathrm{Cap}_{q,N}\left\{w:\ \rho_j\left(\boldsymbol{w}^{(m+1)}\right) > 2^{\frac{m\alpha}{k}-1}\right\}.\end{aligned}$$

By Lemma 2.4, for $\theta \in \left(\left(\frac{p(2h+1)}{6} - 1\right)^+, hp-1\right)$, $\widetilde{N} > \frac{N}{2} \vee \left(2\left(h - \frac{\theta+1}{p}\right)\right)^{-1}$ and $i = 1, 2, 3$, we have

$$\begin{aligned}&\mathrm{Cap}_{q,N}\left\{w:\ \rho_i\left(\boldsymbol{w}^{(m+1)}, \boldsymbol{w}^{(m)}\right) > \frac{1}{2^{m(\beta+\alpha)}}\right\}\\ \leqslant\ &C_i\left(\frac{1}{2^m}\right)^{2i\widetilde{N}\left(h-\frac{\theta+1}{p}\right)-1-2(\beta+\alpha)\widetilde{N}}.\end{aligned}$$

Let $\alpha, \beta > 0$ be such that

$$\frac{2\widetilde{N}\left(h - \frac{\theta+1}{p}\right) - 1}{2\widetilde{N}} > \beta + \alpha > 0. \tag{2.13}$$

It follows easily that

$$\sum_{m=1}^{\infty}\mathrm{Cap}_{q,N}\left\{w:\ \rho_i\left(\boldsymbol{w}^{(m+1)}, \boldsymbol{w}^{(m)}\right) > \frac{1}{2^{m(\beta+\alpha)}}\right\} < \infty. \tag{2.14}$$

Similarly,

$$\mathrm{Cap}_{q,N}\left\{w:\ \rho_j\left(\boldsymbol{w}^{(m)}\right) > 2^{\frac{m\alpha}{k}-1}\right\} \leqslant C_j 2^{-\frac{m\alpha}{k}-1},$$

and hence

$$\sum_{m=1}^{\infty} \mathrm{Cap}_{q,N} \left\{ w : \ \rho_j \left(\boldsymbol{w}^{(m)} \right) > 2^{\frac{m\alpha}{k}-1} \right\} < \infty.$$

Combining with (2.14), we arrive at

$$\sum_{m=1}^{\infty} \mathrm{Cap}_{q,N} \left\{ w : \ \rho_i \left(\boldsymbol{w}^{(m+1)}, \boldsymbol{w}^{(m)} \right) \left(\rho_j \left(\boldsymbol{w}^{(m)} \right) + \rho_j \left(\boldsymbol{w}^{(m+1)} \right) \right)^k > \frac{1}{2^{m\beta}} \right\}$$

$< \infty$.

The case of $k = 0$ follows directly from (2.14), since for all $\alpha > 0$,

$$\left\{ w : \ \rho_i \left(\boldsymbol{w}^{(m+1)}, \boldsymbol{w}^{(m)} \right) > \frac{1}{2^{m\beta}} \right\} \subset \left\{ w : \ \rho_i \left(\boldsymbol{w}^{(m+1)}, \boldsymbol{w}^{(m)} \right) > \frac{1}{2^{m(\beta+\alpha)}} \right\}.$$

Now the proof is complete.

□

3 Large Deviations for Capacities

In this section, we apply the previous technique to prove a large deviation principle for capacities for Gaussian rough paths with long-time memory.

Before stating our main result, we first recall the definition of general LDPs for induced capacities in Polish spaces (see [7], [24]).

Let (B, H, μ) be an abstract Wiener space.

Definition 3.1. Let $q > 1, N \in \mathbb{N}$, and let $\{T^\varepsilon\}$ be a family of $\mathrm{Cap}_{q,N}$-quasi surely defined maps from B to some Polish space (X, d). We say that the family $\{T^\varepsilon\}$ satisfies the $\mathrm{Cap}_{q,N}$-*large deviation principle* (or simply $\mathrm{Cap}_{q,N}$-*LDP*) with good rate function $I : \ X \to [0, \infty]$ if

(1) I is a good rate function on X, i.e. I is lower semi-continuous and for every $\alpha \in [0, \infty)$, the level set $\Psi_I(\alpha) = \{y \in X : \ I(y) \leqslant \alpha\}$ is compact in X;

(2) for every closed subset $C \subset X$, we have

$$\limsup_{\varepsilon \to 0} \varepsilon^2 \log \mathrm{Cap}_{q,N} \{w \in B : \ T^\varepsilon(w) \in C\} \leqslant -\frac{1}{q} \inf_{x \in C} I(x), \tag{3.1}$$

and for ever open subset $G \subset X$, we have

$$\liminf_{\varepsilon \to 0} \varepsilon^2 \log \mathrm{Cap}_{q,N} \{w \in B : \ T^\varepsilon(w) \in G\} \geqslant -\frac{1}{q} \inf_{x \in G} I(x). \tag{3.2}$$

Remark 3.1. The appearance of the factor $1/q$ comes from the definition of $\mathrm{Cap}_{q,N}$, so

$$\mathrm{Cap}_{q,N}(A) \geqslant \mathrm{Cap}_{q,0}(A) = \mathbb{P}(A)^{\frac{1}{q}}, \ \text{for all } A \in \mathcal{B}(B). \tag{3.3}$$

It is consistent with the classical large deviation principle for probability measures.

Due to the properties of (q, N)-capacity, many important results for LDPs can be carried through in the capacity setting without much difficulty, and the proofs are similar to the case of probability measures. Here we present two fundamental results on transformations of LDPs for capacities that are crucial for us, which did not appear in [7],[24] and related literatures.

The first result is the contraction principle.

Theorem 3.1. *Let $\{T^\varepsilon\}$ be a family of $\mathrm{Cap}_{q,N}$-quasi surely defined maps from B to (X, d) satisfying the $\mathrm{Cap}_{q,N}$-LDP with good rate function I. Let F be a continuous map from X to another Polish space (Y, d'). Then the family $\{F \circ T^\varepsilon\}$ of $C\mathrm{ap}_{q,N}$-quasi surely defined maps satisfies the $\mathrm{Cap}_{q,N}$-LDP with good rate function*

$$J(y) = \inf_{x:\ F(x)=y} I(x), \tag{3.4}$$

where we define $\inf \emptyset = \infty$.

Proof. Since I is a good rate function, it is not hard to see that J is lower semi-continuous and also by the continuity of F, if $J(y) < \infty$ then the infimum in (3.4) is attained at some point $x \in F^{-1}(y)$. Therefore, for any $\alpha > 0$, we have

$$\{y \in Y :\ J(y) \leqslant \alpha\} = F\left(\{x \in X :\ I(x) \leqslant \alpha\}\right),$$

and hence J is a good rate function. The $\mathrm{Cap}_{q,N}$-LDP (the upper bound (3.1) and lower bound (3.2)) for the family $\{F \circ T^\varepsilon\}$ under the good rate function J follows easily from the continuity of F. $\square$

The second result is about exponentially good approximations.

We first need the following definition.

Definition 3.2. Let $\{T^{\varepsilon,m}\}$ and $\{T^\varepsilon\}$ be two families of $\mathrm{Cap}_{q,N}$-quasi-surely defined maps from B to (X, d). We say that $\{T^{\varepsilon,m}\}$ are *exponentially good approximations* of $\{T^\varepsilon\}$ under $\mathrm{Cap}_{q,N}$, if for any $\lambda > 0$,

$$\lim_{m\to\infty} \limsup_{\varepsilon\to 0} \varepsilon^2 \log \mathrm{Cap}_{q,N}\{w :\ d(T^{\varepsilon,m}(w), T^\varepsilon(w)) > \lambda\} = -\infty. \tag{3.5}$$

Now we have the following result.

Theorem 3.2. *Suppose that for each $m \geqslant 1$, the family $\{T^{\varepsilon,m}\}$ of $\mathrm{Cap}_{q,N}$-quasi-surely defined maps satisfies the $\mathrm{Cap}_{q,N}$-LDP with good rate function I_m and $\{T^{\varepsilon,m}\}$ are exponentially good approximations of some family $\{T^\varepsilon\}$ of $\mathrm{Cap}_{q,N}$-quasi-surely defined maps. Suppose further that the function I defined by*

$$I(x) = \sup_{\lambda>0} \liminf_{m\to\infty} \inf_{y\in B_{x,\lambda}} I_m(y), \tag{3.6}$$

where $B_{x,\lambda}$ denotes the open ball $\{y \in X :\ d(y, x) < \lambda\}$, is a good rate function and for every closed set $C \subset X$,

$$\inf_{x\in C} I(x) \leqslant \limsup_{m\to\infty} \inf_{x\in C} I_m(x). \tag{3.7}$$

Then $\{T^\varepsilon\}$ satisfies the $\mathrm{Cap}_{q,N}$-LDP with good rate function I.

Proof. Upper bound. Let C be a closed subset of X. For any $\lambda > 0$, let $C_\lambda = \{x : d(x, C) \leqslant \lambda\}$. Since

$$\begin{aligned}&\{w : T^\varepsilon(w) \in C\}\\ \subset &\{w : T^{\varepsilon,m}(w) \in C_\lambda\} \bigcup \{w : d(T^{\varepsilon,m}(w), T^\varepsilon(w)) > \lambda\},\end{aligned}$$

it follows from the $\mathrm{Cap}_{q,N}$-LDP for $\{T^{\varepsilon,m}\}$ (the upper bound) that

$$\begin{aligned}&\limsup_{\varepsilon\to 0} \varepsilon^2 \log \mathrm{Cap}_{q,N} \{w : T^\varepsilon(w) \in C\}\\ \leqslant &\limsup_{\varepsilon\to 0} \varepsilon^2 \log \mathrm{Cap}_{q,N} \{w : T^{\varepsilon,m}(w) \in C_\lambda\}\\ &\vee \limsup_{\varepsilon\to 0} \varepsilon^2 \log \mathrm{Cap}_{q,N} \{w : d(T^{\varepsilon,m}(w), T^\varepsilon(w)) > \lambda\}\\ \leqslant &\left(-\frac{1}{q} \inf_{x\in C_\lambda} I_m(x)\right) \vee \limsup_{\varepsilon\to 0} \varepsilon^2 \log \mathrm{Cap}_{q,N} \{w : d(T^{\varepsilon,m}(w), T^\varepsilon(w)) > \lambda\}.\end{aligned}$$

By letting $m \to \infty$, we obtain from (3.5) and (3.7) that

$$\begin{aligned}\limsup_{\varepsilon\to 0} \varepsilon^2 \log \mathrm{Cap}_{q,N} \{w : T^\varepsilon(w) \in C\} &\leqslant -\frac{1}{q} \limsup_{m\to\infty} \inf_{x\in C_\lambda} I_m(x)\\ &\leqslant -\frac{1}{q} \inf_{x\in C_\lambda} I(x).\end{aligned}$$

Now the upper bound (3.1) follows from a basic property for good rate functions (see [4], Lemma 4.1.6) that

$$\lim_{\lambda\to 0} \inf_{x\in C_\lambda} I(x) = \inf_{x\in C} I(x).$$

To prove the lower bound (3.2), we first show that

$$\begin{aligned}-\frac{1}{q} I(x) &= \inf_{\lambda>0} \limsup_{\varepsilon\to 0} \varepsilon^2 \log \mathrm{Cap}_{q,N} \{w : T^\varepsilon(w) \in B_{x,\lambda}\}\\ &= \inf_{\lambda>0} \liminf_{\varepsilon\to 0} \varepsilon^2 \log \mathrm{Cap}_{q,N} \{w : T^\varepsilon(w) \in B_{x,\lambda}\}. \qquad (3.8)\end{aligned}$$

In fact, since

$$\begin{aligned}&\{w : T^{\varepsilon,m}(w) \in B_{x,\lambda}\}\\ \subset &\{w : T^\varepsilon(w) \in B_{x,2\lambda}\} \bigcup \{w : d(T^{\varepsilon,m}(w), T^\varepsilon(w)) > \lambda\}, \qquad (3.9)\end{aligned}$$

we have

$$\begin{aligned}&\mathrm{Cap}_{q,N} \{w : T^{\varepsilon,m}(w) \in B_{x,\lambda}\}\\ \leqslant &\mathrm{Cap}_{q,N} \{w : T^\varepsilon(w) \in B_{x,2\lambda}\} + \{w : d(T^{\varepsilon,m}(w), T^\varepsilon(w)) > \lambda\}.\end{aligned}$$

It follows from the $\mathrm{Cap}_{q,N}$-LDP (the lower bound) for $\{T^{\varepsilon,m}\}$ that

$$
\begin{aligned}
-\frac{1}{q}\inf_{y\in B_{x,\lambda}} I_m(y) &\leqslant \liminf_{\varepsilon\to 0}\varepsilon^2\log \mathrm{Cap}_{q,N}\{w:\ T^{\varepsilon,m}(w)\in B_{x,\lambda}\} \\
&\leqslant \liminf_{\varepsilon\to 0}\varepsilon^2\Big(\log \mathrm{Cap}_{q,N}\{w:\ T^{\varepsilon}(w)\in B_{x,2\lambda}\} \\
&\qquad \vee \log \mathrm{Cap}_{q,N}\{w:\ d(T^{\varepsilon,m}(w),T^{\varepsilon}(w))>\lambda\}\Big) \\
&\leqslant \liminf_{\varepsilon\to 0}\varepsilon^2\log \mathrm{Cap}_{q,N}\{w:\ T^{\varepsilon}(w)\in B_{x,2\lambda}\} \\
&\qquad \vee \limsup_{\varepsilon\to 0}\varepsilon^2\log \mathrm{Cap}_{q,N}\{w:\ d(T^{\varepsilon,m}(w),T^{\varepsilon}(w))>\lambda\},
\end{aligned}
$$

and (3.5) implies that

$$
-\frac{1}{q}\liminf_{m\to\infty}\inf_{y\in B_{x,\lambda}} I_m(y) \leqslant \liminf_{\varepsilon\to 0}\varepsilon^2\log \mathrm{Cap}_{q,N}\{w:\ T^{\varepsilon}(w)\in B_{x,2\lambda}\}.
$$

By taking infimum over $\lambda>0$, we obtain

$$
-\frac{1}{q}I(x) \leqslant \inf_{\lambda>0}\liminf_{\varepsilon\to 0}\varepsilon^2\log \mathrm{Cap}_{q,N}\{w:\ T^{\varepsilon}(w)\in B_{x,2\lambda}\}.
$$

On the other hand, by exchanging $T^{\varepsilon,m}$ and T^{ε} in (3.9), the same argument yields that (using the upper bound in the $\mathrm{Cap}_{q,N}$-LDP)

$$
\inf_{\lambda>0}\limsup_{\varepsilon\to 0}\varepsilon^2\log \mathrm{Cap}_{q,N}\{w:\ T^{\varepsilon}(w)\in B_{x,\lambda}\} \leqslant -\frac{1}{q}I(x).
$$

Therefore, (3.8) follows.

Lower bound. Let G be an open subset of X. For any fixed $x\in G$, take $\lambda>0$ such that $B_{x,\lambda}\subset G$. It follows from (3.8) that

$$
\begin{aligned}
&\liminf_{\varepsilon\to 0}\varepsilon^2\log \mathrm{Cap}_{q,N}\{w:\ T^{\varepsilon}(w)\in G\} \\
\geqslant &\liminf_{\varepsilon\to 0}\varepsilon^2\log \mathrm{Cap}_{q,N}\{w:\ T^{\varepsilon}(w)\in B_{x,\lambda}\} \\
\geqslant &-\frac{1}{q}I(x).
\end{aligned}
$$

Therefore, the lower bound (3.2) holds. □

Consider the abstract Wiener space $(W,\mathcal{H},\mathbb{P})$ associated with a Gaussian process satisfying the assumptions in Theorem 2.1. According to [6], the covariance function of the process has finite $(1/2h)$-variation in the 2D sense, and $\mathcal{H}$ is continuously embedded in the space of continuous paths with finite $(1/2h)$-variation. Therefore, every $h\in\mathcal{H}$ admits a natural lifting $\boldsymbol{h}$ in $G\Omega_p(\mathbb{R}^d)$ in the sense of iterated Young's integrals (see [25]).

Recall that $\mathcal{A}_p$ is the set of paths $w\in W$ such that the lifting $\boldsymbol{w}^{(m)}$ of the dyadic piecewise linear interpolation of w is a Cauchy sequence under d_p, and the map

$$
F:\ w\in\mathcal{A}_p\mapsto \boldsymbol{w}=\left(1,w^1,\cdots,w^{[p]}\right):=\lim_{m\to\infty}\boldsymbol{w}^{(m)}\in G\Omega_p\left(\mathbb{R}^d\right)
$$

is well-defined. For $\varepsilon > 0$, let $T^{\varepsilon} : \mathcal{A}_p \to G\Omega_p\left(\mathbb{R}^d\right)$ be the map defined by

$$T^{\varepsilon}(w) = \delta_{\varepsilon}\boldsymbol{w} := (1, \varepsilon w^1, \cdots, \varepsilon^{[p]} w^{[p]}).$$

By Theorem 2.1, $\mathcal{A}_p^c$ is a slim set. Therefore, T^{ε} is quasi-surely well-defined.

Let

$$\Lambda(w) = \begin{cases} \frac{1}{2}\|w\|_{\mathcal{H}}^2, & w \in \mathcal{H}; \\ \infty, & \text{otherwise}, \end{cases} \tag{3.10}$$

and define $I : G\Omega_p\left(\mathbb{R}^d\right) \to [0, \infty]$ by

$$I(\boldsymbol{w}) = \inf\{\Lambda(w) : w \in \mathcal{A}_p, F(w) = \boldsymbol{w}\}. \tag{3.11}$$

We will see later in Lemma 3.2 that $\mathcal{H} \subset \mathcal{A}_p$ and hence

$$I(\boldsymbol{w}) = \begin{cases} \frac{1}{2}\left\|\pi_1(\boldsymbol{w})_{0,\cdot}\right\|_{\mathcal{H}}^2, & \text{if } \pi_1(\boldsymbol{w})_{0,\cdot} \in \mathcal{H} \text{ and } \boldsymbol{w} = F\left(\pi_1(\boldsymbol{w})_{0,\cdot}\right); \\ \infty, & \text{otherwise}, \end{cases}$$

where π_1 is the projection onto the first level path.

Now we can state our main result of this section.

Theorem 3.3. *For any $q > 1$, $N \in \mathbb{N}$, the family $\{T^{\varepsilon}\}$ of $\mathrm{Cap}_{q,N}$-quasi-surely defined maps from W to $G\Omega_p\left(\mathbb{R}^d\right)$ satisfies the $\mathrm{Cap}_{q,N}$-LDP with good rate function I.*

In particular, since the projection map from $G\Omega_p\left(\mathbb{R}^d\right)$ onto the first level path is continuous, we immediately obtain the following result of Yoshida [24] in the case of Gaussian processes with long-time memory.

Corollary 3.1. *The family of maps $\{\varepsilon w\}$ satisfies the $\mathrm{Cap}_{q,N}$-LDP with good rate function Λ.*

Moreover, according to the universal limit theorem (Theorem 1.1) and the contraction principle (Theorem 3.1), a direct corollary of Theorem 3.3 is the LDPs for capacities for solutions to differential equations driven by Gaussian rough paths with long-time memory. This generalizes the classical Freidlin-Wentzell theory for diffusion measures to the quasi-sure and rough path setting and in particular recovers a result of Gao and Ren [7]. Here we are again taking the advantage of working in the stronger topology (the p-variation topology), under which we have nice stability for differential equations.

It should be pointed out that the lifting map F, which can be regarded as the pathwise solution to a differential equation driven by w with a polynomial one form, is *not* continuous under the uniform topology (see [14], [15]). Therefore the contraction principle cannot be applied directly in our context. A standard way of getting around this difficulty, as in [12] for Brownian motion and [19] for fractional Brownian motion in the case of LDPs for probability measures, is to construct exponentially good approximations by using dyadic piecewise linear interpolation. Here we adopt the same idea in the capacity setting.

Let $T^{\varepsilon,m}: W \to G\Omega_p(\mathbb{R}^d)$ be the map given by $T^{\varepsilon,m}(w) = \delta_\varepsilon \boldsymbol{w}^{(m)}$. The proof of Theorem 3.3 essentially consists of two parts: show that the family $\{T^{\varepsilon,m}\}$ satisfies a $\mathrm{Cap}_{q,N}$-LDP and show that $\{T^{\varepsilon,m}\}$ are exponentially good approximations of $\{T^\varepsilon\}$ under $\mathrm{Cap}_{q,N}$.

We first need to establish the $\mathrm{Cap}_{q,N}$-LDP for $\{T^{\varepsilon,m}\}$, and we begin with considering the standard finite dimensional abstract Wiener space.

Let μ be the standard Gaussian measure on $\mathbb{R}^n$. In this case, the Cameron-Martin space is just $\mathbb{R}^n$ equipped with the standard Euclidean inner product. For clarity we use the notation $\mathrm{Cap}_{q,N}^{(n)}$ to emphasize that the capacity is defined on $\mathbb{R}^n$. Now we have the following result.

Proposition 3.1. *The family $\{\varepsilon x\}$ satisfies the $\mathrm{Cap}_{q,N}^{(n)}$-LDP with good rate function*

$$J(x) = \frac{|x|^2}{2},\ x \in \mathbb{R}^n.$$

Proof. The lower bound follows immediately from the simple relation in (3.3) and the classical LDP for the family $\{\mu\left(\varepsilon^{-1}dx\right)\}$ of probability measures. It suffices to establish the upper bound.

We first prove the following inequality for the one dimensional case:

$$\limsup_{\varepsilon\to 0} \varepsilon^2 \log \mathrm{Cap}_{q,N}^{(1)}\{x:\ \varepsilon x > b\} \leqslant -\frac{1}{2q}b^2, \tag{3.12}$$

for any $b > 0$. In fact, for any $\lambda > 0$, define the non-negative function

$$f(x) = \mathrm{e}^{\lambda\varepsilon x - \lambda b},\ x \in \mathbb{R}^1.$$

Obviously $f \in \mathbb{D}_N^q$, and $f \geqslant 1$ on $\{x:\ \varepsilon x > b\}$. Therefore, by the definition of capacity we have

$$\begin{aligned} \mathrm{Cap}_{q,N}^{(1)}\{x:\ \varepsilon x > b\} &\leqslant \|f\|_{q,N} \\ &\leqslant \sum_{i=0}^{N}\left(\int_{\mathbb{R}^1}\left|f^{(i)}\right|^q \mu(dx)\right)^{\frac{1}{q}} \\ &= \sum_{i=0}^{N}\left(\int_{\mathbb{R}^1}(\lambda\varepsilon)^{qi}\mathrm{e}^{q\lambda\varepsilon x - q\lambda b}\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{x^2}{2}}dx\right)^{\frac{1}{q}} \\ &= \sum_{i=0}^{N}(\lambda\varepsilon)^i \mathrm{e}^{\frac{q}{2}(\lambda\varepsilon)^2 - \lambda b}. \end{aligned}$$

It follows that

$$\varepsilon^2 \log \mathrm{Cap}_{q,N}^{(1)}\{x:\ \varepsilon x > b\} \leqslant \varepsilon^2 \log N + \max_{0\leqslant i\leqslant N}\left\{i\varepsilon^2\log(\lambda\varepsilon)\right\} + \frac{q}{2}(\lambda\varepsilon^2)^2 - \lambda\varepsilon^2 b.$$

Now take $\lambda = b/(q\varepsilon^2)$, then we have

$$\varepsilon^2 \log \mathrm{Cap}_{q,N}^{(1)}\{x:\ \varepsilon x > b\} \leqslant \varepsilon^2 \log N + \max_{0\leqslant i\leqslant N}\left\{i\varepsilon^2\log\left(\frac{b}{q\varepsilon}\right)\right\} - \frac{b^2}{2q},$$

and therefore (3.12) holds. Apparently (3.12) still holds if $\{x: \ \varepsilon x > b\}$ is replaced by $\{x: \ \varepsilon x \geqslant b\}$, and a similar inequality holds for $\{x: \ \varepsilon x \leqslant a\}$ for $a < 0$.

Now we come back to the n-dimensional case.

Firstly, consider an open ball $B(a,r) \subset \mathbb{R}^n$. For any $\lambda \in \mathbb{R}^n$, consider the non-negative function

$$f(x) = \mathrm{e}^{\langle \lambda, \varepsilon x\rangle + |\lambda| r - \langle \lambda, a\rangle}, \ x \in \mathbb{R}^n.$$

Then apparently we have $f \in \mathbb{D}_N^q$. Moreover, from the fact that

$$\langle \lambda, a\rangle - |\lambda| r = \inf_{y \in B(a,r)} \langle \lambda, y\rangle,$$

we have $f \geqslant 1$ on $\{x: \ \varepsilon x \in B(a,r)\}$. Therefore, similarly as before we have

$$\begin{aligned} \mathrm{Cap}_{q,N}^{(n)}\{x: \ \varepsilon x \in B(a,r)\} &\leqslant \|f\|_{q,N} \\ &\leqslant \sum_{i=0}^{N} \left(\int_{\mathbb{R}^n} \left| D^i f \right|^q \mu(dx) \right)^{\frac{1}{q}} \\ &\leqslant \sum_{i=0}^{N} (n|\lambda|\varepsilon)^i \mathrm{e}^{\frac{1}{2} q (|\lambda| \varepsilon)^2 + |\lambda| r - \langle \lambda, a\rangle} \end{aligned}$$

and

$$\begin{aligned} & \varepsilon^2 \log \mathrm{Cap}_{q,N}^{(n)}\{x: \ \varepsilon x \in B(a,r)\} \\ \leqslant \ & \varepsilon^2 \log N + \max_{0 \leqslant i \leqslant N} \left\{ i\varepsilon^2 \log(n|\lambda\varepsilon|) \right\} + \frac{q}{2} \left(|\lambda| \varepsilon^2 \right)^2 \\ & + |\lambda| \varepsilon^2 r - \langle \varepsilon^2 \lambda, a\rangle. \end{aligned}$$

Note that the function $\frac{q}{2} \left(|\lambda| \varepsilon^2 \right)^2 + |\lambda| \varepsilon^2 r - \langle \varepsilon^2 \lambda, a\rangle$ attains its minimum at

$$\lambda = \frac{(|a| - r)^+}{q \varepsilon^2 |a|} a,$$

By taking this λ and letting $\varepsilon \to 0$, we arrive at

$$\limsup_{\varepsilon \to 0} \varepsilon^2 \log \mathrm{Cap}_{q,N}^{(n)}\{x: \ \varepsilon x \in B(a,r)\} \leqslant -\frac{1}{2q} \left((|a| - r)^+ \right)^2 = -\frac{1}{q} \inf_{y \in B(a,r)} J(y).$$

Secondly, let K be a compact subset of $\mathbb{R}^n$. Then for any $\delta > 0$, we can find a finite cover of K by open balls $\{B(a_i, \delta)\}_{1 \leqslant i \leqslant k(\delta)}$ where each $a_i \in K$. It follows that

$$\begin{aligned}
&\limsup_{\varepsilon\to 0}\varepsilon^2\log \mathrm{Cap}_{q,N}\{x:\ \varepsilon x\in K\}\\
&\leqslant \limsup_{\varepsilon\to 0}\varepsilon^2\left(\log k(\delta)+\max_{1\leqslant i\leqslant k(\delta)}\log \mathrm{Cap}_{q,N}^{(n)}\{x:\ \varepsilon x\in B(a_i,\delta)\}\right)\\
&\leqslant \max_{1\leqslant i\leqslant k(\delta)}\left(-\frac{1}{q}\inf_{y\in B(a_i,\delta)}J(y)\right)\\
&\leqslant -\frac{1}{q}\inf_{y\in B(K,\delta)}J(y),
\end{aligned}$$

where $B(K,\delta):=\{x:\ \mathrm{dist}(x,K)<\delta\}$. By letting $\delta\to 0$ we obtain the upper bound result for the compact set K.

Finally, let C be an arbitrary closed subset of $\mathbb{R}^n$. For $\rho>0$, let

$$H_\rho=\{x:\ \left|x^i\right|\leqslant \rho \text{ for all } i\}.$$

Then we have

$$\mathrm{Cap}_{q,N}^{(n)}\{x:\ \varepsilon x\in C\}\leqslant \mathrm{Cap}_{q,N}^{(n)}\{x:\ \varepsilon x\in C\bigcap H_\rho\}+\sum_{i=1}^{n}\mathrm{Cap}_{q,N}^{(n)}\{x:\ \varepsilon\left|x^i\right|>\rho\}.$$

On the other hand, from the definition of capacity, we have (see also the proof of the following Corollary 3.2):

$$\mathrm{Cap}_{q,N}^{(n)}\{x:\ \varepsilon\left|x^i\right|>\rho\}\leqslant \mathrm{Cap}_{q,N}^{(1)}\{x\in\mathbb{R}^1:\ \varepsilon|x|>\rho\}.$$

Combining with the upper bound result for compact sets and (3.12), we arrive at

$$\limsup_{\varepsilon\to 0}\varepsilon^2\log \mathrm{Cap}_{q,N}^{(n)}\{x:\ \varepsilon x\in C\}\leqslant \max\left\{-\frac{1}{q}\inf_{y\in C\bigcap H_\rho}J(y),-\frac{1}{q}\rho^2\right\}$$

for all $\rho>0$. The upper bound result for C follows from letting $\rho\to\infty$. □

Now consider the situation where ν is a general non-degenerate Gaussian measure on $\mathbb{R}^n$ with covariance matrix Σ. In this case the Cameron-Martin space $\mathcal{H}=\mathbb{R}^n$ but with inner product

$$\langle h_1,h_2\rangle=h_1^T\Sigma^{-1}h_2.$$

Moreover, the Cameron-Martin embedding $\iota:\ \mathcal{H}\to\mathbb{R}^n$ is just the identity map but the dual embedding $\iota^*:\ \mathbb{R}^n\to\mathcal{H}^*\cong\mathcal{H}$ is given by

$$\iota^*(\lambda)=\Sigma\lambda,\ \lambda\in\mathbb{R}^n.$$

Therefore, if we write $\Sigma=QQ^T$ for some non-degenerate matrix Q, it follows from the definition of Sobolev spaces and change of variables that

$$\mathrm{Cap}_{q,N}^{\nu}(A)=\mathrm{Cap}_{q,N}^{\mu}\left(Q^{-1}A\right)$$

for all $A \subset \mathbb{R}^n$, where the left hand side is the capacity for ν and the right hand side is the capacity for the standard Gaussian measure μ. In other words, capacities for non-degenerate Gaussian measures on $\mathbb{R}^n$ are all equivalent. As a consequence, we conclude that the family $\{\varepsilon x\}$ satisfies the $\mathrm{Cap}^{\nu}_{q,N}$-LDP with good rate function

$$J(y) = \frac{1}{2}\|y\|^2_{\mathcal{H}} = \frac{1}{2}y^T\Sigma^{-1}y,\ y \in \mathbb{R}^n.$$

The case of degenerate Gaussian measures follows easily by restriction on the maximal invariant subspace on which the covariance matrix is positive definite.

A direct consequence of the previous discussion is the following.

Corollary 3.2. *For each $m \geqslant 1$, the family $\{T^{\varepsilon,m}\}$ satisfies the $\mathrm{Cap}_{q,N}$-LDP with good rate function*

$$I_m(\boldsymbol{w}) = \inf\left\{J_m(x):\ x \in \left(\mathbb{R}^d\right)^{2^m}:\ \Phi_m(x) = \boldsymbol{w}\right\},\ \boldsymbol{w} \in G\Omega_p\left(\mathbb{R}^d\right), \tag{3.13}$$

where $J_m(x)$ is the good rate function for the Gaussian measure ν_m on $\left(\mathbb{R}^d\right)^{2^m}$ induced by $\left(w_{t^1_m}, \cdots, w_{t^{2^m}_m}\right)$, and Φ_m is the map sending each $x \in \left(\mathbb{R}^d\right)^{2^m}$ to the lifting of the dyadic piecewise linear interpolation associated with x.

Proof. Since Φ_m is continuous under the Euclidean and p-variation topology respectively, the result follows immediately from the contraction principle (Theorem 3.1) once we have established the $\mathrm{Cap}_{q,N}$-LDP for the family $\varepsilon\pi_m:\ W \to \left(\mathbb{R}^d\right)^{2^m}$ where π^m is defined by

$$\pi_m(w) = \left(w_{t^1_m}, \cdots, w_{t^{2^m}_m}\right),\ w \in W,$$

with good rate function J_m.

To see this, first notice again that the lower bound follows from the relation (3.3) and the classical LDP for finite dimensional Gaussian measures. Moreover, let U be an open subset of $\left(\mathbb{R}^d\right)^{2^m}$ and let $f \in \mathbb{D}^q_N\left(\nu_m\right)$ be a function such that for ν_m-almost surely

$$f \geqslant 1 \text{ on } U,\ f \geqslant 0 \text{ on } \left(\mathbb{R}^d\right)^{2^m},$$

$\mathbb{D}^q_N\left(\nu_m\right)$ is the Sobolev space over $\left(\mathbb{R}^d\right)^{2^m}$ associated with ν_m. Define

$$g(w) = f\left(w_{t^1_m}, \cdots, w_{t^{2^m}_m}\right),\ w \in W.$$

Apparently $g \in \mathbb{D}^q_N$, and for $\mathbb{P}$-almost surely

$$g \geqslant 1 \text{ on } \pi_m^{-1}U,\ g \geqslant 0 \text{ on } W.$$

Moreover, since $\|g\|_{q,N} = \|f\|_{q,N;\nu_m}$, we know that

$$\mathrm{Cap}_{q,N}\left(\pi_m^{-1}U\right) \leqslant \|f\|_{q,N;\nu_m}.$$

By taking infimum over all such f, we obtain

$$\mathrm{Cap}_{q,N}\left(\pi_m^{-1}U\right) \leqslant \mathrm{Cap}_{q,N}^{\nu_m}(U).$$

Now the upper bound result follows from the $\mathrm{Cap}_{q,N}$-LDP for the family $\{\nu_{m,\varepsilon} := \nu_m(\varepsilon^{-1}dx)\}$ of probability measure and a simple limiting argument. □

Remark 3.2. There is an equivalent way of expressing the rate function I_m, which is very convenient for us to prove our main result of Theorem 3.3. In fact, from classical LDP results for Gaussian measures (see for example [5]), we know that the family $\left\{\mathbb{P}_\varepsilon := \mathbb{P}\left(\varepsilon^{-1}dw\right)\right\}$ of probability measures on W satisfies the LDP with good rate function Λ given by (3.10). Moreover, the map $\Psi_m : W \to G\Omega_p(\mathbb{R}^d)$ defined by $\Psi_m(w) = \boldsymbol{w}^{(m)}$ is continuous under the uniform and p-variation topology respectively. Therefore, according to the classical contraction principle, the family $\{\mathbb{P}_\varepsilon \circ \Psi_m^{-1}\}$ of probability measures on $G\Omega_p\left(\mathbb{R}^d\right)$ satisfies the LDP with good rate function

$$I_m'(\boldsymbol{w}) = \inf\left\{\Lambda(w) : \ w \in W, \ \Psi_m(w) = \boldsymbol{w}\right\}, \ \boldsymbol{w} \in G\Omega_p\left(\mathbb{R}^d\right). \tag{3.14}$$

On the other hand, the same argument implies that the family $\left\{\nu_{m,\varepsilon} \circ \Phi_m^{-1}\right\}$ of probability measures on $G\Omega_p\left(\mathbb{R}^d\right)$ satisfies the LDP with good rate function I_m given by (3.13). Observe that $\mathbb{P}_\varepsilon \circ \Psi_m^{-1} = \nu_{m,\varepsilon} \circ \Phi_m^{-1}$. By the uniqueness of rate functions (see [4], Chapter 4, Lemma 4.1.4), we conclude that $I_m = I_m'$.

Remark 3.3. Of course we can apply Yoshida's result directly with the contraction principle to obtain the $\mathrm{Cap}_{q,N}$-LDP for the family $\{T^{\varepsilon,m}\}$ with good rate function I_m'. Here we do not proceed in this way so that in the end our result yields Yoshida's one as a corollary, and our proof relies only on basic properties of capacities and finite dimensional Gaussian spaces.

The second main ingredient of proving Theorem 3.3 is the following.

Lemma 3.1. *For any $q > 1, N \in \mathbb{N}$ and $\lambda > 0$, we have*

$$\lim_{m\to\infty} \limsup_{\varepsilon\to 0} \varepsilon^2 \log \mathrm{Cap}_{q,N}\left\{w : \ d_p\left(\delta_\varepsilon \boldsymbol{w}^{(m)}, \delta_\varepsilon \boldsymbol{w}\right) > \lambda\right\} = -\infty.$$

Therefore, $\{T^{\varepsilon,m}\}$ are exponentially good approximations of $\{T^\varepsilon\}$ under $\mathrm{Cap}_{q,N}$.

Proof. For any $\beta > 0$, since

$$\begin{aligned}
&\left\{w : \ d_p\left(\delta_\varepsilon \boldsymbol{w}^{(m)}, \delta_\varepsilon \boldsymbol{w}\right) > \lambda\right\} \\
\subset &\left\{w : \ \sum_{l=m}^{\infty} d_p\left(\delta_\varepsilon \boldsymbol{w}^{(l)}, \delta_\varepsilon \boldsymbol{w}^{(l+1)}\right) > \lambda\right\} \\
\subset &\bigcup_{l=m}^{\infty}\left\{w : \ d_p\left(\delta_\varepsilon \boldsymbol{w}^{(l)}, \delta_\varepsilon \boldsymbol{w}^{(l+1)}\right) > \frac{\lambda}{C_\beta} \cdot \frac{1}{2^{(l-m)\beta}}\right\},
\end{aligned}$$

we have

$$
\begin{aligned}
&\mathrm{Cap}_{q,N}\left\{w:\ d_p\left(\delta_\varepsilon \boldsymbol{w}^{(m)},\delta_\varepsilon \boldsymbol{w}\right)>\lambda\right\}\\
&\leqslant \sum_{l=m}^{\infty}\mathrm{Cap}_{q,N}\left\{w:\ d_p\left(\delta_\varepsilon \boldsymbol{w}^{(l)},\delta_\varepsilon \boldsymbol{w}^{(l+1)}\right)>\frac{\lambda}{C_\beta}\cdot\frac{1}{2^{(l-m)\beta}}\right\},
\end{aligned}
$$

where $C_\beta:=\sum_{k=0}^{\infty}2^{-\beta k}$. It then follows from (2.11) that for any $\alpha>0$,

$$
\begin{aligned}
&\mathrm{Cap}_{q,N}\left\{w:\ d_p\left(\delta_\varepsilon \boldsymbol{w}^{(m)},\delta_\varepsilon \boldsymbol{w}\right)>\lambda\right\}\\
&\leqslant \sum_{i=1}^{3}\sum_{l=m}^{\infty}\mathrm{Cap}_{q,N}\left\{w:\ \rho_i\left(\delta_\varepsilon \boldsymbol{w}^{(l)},\delta_\varepsilon \boldsymbol{w}^{(l+1)}\right)>\frac{\lambda}{C_{d,p,\gamma,\beta}}\frac{1}{2^{(l-m)\beta}}\right\}\\
&\quad+\sum_{\substack{i,j,k\geqslant 1\\ i+jk\leqslant 3}}\sum_{l=m}^{\infty}\left(\mathrm{Cap}_{q,N}\left\{w:\ \rho_i\left(\delta_\varepsilon \boldsymbol{w}^{(l)},\delta_\varepsilon \boldsymbol{w}^{(l+1)}\right)>\frac{\lambda}{C_{d,p,\gamma,\beta}}\frac{2^{m\beta}}{2^{l(\alpha+\beta)}}\right\}\right.\\
&\quad+\mathrm{Cap}_{q,N}\left\{w:\ \rho_j\left(\delta_\varepsilon \boldsymbol{w}^{(l)}\right)>2^{\frac{l\alpha}{k}-1}\right\}\\
&\quad+\mathrm{Cap}_{q,N}\left\{w:\ \rho_j\left(\delta_\varepsilon \boldsymbol{w}^{(l+1)}\right)>2^{\frac{l\alpha}{k}-1}\right\},
\end{aligned}
\tag{3.15}
$$

where $C_{d,p,\gamma,\beta}$ is a constant depending only on p,d,γ,β.

Similar to the proof of Theorem 2.1, we estimate each term on the right hand side of (3.15). Here we choose α,β in exactly the same way as in the proof of Theorem 2.1, namely, by (2.13). It should be pointed out that the choice of α,β can be made independent of $\widetilde{N}$, since $\theta\in\left(\left(\frac{p(2h+1)}{6}-1\right)^{+},hp-1\right)$.

Firstly, it follows from Lemma 2.4 that for $i=1,2,3$,

$$
\begin{aligned}
&\mathrm{Cap}_{q,N}\left\{w:\ \rho_i\left(\delta_\varepsilon \boldsymbol{w}^{(l)},\delta_\varepsilon \boldsymbol{w}^{(l+1)}\right)>\frac{\lambda}{C_{d,p,\gamma,\beta}}\frac{1}{2^{(l-m)\beta}}\right\}\\
&=\mathrm{Cap}_{q,N}\left\{w:\ \rho_i\left(\boldsymbol{w}^{(l)},\boldsymbol{w}^{(l+1)}\right)>\frac{\lambda}{C_{d,p,\gamma,\beta}}\frac{\varepsilon^{-i}}{2^{(l-m)\beta}}\right\}\\
&\leqslant C_1C_2^{\widetilde{N}}g\left(\widetilde{N};N\right)\widetilde{N}^{i\widetilde{N}}\cdot\left(\frac{\lambda}{C_{d,p,\gamma,\beta}}\frac{\varepsilon^{-i}}{2^{(l-m)\beta}}\right)^{-2\widetilde{N}}\cdot\left(\frac{1}{2^l}\right)^{2i\widetilde{N}\left(h-\frac{\theta+1}{p}\right)-1}\\
&=C_1C_3^{\widetilde{N}}g\left(\widetilde{N};N\right)\left(\widetilde{N}\varepsilon^2\right)^{i\widetilde{N}}\cdot\frac{1}{2^{2m\widetilde{N}\beta}}\left(\frac{1}{2^l}\right)^{2i\widetilde{N}\left(h-\frac{\theta+1}{p}\right)-1-2\widetilde{N}\beta},
\end{aligned}
\tag{3.16}
$$

where $C_3=C_2\left(\frac{\lambda}{C_{d,p,\gamma,\beta}}\right)^{-2}$. Note that by the choice of β, the right hand side of (3.16) is summable over l, and it follows that

$$
\begin{aligned}
&\sum_{l=m}^{\infty}\mathrm{Cap}_{q,N}\left\{w:\ \rho_i\left(\delta_\varepsilon \boldsymbol{w}^{(l)},\delta_\varepsilon \boldsymbol{w}^{(l+1)}\right)>\frac{\lambda}{C_{d,p,\gamma,\beta}}\frac{1}{2^{(l-m)\beta}}\right\}\\
&\leqslant C_4C_3^{\widetilde{N}}g\left(\widetilde{N};N\right)\left(\widetilde{N}\varepsilon^2\right)^{i\widetilde{N}}\cdot\left(\frac{1}{2^m}\right)^{2i\widetilde{N}\left(h-\frac{\theta+1}{p}\right)-1},
\end{aligned}
$$

where $C_4 = C_1 \left(1 - 2^{-\left(2i\widetilde{N}\left(h-\frac{\theta+1}{p}\right)-1-2\widetilde{N}\beta\right)}\right)^{-1}$. By taking $\widetilde{N} = \left[\varepsilon^{-2}\right]$ for ε small enough, it is easy to see that

$$\begin{aligned}
&\limsup_{\varepsilon\to 0} \varepsilon^2 \log \left(\sum_{l=m}^{\infty} \mathrm{Cap}_{q,N} \left\{ w:\ \rho_i \left(\delta_\varepsilon \boldsymbol{w}^{(l)}, \delta_\varepsilon \boldsymbol{w}^{(l+1)} \right) > \frac{\lambda}{C_{d,p,\gamma,\beta}} \frac{1}{2^{(l-m)\beta}} \right\} \right) \\
=& \log C_3 + 2i \left(h - \frac{\theta+1}{p} \right) \log \left(\frac{1}{2^m} \right).
\end{aligned}$$

Therefore, we have

$$\begin{aligned}
&\lim_{m\to\infty} \limsup_{\varepsilon\to 0} \varepsilon^2 \log \left(\sum_{l=m}^{\infty} \mathrm{Cap}_{q,N} \left\{ w:\ \rho_i \left(\delta_\varepsilon \boldsymbol{w}^{(l)}, \delta_\varepsilon \boldsymbol{w}^{(l+1)} \right) > \frac{\lambda}{C_{d,p,\gamma,\beta}} \frac{1}{2^{(l-m)\beta}} \right\} \right) \\
=& -\infty.
\end{aligned}$$

Again by the choice of α, β and by taking $\widetilde{N} = \left[\varepsilon^{-2}\right]$, the same computation based on Lemma 2.4 yields that

$$\begin{aligned}
&\lim_{m\to\infty} \limsup_{\varepsilon\to 0} \varepsilon^2 \log \left(\sum_{l=m}^{\infty} \mathrm{Cap}_{q,N} \left\{ w:\ \rho_i \left(\delta_\varepsilon \boldsymbol{w}^{(l)}, \delta_\varepsilon \boldsymbol{w}^{(l+1)} \right) > \frac{\lambda}{C_{d,p,\gamma,\beta}} \frac{2^{m\beta}}{2^{l(\alpha+\beta)}} \right\} \right) \\
=& \lim_{m\to\infty} \limsup_{\varepsilon\to 0} \varepsilon^2 \log \left(\sum_{l=m}^{\infty} \mathrm{Cap}_{q,N} \left\{ w:\ \rho_j \left(\delta_\varepsilon \boldsymbol{w}^{(l)} \right) > 2^{\frac{l\alpha}{k}-1} \right\} \right) \\
=& \lim_{m\to\infty} \limsup_{\varepsilon\to 0} \varepsilon^2 \log \left(\sum_{l=m}^{\infty} \mathrm{Cap}_{q,N} \left\{ w:\ \rho_j \left(\delta_\varepsilon \boldsymbol{w}^{(l+1)} \right) > 2^{\frac{l\alpha}{k}-1} \right\} \right) \\
=& -\infty,
\end{aligned}$$

for $i, j, k \geqslant 1$ with $i + jk \leqslant 3$.

Now the desired result follows easily. □

In order to apply Theorem 3.2, we need the following convergence result in [6] for Cameron-Martin paths.

Lemma 3.2. *For any $\alpha > 0$, we have*

$$\lim_{m\to\infty} \sup_{\{h\in\mathcal{H}:\ \|h\|_{\mathcal{H}}\leqslant\alpha\}} d_p\left(\boldsymbol{h}^{(m)}, \boldsymbol{h}\right) = 0.$$

In particular, $\mathcal{H}$ is contained in $\mathcal{A}_p$.

Now we are in a position to prove Theorem 3.3.

Proof of Theorem 3.3. It suffices to show that the function I given by (3.11) coincides with the one given by (3.6), and it satisfies all conditions in Theorem 3.2. Here we use I'_m given by (3.14) for the rate function of $\{T^{\varepsilon,m}\}$.

Firstly, by Lemma 3.2 it is easy to see that the lifting map F is continuous on each level set $\{w: \Lambda(w) \leqslant \alpha\} \subset \mathcal{H} \subset \mathcal{A}_p$ of Λ. It follows from the definition of I that

$$F(\{w: \Lambda(w) \leqslant \alpha\}) = \{\boldsymbol{w}: I(\boldsymbol{w}) \leqslant \alpha\},$$

which then implies that I is a good rate function.

Now we show that for any closed subset $C \subset G\Omega_p(\mathbb{R}^d)$, we have

$$\inf_{\boldsymbol{w}\in C} I(\boldsymbol{w}) \leqslant \liminf_{m\to\infty} \inf_{\boldsymbol{w}\in C} I'_m(\boldsymbol{w}). \tag{3.17}$$

In fact, let $\gamma_m = \inf_{\boldsymbol{w}\in C} I'_m(\boldsymbol{w}) = \inf_{w\in\Psi_m^{-1}(C)} \Lambda(w)$. We only consider the nontrivial case $\liminf_{m\to\infty} \gamma_m = \alpha < \infty$, and without loss of generality we assume that $\lim_{m\to\infty} \gamma_m = \alpha$. Since Λ is a good rate function, we know that the infimum over the closed subset $\Psi_m^{-1}(C) \subset W$ is attainable. Therefore, there exists $w_m \in W$ such that $\Psi_m(w_m) \in C$ and $\gamma_m = \Lambda(w_m)$. It follows from Lemma 3.2 that for any fixed $\lambda > 0$, $F(w_m) \in C_\lambda$ when m is large, where $C_\lambda := \{\boldsymbol{w}: d_p(\boldsymbol{w}, C) \leqslant \lambda\}$. Consequently, when m is large, we have

$$\inf_{\boldsymbol{w}\in C_\lambda} I(\boldsymbol{w}) \leqslant I(F(w_m)) = \Lambda(w_m) = \gamma_m,$$

and hence

$$\inf_{\boldsymbol{w}\in C_\lambda} I(\boldsymbol{w}) \leqslant \alpha.$$

(3.17) then follows easily from [4], Chapter 4, Lemma 4.1.6. by taking $\lambda \to 0$.

A direct consequence of (3.17) is the condition (3.7) in Theorem 3.2. Moreover, if we let $C = \overline{B_{\boldsymbol{w},\lambda}}$ in (3.17), by taking $\lambda \to 0$ we easily obtain that $I(\boldsymbol{w}) \leqslant \overline{I}(\boldsymbol{w})$, where $\overline{I}$ is the function given by (3.6).

It remains to show that $\overline{I}(\boldsymbol{w}) \leqslant I(\boldsymbol{w})$, and we only consider the nontrivial case $I(\boldsymbol{w}) = \alpha < \infty$. It follows that $I(\boldsymbol{w}) = \Lambda(w)$, where $w \in \mathcal{H} \subset \mathcal{A}_p$ with $F(w) = \boldsymbol{w}$. Let $\boldsymbol{w}_m = \Psi_m(w)$. By Lemma 3.2 we know that $\boldsymbol{w}_m \to \boldsymbol{w}$ under d_p. Therefore, for any fixed $\lambda > 0$,

$$\inf_{\boldsymbol{w}'\in B_{\boldsymbol{w},\lambda}} I'_m(\boldsymbol{w}') \leqslant I'_m(\boldsymbol{w}_m) \leqslant \Lambda(w) = I(\boldsymbol{w})$$

when m is large. By taking "$\liminf_{m\to\infty}$" and "$\sup_{\lambda>0}$", we obtain that $\overline{I}(\boldsymbol{w}) \leqslant I(\boldsymbol{w})$.

Now the proof is complete.

□

Remark 3.4. In some literature (in particular, in [24]), the Sobolev norms over $(W, \mathcal{H}, \mathbb{P})$ are defined in terms of the Ornstein-Uhlenbeck operator, which can be regarded as the infinite dimensional Laplacian under the Gaussian measure $\mathbb{P}$. An advantage of using such norms is that they can be easily extended to the fractional case. According to the well known Meyer's inequalities, such norms are equivalent to the ones we have used here which are defined in terms of the Malliavan derivatives. Therefore, the LDP for the corresponding capacities under these Sobolev norms holds in exactly the same way.

Acknowledgements

The research of the authors is supported by the Oxford-Man Institute in University of Oxford, and the first two authors are also supported by ERC (Grant Agreement No.291244 Esig). We would like to thank Professor Inahama for bringing the papers [1], [9], [10] and [11] to our attention.

References

[1] S. Aida, Vanishing of one-dimensional L^2-cohomologies of loop groups, *J. Funct. Anal.* 261 (8), 2164–2213, 2011.

[2] R. F. Bass, B. M. Hambly and T. Lyons, Extending the Wong-Zakai theorem to reversible Markov processes, *J. Eur. Math. Soc.* 4 (3), 237–269, 2002.

[3] L. Coutin and Z. Qian, Stochastic analysis, rough path analysis and fractional Brownian motions, *Probab. Theory Relat. Fields* 122 (1), 108–140, 2002.

[4] A. Dembo and O. Zeitouni, *Large deviations techniques and applications,* 2nd Edition, Springer-Verlag, Berlin, 1998.

[5] J.D. Deuschel and D.W. Stroock, *Large deviations,* Academic Press, New York, 1989.

[6] P. K. Friz and N. B. Victoir, *Multidimensional stochastic processes as rough paths*, Cambridge University Press, Cambridge, 2010.

[7] F. Gao and J. Ren, Large deviations for stochastic flows and their applications, *Science in China Series A: Mathematics* 44 (8), 1016–1033, 2001.

[8] B. M. Hambly and T. Lyons, Stochastic area for Brownian motion on the Sierpinski gasket, *Ann. Probab.* 26 (1), 132–148, 1998.

[9] Y. Higuchi, Master Thesis, Japanese Version, Graduate School of Engineering Sciences, Osaka University, 2006.

[10] Y. Inahama, Quasi-sure existence of Brownian Rough Paths and a construction of Brownian Pants, Infin. Dimens. *Anal. Quantum Probab. Relat. Top.* 9 (4), 513–528, 2006.

[11] Y. Inahama, Large deviation principle of Freidlin-Wentzell type for pinned diffusion processes, *arXiv preprint*, arXiv:1203.5177, 2012.

[12] M. Ledoux, Z. Qian, and T. Zhang, Large deviations and support theorem for diffusion processes via rough paths, *Stochastic Process. Appl.* 102 (2), 265–283, 2002.

[13] T. Lyons, Differential equations driven by rough signals, *Rev. Mat. Iberoamericana* 14 (2), 215–310, 1998.

[14] T. Lyons, M. Caruana and T. Lévy, *Differential equations driven by rough paths*, Springer, 2007.

[15] T. Lyons and Z. Qian, *System control and rough paths*, Oxford Mathematical Monographs, Oxford University Press, Oxford, 2002.

[16] P. Malliavin, Implicit function in finite co-rank on the Wiener space, *Stochastic analysis*, 369–386, 1982.

[17] P. Malliavin, *Stochastic analysis*, Springer-Verlag, Berlin, 1997.

[18] P. Malliavin and D. Nualart, Quasi-sure analysis and Stratonovich anticipative stochastic differential equations, *Probab. Theory Relat. Fields* 96 (1), 45–55, 1993.

[19] A. Millet and M. Sanz-Solé, Large deviations for rough paths of the fractional Brownain motion, *Ann. Inst. Henri Poincaré Probab. Stat.* 42 (2), 245–271, 2006

[20] D. Nualart, *The Malliavin calculus and related topics*, 2nd Edition, Springer-Verlag, Berlin, 1995.

[21] J. Ren, Analyse quasi-sûre des équations différentielles stochastiques, *Bull. Sci. Math.* 114 (2), 187–213, 1990.

[22] E. M. Sipiläinen, *A pathwise view of solutions of stochastic differential equations*, Ph.D. thesis, University of Edinburgh, 1993.

[23] S. Watanabe, Itô calculus and Malliavin calculus, *Stoch. Anal. Appl.* Abel Symp. (2), 623–639, 2007.

[24] N. Yoshida, A large deviation principle for (r, p)-capacities on the Wiener space, *Probab. Theory Relat. Fields* 94 (4), 473–488, 1993.

[25] L. C. Young, An inequality of Hölder type connected with Stieltjes integration, *Acta. Math.* 67, 251–282, 1936.