

***G*-Brownian Motion as Rough Paths and Differential Equations Driven by *G*-Brownian Motion**

Xi Geng, Zhongmin Qian, and Danyu Yang

Abstract The present article is devoted to the study of sample paths of *G*-Brownian motion and stochastic differential equations (SDEs) driven by *G*-Brownian motion from the view of rough path theory. As the starting point, by using techniques in rough path theory, we show that quasi-surely, sample paths of *G*-Brownian motion can be enhanced to the second level in a canonical way so that they become geometric rough paths of roughness $2 < p < 3$. This result enables us to introduce the notion of rough differential equations (RDEs) driven by *G*-Brownian motion in the pathwise sense under the general framework of rough paths. Next we establish the fundamental relation between SDEs and RDEs driven by *G*-Brownian motion. As an application, we introduce the notion of SDEs on a differentiable manifold driven by *G*-Brownian motion and construct solutions from the RDE point of view by using pathwise localization technique. This is the starting point of developing *G*-Brownian motion on a Riemannian manifold, based on the idea of Eells-Elworthy-Malliavin. The last part of this article is devoted to such construction for a wide and interesting class of *G*-functions whose invariant group is the orthogonal group. In particular, we establish the generating nonlinear heat equation for such *G*-Brownian motion on a Riemannian manifold. We also develop the Euler-Maruyama approximation for SDEs driven by *G*-Brownian motion of independent interest.

Xi Geng

Mathematical Institute, University of Oxford, Oxford OX2 6GG, the United Kingdom and the Oxford-Man Institute, University of Oxford, Oxford OX2 6ED, the United Kingdom
e-mail: xi.geng@maths.ox.ac.uk

Zhongmin Qian (✉)

Exeter College, University of Oxford, OX1 3DP, the United Kingdom
e-mail: qianz@maths.ox.ac.uk

Danyu Yang

Mathematical Institute, University of Oxford, Oxford OX2 6GG, the United Kingdom and the Oxford-Man Institute, University of Oxford, Oxford OX2 6ED, the United Kingdom
e-mail: danyu.yang@maths.ox.ac.uk

Keywords G -Brownian motion • G -expectation • Euler-Maruyama approximation • Geometric rough paths • Rough differential equations • Nonlinear diffusion processes • Nonlinear heat flow

AMS 2010 subject classifications. 60H10, 34A12, 58J65.

1 Introduction

The classical Feynman-Kac formula (see [14], [15]) provides us with a way to represent the solution of a linear parabolic PDE in terms of the conditional expectation of certain functional of a diffusion process (solution of an SDE). However, it works only for the linear case, mainly due to the linear nature of diffusion processes. To understand nonlinear parabolic PDEs from the probabilistic point of view, Peng and Pardoux (see [21], [22], [23]) initiated the study of backward stochastic differential equations (BSDEs) and showed that the solution of a certain type of quasilinear parabolic PDEs can be expressed in terms of the solution of BSDE. This result suggests that BSDE reveals a certain type of nonlinear dynamics, and was made explicit by Peng [24]. More precisely, Peng introduced a notion of nonlinear expectation called the g -expectation in terms of the solution of BSDE which is filtration consistent. However, it was developed under the framework of classical Itô calculus and did not capture the fully nonlinear situation.

Motivated from the study of fully nonlinear dynamics, Peng [25] introduced the notion of G -expectation in an intrinsic way which does not rely on any particular probability space. It reveals the probability distribution uncertainty in a fundamental way which is crucial in many situations such as modeling risk uncertainty in mathematical finance. The underlying mechanism corresponding to such kind of uncertainty is a fully nonlinear parabolic PDE. In [25], [26], he also introduced the concept of G -Brownian motion which is generated by the so-called nonlinear G -heat equation and related stochastic calculus such as G -Itô integral, G -Itô formula, SDEs driven by G -Brownian motion, etc. One of the major significance of such theory is the corresponding nonlinear Feynman-Kac formula proved by Peng [27], which gives us a way to represent the solution of a fully nonlinear parabolic PDE via the solution of a forward-backward SDE under the framework of G -expectation.

On the other hand, motivated from the study of integration against irregular paths and differential equations driven by rough signals, Lyons [17] proposed a theory of rough paths which reveals the fundamental way of understanding the roughness of a continuous path. He pointed out that to understand the evolution of a system whose input signal (driven path) is rough, a finite sequence of “iterated integrals” (higher levels) of the driving path which satisfy a certain type of algebraic relation (Chen identity) should be specified in advance. Such point of view is fundamental, if we look at the Taylor expansion for the solution of an ODE whose driving path is of bounded variation (see (6) and a more detailed introduction in the next section). In other words, it is essential to regard a path as an object valued in some tensor

algebra which records the information of higher levels if we wish to understand the “differential” of the path. Moreover, Lyons [17] proved the so-called universal limit theorem (see Theorem 6 in the next section), which allows us to introduce the notion of differential equations driven by rough paths (simply called RDEs) in a rigorous way. The theory of rough paths has significant applications in classical stochastic analysis, as we can prove that the sample paths of many stochastic processes we’ve encountered are essentially rough paths with certain roughness. According to Lyons’ universal limit theorem, we are able to establish RDEs driven by the sample paths of those stochastic processes in a pathwise manner. It provides us with a new way to understand SDEs, especially when the driving process is not the classical Brownian motion in which case a well-developed Itô SDE theory is still not available.

The case of classical Brownian motion is quite special, since we have a complete SDE theory in the L^2 -sense, as well as the notion of Stratonovich type integrals and differential equations. The fundamental relation between the two types of stochastic differentials (one-dimensional case) can be expressed by

$$X \circ dY = X dY + \frac{1}{2} dX \cdot dY.$$

It is proved in the rough path theory (see [9], [19], and also [13], [28] from the view of Wong-Zakai type approximation) that the Stratonovich type integrals and differential equations are equivalent to the pathwise integrals and RDEs in the sense of rough paths. In other words, the following two types of differential equations driven by Brownian motion

$$\begin{aligned} dX_t &= \sum_{\alpha=1}^d V_\alpha(X_t) dW_t^\alpha + b(X_t) dt, \quad (\text{Itô type SDE}) \\ dY_t &= \sum_{\alpha=1}^d V_\alpha(Y_t) dW_t^\alpha + (b(Y_t) - \sum_{\alpha=1}^d \frac{1}{2} DV_\alpha(Y_t) \cdot V_\alpha(Y_t)) dt, \quad (\text{RDE}) \end{aligned}$$

which are both well-defined under some regularity assumptions on the generating vector fields, are equivalent in the sense that if their solutions X_t and Y_t satisfy $X_0 = Y_0$, then $X = Y$ almost surely.

Under the framework of G -expectation, SDEs driven by G -Brownian motion introduced by Peng, can be regarded as nonlinear diffusion processes in Euclidean spaces. The idea of constructing G -Itô integrals and SDEs driven by G -Brownian motion is similar to the classical Itô calculus, which is also an L^2 -theory but under the G -expectation instead of probability measures. What is missing is the notion of Stratonovich type integrals, mainly due to the reason that the theory of G -martingales is still not well understood. In particular, we don’t have the corresponding nonlinear Doob-Meyer type decomposition theorem and the notion of quadratic variation processes for G -martingales. However, by the key observation in the classical case that the Stratonovich type integrals and the pathwise integrals are essentially equivalent in the sense of rough paths, we can study the sample paths of G -Brownian motion and SDEs driven by G -Brownian motion from the view of rough path the-

ory, once we prove that the sample paths of G -Brownian motion can be regarded as objects in some rough path space with certain roughness. This is in fact what the present article is mainly focused on. The basic language to describe path structure under the G -expectation is quasi-sure analysis and capacity theory, which was developed by Denis, Hu and Peng [7]. They generalized the Kolmogorov continuity theorem and studied sample path properties of G -Brownian motion. In particular, they also studied the relation between G -expectation and upper expectation associated to a family of probability measures which defines a Choquet capacity and the relation between the corresponding two types of L^p -spaces. The pathwise properties and homeomorphic flows for SDEs driven by G -Brownian motion in the quasi-sure setting was studied by Gao [10].

There are two main goals of the present article. This first one is to study the geometric rough path nature of sample paths of G -Brownian motion so that we can define RDEs driven by G -Brownian motion (the Stratonovich counterpart in the classical case) in the pathwise sense, and establish the fundamental relation between two types of differential equations driven by G -Brownian motion. The second one is to understand nonlinear diffusion processes in a (Riemannian) geometric setting, from the view of paths and distributions (the generating nonlinear PDE).

The present article is organized in the following way. Section 2 is a basic review of the theory of G -expectation and rough paths, which provides us with the general framework and basic tools for our study. In Section 3 we study the Euler-Maruyama approximation scheme for SDEs driven by G -Brownian motion. In Section 4 we show that for quasi-surely, the sample paths of G -Brownian motion can be enhanced to the second level in a canonical way so that they become geometric rough paths of roughness $2 < p < 3$ by using techniques in rough path theory. In Section 5 we establish the fundamental relation between SDEs and RDEs driven by G -Brownian motion by using rough Taylor expansions. In section 6 we introduce the notion of SDEs on a differentiable manifold driven by G -Brownian motion from the RDE point of view by using pathwise localization technique. In the last section, we study the infinitesimal diffusive nature and the generating PDE for nonlinear diffusion processes in a (Riemannian) geometric setting, which leads to the construction of G -Brownian motion on a Riemannian manifold. We restrict ourselves to compact manifolds only, although the general case can be treated in a similar way with more technical complexity.

Throughout the rest of this article, we will use standard geometric notation for differential equations. Moreover, we will use the Einstein convention of summation, that is, when an index α appears as both subscript and superscript in the same expression, summation over α is taken automatically.

2 Preliminaries on G -expectation and Rough Path Theory

2.1 G -expectation and Related Stochastic Calculus

We first introduce some fundamentals on G -expectation and related stochastic calculus. For a systematic introduction, see [25], [26], [27].

Let Ω be a nonempty set, and \mathcal{H} be a vector space of functionals on Ω such that \mathcal{H} contains all constant functionals and for any $X_1, \dots, X_n \in \mathcal{H}$ and any $\varphi \in C_{l,Lip}(\mathbb{R}^n)$,

$$\varphi(X_1, \dots, X_n) \in \mathcal{H},$$

where $C_{l,Lip}(\mathbb{R}^n)$ denotes the space of functions φ on \mathbb{R}^n satisfying

$$|\varphi(x) - \varphi(y)| \leq C(1 + |x|^m + |y|^m)(|x - y|), \quad \forall x, y \in \mathbb{R}^n,$$

for some constant $C > 0$ and $m \in \mathbb{N}$ depending on φ . \mathcal{H} can be regarded as the space of random variables.

Definition 1. A sublinear expectation \mathbb{E} on (Ω, \mathcal{H}) is a functional $\mathbb{E} : \mathcal{H} \rightarrow \mathbb{R}$ such that

- (1) if $X \leq Y$, then $\mathbb{E}[X] \leq \mathbb{E}[Y]$;
- (2) for any constant c , $\mathbb{E}[c] = c$;
- (3) for any $X, Y \in \mathcal{H}$, $\mathbb{E}[X + Y] \leq \mathbb{E}[X] + \mathbb{E}[Y]$;
- (4) for any $\lambda \geq 0$ and $X \in \mathcal{H}$, $\mathbb{E}[\lambda X] = \lambda \mathbb{E}[X]$.

The triple $(\Omega, \mathcal{H}, \mathbb{E})$ is called a sublinear expectation space.

The relation between sublinear expectations and linear expectations, which was proved by Peng [27], is contained in the following representation theorem.

Theorem 1. Let $(\Omega, \mathcal{H}, \mathbb{E})$ be a sublinear expectation space. Then there exists a family of linear expectations (linear functionals) $\{\mathbb{E}_\theta : \theta \in \Theta\}$ on \mathcal{H} , such that

$$\mathbb{E}[X] = \sup_{\theta \in \Theta} \mathbb{E}_\theta[X], \quad \forall X \in \mathcal{H}.$$

Under the frame work of sublinear expectation space, we also have the notion of independence and distribution (law).

Definition 2. (1) A random vector $Y \in \mathcal{H}^n$ is said to be independent from another random vector $X \in \mathcal{H}^m$ under the sublinear expectation \mathbb{E} , if for any $\varphi \in C_{l,Lip}(\mathbb{R}^m \times \mathbb{R}^n)$,

$$\mathbb{E}[\varphi(X, Y)] = \mathbb{E}[\mathbb{E}[\varphi(x, Y)]_{x=X}].$$

(2) Given a random vector $X \in \mathcal{H}^n$, the distribution (or the law) of X is defined as the sublinear expectation

$$\mathbb{F}_X[\varphi] := \mathbb{E}[\varphi(X)], \quad \varphi \in C_{l,Lip}(\mathbb{R}^n),$$

on $(\mathbb{R}^n, C_{l,Lip}(\mathbb{R}^n))$. By saying that two random vectors X, Y (possibly defined on different sublinear expectation spaces) are identically distributed, we mean that their distributions are the same.

Now we introduce the notion of G -distribution, which is the generalization of degenerate distributions and normal distributions. It captures the uncertainty of probability distributions and plays a fundamental role in the theory of sublinear expectation.

Let $S(d)$ be the space of $d \times d$ symmetric matrices, and let $G : \mathbb{R}^d \times S(d) \rightarrow \mathbb{R}$ be a continuous and sublinear function monotonic in $S(d)$ in the sense that:

- (1) $G(p + \bar{p}, A + \bar{A}) \leq G(p, A) + G(\bar{p}, \bar{A})$, $\forall p, \bar{p} \in \mathbb{R}^d, A, \bar{A} \in S(d)$;
- (2) $G(\lambda p, \lambda A) = \lambda G(p, A)$, $\forall \lambda \geq 0$;
- (3) $G(p, A) \leq G(p, \bar{A})$, $\forall A \leq \bar{A}$.

Definition 3. Let $X, \eta \in \mathcal{H}^d$ be two random vectors. (X, η) is called G -distributed if for any $\varphi \in C_{l,Lip}(\mathbb{R}^d \times \mathbb{R}^d)$, the function

$$u(t, x, y) := \mathbb{E}[\varphi(x + \sqrt{t}X, y + t\eta)], \quad (t, x, y) \in [0, \infty) \times \mathbb{R}^d \times \mathbb{R}^d,$$

is a viscosity solution of the following parabolic PDE (called a G -heat equation):

$$\partial_t u - G(D_y u, D_x^2 u) = 0, \quad (1)$$

with Cauchy condition $u|_{t=0} = \varphi$.

Remark 1. From the general theory of viscosity solutions (see [4], [27]), the G -heat equation (1) has a unique viscosity solution. By solving the G -heat equation (1) (in some special cases, it is explicitly solvable), we can compute the sublinear expectation of some functionals of a G -distributed random vector. The case of convex functionals, for instance, the power function $|x|^k$, is quite interesting.

It can be proved that for such a function G , there exists a bounded, closed and convex subset $\Gamma \subset \mathbb{R}^d \times \mathbb{R}^{d \times d}$, such that G has the following representation:

$$G(p, A) = \sup_{(q, Q) \in \Gamma} \left\{ \frac{1}{2} \text{tr}(AQQ^T) + \langle p, q \rangle \right\}, \quad \forall (p, A) \in \mathbb{R}^d \times S(d).$$

The set Γ captures the uncertainty of probability distribution (mean uncertainty and variance uncertainty) of a G -distributed random vector.

In particular, if G only depends on $p \in \mathbb{R}^d$, then there exists some bounded, closed and convex subset $\Lambda \subset \mathbb{R}^d$, such that

$$G(p) = \sup_{q \in \Lambda} \langle p, q \rangle.$$

In this case a G -distributed random vector η is called maximal distributed and is denoted by $\eta \sim N(\Lambda, \{0\})$. Similarly, if G only depends on $A \in S(d)$, then there exists some bounded, closed and convex subset $\Sigma \subset S_+(d)$ (the space of symmetric and nonnegative definite matrices) such that

$$G(A) = \frac{1}{2} \sup_{B \in \Sigma} \text{tr}(AB), \quad \forall A \in S(d). \quad (2)$$

A G -distributed random vector X for such G is called G -normal distributed and is denoted by $X \sim N(\{0\}, \Sigma)$.

Now we introduce the concept of G -Brownian motion and related stochastic calculus.

From now on, let $G : S(d) \rightarrow \mathbb{R}$ be a function given by (2).

Definition 4. A d -dimensional process B_t is called a G -Brownian motion if

- (1) $B_0(\omega) = 0, \forall \omega \in \Omega$;
- (2) for each $s, t \geq 0, B_{t+s} - B_t \sim N(\{0\}, s\Sigma)$ and is independent from

$$(B_{t_1}, \dots, B_{t_n})$$

for any $n \geq 1$ and $0 \leq t_1 < \dots < t_n \leq t$.

Similar to the classical situation, a G -Brownian motion can be constructed explicitly on the canonical path space by using independent G -normal random vectors. We refer the readers to [27] for a detailed construction.

In summary, let $\Omega = C_0([0, \infty); \mathbb{R}^d)$ be the space of \mathbb{R}^d -valued continuous paths starting at the origin, and let $B_t(\omega) := \omega_t$ be the coordinate process. For any $T \geq 0$, define

$$Lip(\Omega_T) := \{\varphi(B_{t_1}, \dots, B_{t_n}) : n \geq 1, t_1, \dots, t_n \in [0, T], \varphi \in C_{l, Lip}(\mathbb{R}^{d \times n})\},$$

and

$$Lip(\Omega) := \bigcup_{n=1}^{\infty} Lip(\Omega_n).$$

Then on $(\Omega, Lip(\Omega))$ we can define the canonical sublinear expectation \mathbb{E} such that the coordinate process B_t becomes a G -Brownian motion, which is usually called the G -expectation and denoted by \mathbb{E}^G . $(\Omega, Lip(\Omega), \mathbb{E}^G)$ is also called the canonical G -expectation space. Throughout the rest of this article, we will restrict ourselves on the canonical G -expectation space and its completion (to be defined later on).

On $(\Omega, Lip(\Omega), \mathbb{E}^G)$ we can introduce the notion of conditional G -expectation. More precisely, for

$$X = \varphi(B_{t_1}, B_{t_2} - B_{t_1}, \dots, B_{t_n} - B_{t_{n-1}}) \in Lip(\Omega),$$

where $0 \leq t_1 < t_2 < \dots < t_n$, the G -conditional expectation of X under Ω_{t_j} is defined by

$$\mathbb{E}^G[X | \Omega_{t_j}] := \psi(B_{t_1}, B_{t_2} - B_{t_1}, \dots, B_{t_j} - B_{t_{j-1}}),$$

where

$$\psi(x_1, \dots, x_j) := \mathbb{E}^G[\varphi(x_1, \dots, x_j, B_{t_{j+1}} - B_{t_j}, \dots, B_{t_n} - B_{t_{n-1}})], \quad x_1, \dots, x_j \in \mathbb{R}^d.$$

The conditional G -expectation $\mathbb{E}^G[\cdot|\Omega_t]$ has the following properties: for any $X, Y \in L_{ip}(\Omega)$,

- (1) if $X \leq Y$, then $\mathbb{E}^G[X|\Omega_t] \leq \mathbb{E}^G[Y|\Omega_t]$;
- (2) $\mathbb{E}^G[X+Y|\Omega_t] \leq \mathbb{E}^G[X|\Omega_t] + \mathbb{E}^G[Y|\Omega_t]$;
- (3) for any $\eta \in L_{ip}(\Omega_t)$,

$$\begin{aligned}\mathbb{E}^G[\eta|\Omega_t] &= \eta, \\ \mathbb{E}^G[\eta X|\Omega_t] &= \eta^+ \mathbb{E}^G[X|\Omega_t] + \eta^- \mathbb{E}[-X|\Omega_t];\end{aligned}$$

- (4) $\mathbb{E}^G[\mathbb{E}^G[X|\Omega_t]|\Omega_s] = \mathbb{E}^G[X|\Omega_{t \wedge s}]$. In particular, $\mathbb{E}^G[\mathbb{E}^G[X|\Omega_t]] = \mathbb{E}^G[X]$.

For any $p \geq 1$, let L_G^p (respectively, $L_G^p(\Omega_t)$) be the completion of $L_{ip}(\Omega)$ (respectively, $L_{ip}(\Omega_t)$) under the semi-norm $\|X\|_p := (\mathbb{E}^G[|X|^p])^{\frac{1}{p}}$. Then \mathbb{E}^G can be continuously extended to a sublinear expectation on $L_G^p(\Omega)$ (respectively, $L_G^p(\Omega_t)$), still denoted by \mathbb{E}^G .

For $t < T \leq \infty$, the conditional G -expectation $\mathbb{E}^G[\cdot|\Omega_t] : L_{ip}(\Omega_T) \rightarrow L_{ip}(\Omega_t)$ is a continuous mapping under $\|\cdot\|_1$ and can be continuously extended to a mapping

$$\mathbb{E}^G[\cdot|\Omega_t] : L_G^1(\Omega_T) \rightarrow L_G^1(\Omega_t),$$

which can still be interpreted as the conditional G -expectation. It is easy to show that the properties (1) to (4) for the conditional G -expectation still hold true on $L_G^1(\Omega_T)$ as long as it is well-defined.

Now we introduce the related stochastic calculus for G -Brownian motion and (Itô type) stochastic differential equations (SDEs) driven by G -Brownian motion.

First of all, similar to the idea in the classical case, we still have the notion of Itô integral with respect to a 1-dimensional G -Brownian motion. More precisely, consider $d = 1$, we can first define Itô integral of simple processes and then pass limit under the G -expectation \mathbb{E}^G in some suitable functional spaces. Let $M_G^{p,0}(0, T)$ be the space of simple processes $\eta_t(\omega)$ on $[0, T]$ of the form

$$\eta_t(\omega) = \sum_{k=1}^N \xi_{k-1}(\omega) \mathbf{1}_{[t_{k-1}, t_k)}(t),$$

where $\pi_T^N := \{t_0, t_1, \dots, t_N\}$ is a partition of $[0, T]$ and $\xi_k \in L_G^p(\Omega_{t_k})$, and introduce the semi-norm

$$\|\eta\|_{M_G^p(0, T)} := (\mathbb{E}^G[\int_0^T |\eta_t|^p dt])^{\frac{1}{p}}$$

on $M_G^{p,0}(0, T)$. Let $M_G^p(0, T)$ be the completion of $M_G^{p,0}(0, T)$ under $\|\cdot\|_{M_G^p(0, T)}$. It is straight forward to define Itô integral $\int_0^T \eta_t dB_t$ of simple processes. Moreover, such an integral operator is linear and continuous under $\|\cdot\|_{M_G^p(0, T)}$ and hence can be extended to a bounded linear operator

$$I : M_G^2(0, T) \rightarrow L_G^2(0, T).$$

The operator I is defined as the Itô integral operator with respect to a G -Brownian motion. For $0 \leq s < t \leq T$, define

$$\int_s^t \eta_u dB_u := \int_0^T \mathbf{1}_{[s,t]}(u) \eta_u dB_u.$$

We list some important properties of G -Itô integral in the following.

Proposition 1. *Let $\eta, \theta \in M_G^2(0, T)$ and let $0 \leq s \leq r \leq t \leq T$. Then*

(1)

$$\int_s^t \eta_u dB_u = \int_s^r \eta_u dB_u + \int_r^t \eta_u dB_u;$$

(2) if α is bounded in $L_G^1(\Omega_s)$, then

$$\int_s^t (\alpha \eta_u + \theta_u) dB_u = \alpha \int_s^t \eta_u dB_u + \int_s^t \theta_u dB_u;$$

(3) for any $X \in L_G^1(\Omega)$,

$$\mathbb{E}^G[X + \int_r^T \eta_u dB_u | \Omega_s] = \mathbb{E}^G[X | \Omega_s];$$

(4)

$$\underline{\sigma}^2 \mathbb{E}^G[\int_0^T \eta_t^2 dt] \leq \mathbb{E}^G[(\int_0^T \eta_t dB_t)^2] \leq \overline{\sigma}^2 \mathbb{E}^G[\int_0^T \eta_t^2 dt],$$

where $\overline{\sigma}^2 := \mathbb{E}^G[B_1^2]$ and $\underline{\sigma}^2 := -\mathbb{E}^G[-B_1^2]$.

Secondly, we have the notion of quadratic variation process of G -Brownian motion. In the case of 1-dimensional G -Brownian motion, the quadratic variation process $\langle B \rangle_t$ is defined as

$$\langle B \rangle_t := B_t^2 - 2 \int_0^t B_s dB_s,$$

which can be regarded as the L_G^2 -limit of the sum $\sum_{j=1}^{k_N} (B_{t_j^N} - B_{t_{j-1}^N})^2$ as $\mu(\pi_t^N) \rightarrow 0$, where $\pi_t^N := \{t_j^N\}_{j=0}^{k_N}$ is a sequence of partitions of $[0, t]$ and

$$\mu(\pi_t^N) := \max\{t_j^N - t_{j-1}^N : j = 1, 2, \dots, k_N\}.$$

It follows that $\langle B \rangle_t$ is an increasing process with $\langle B \rangle_0 = 0$.

Similar to the definition of G -Itô integral, we can define the integration with respect to $\langle B \rangle_t$ where B_t is a 1-dimensional G -Brownian motion. We refer the readers to [27] for a detailed construction but we remark that the integral operator with respect to $\langle B \rangle_t$ is a continuous linear mapping

$$Q_{0,T} : M_G^1(0, T) \rightarrow L_G^1(\Omega_T).$$

The following identity can be regarded as the G -Itô isometry.

Proposition 2. *Let $\eta \in M_G^2(0, T)$, then*

$$\mathbb{E}^G[(\int_0^T \eta_t dB_t)^2] = \mathbb{E}^G[\int_0^T \eta_t^2 d\langle B \rangle_t].$$

Now consider the multi-dimensional case. Let B_t is a d -dimensional G -Brownian motion, and for any $v \in \mathbb{R}^d$, denote

$$B_t^v := \langle v, B_t \rangle,$$

where $\langle \cdot, \cdot \rangle$ is the Euclidean inner product. Then for $a, \bar{a} \in \mathbb{R}^d$, the cross variation process $\langle B^a, B^{\bar{a}} \rangle_t$ is defined as

$$\langle B^a, B^{\bar{a}} \rangle_t = \frac{1}{4}(\langle B^{a+\bar{a}}, B^{a+\bar{a}} \rangle_t - \langle B^{a-\bar{a}}, B^{a-\bar{a}} \rangle_t).$$

Similar to the case of quadratic variation process, we have

$$\begin{aligned} \langle B^a, B^{\bar{a}} \rangle_t &= (L_G^2 -) \lim_{\mu(\pi_t^N) \rightarrow 0} \sum_{j=1}^{k_N} (B_{t_j^N}^a - B_{t_{j-1}^N}^a)(B_{t_j^N}^{\bar{a}} - B_{t_{j-1}^N}^{\bar{a}}) \\ &= B_t^a B_t^{\bar{a}} - \int_0^t B_s^a dB_s^{\bar{a}} - \int_0^t B_s^{\bar{a}} dB_s^a. \end{aligned}$$

Note that unlike the classical case, the cross variation process is not deterministic. The following results characterizes the distribution of $\langle B \rangle_t := (\langle B^\alpha, B^\beta \rangle_t)_{\alpha, \beta=1}^d$, where B_t is a d -dimensional G -Brownian motion and B_t^α is the α -th component of B_t .

Proposition 3. *Recall that the function G has the representation (2). Then $\langle B \rangle_t \sim N(t\Sigma, \{0\})$.*

As in the classical case, we also have the important G -Itô formula under G -expectation, which takes a similar form to the classical one. The main difference is that $dB_t^\alpha \cdot dB_t^\beta$ should be $d\langle B^\alpha, B^\beta \rangle_t$ instead of $\delta_{\alpha\beta} dt$. We are not going to state the full result of G -Itô formula here. See [27] for a detailed discussion.

Now we introduce the notion of SDEs driven by G -Brownian motion.

For $p \geq 1$, let $\overline{M}_G^p(0, T; \mathbb{R}^n)$ be the completion of $M_G^{p,0}(0, T; \mathbb{R}^n)$ under the norm

$$\|\eta\|_{\overline{M}_G^p(0, T; \mathbb{R}^n)} := (\int_0^T \mathbb{E}^G[|\eta_t|^p] dt)^{\frac{1}{p}}.$$

It is easy to see that $\overline{M}_G^p(0, T; \mathbb{R}^n) \subset M_G^p(0, T; \mathbb{R}^n)$.

Consider the following N -dimensional SDE driven by G -Brownian motion over $[0, T]$:

$$dX_t = b(t, X_t)dt + \sum_{\alpha, \beta=1}^d h_{\alpha\beta}(t, X_t)d\langle B^\alpha, B^\beta \rangle_t + \sum_{\alpha=1}^d V_\alpha(t, X_t)dB_t^\alpha \quad (3)$$

with initial condition $\xi \in \mathbb{R}^N$. Here we assume that the coefficients $b^i, h_{\alpha\beta}^i, V_\alpha^i$ are Lipschitz functions in the space variable, uniformly in time. A solution of (3) is a process in $\overline{M}_G^2(0, T; \mathbb{R}^N)$ satisfying the equation (3) in its integral form.

The existence and uniqueness of (3) was studied by Peng [27].

Theorem 2. *There exists a unique solution $X \in \overline{M}_G^2(0, T; \mathbb{R}^N)$ to the SDE (3).*

Finally, we introduce the notion of quasi-sure analysis for G -expectation. It plays an important role in studying pathwise properties of stochastic processes under the framework of G -expectation.

First of all, on the canonical sublinear expectation space $(\Omega, L_{ip}(\Omega), \mathbb{E}^G)$, we can prove a refinement of Theorem 1: there exists a weakly compact family \mathcal{P} of probability measures on $(\Omega, \mathcal{B}(\Omega))$, such that for any $X \in L_{ip}(\Omega)$ and $P \in \mathcal{P}$, $\mathbb{E}_P[X]$ is well-defined and

$$\mathbb{E}^G[X] = \max_{P \in \mathcal{P}} \mathbb{E}_P[X], \quad \forall X \in L_{ip}(\Omega),$$

where “max” means that the supremum is attainable (for each X). Moreover, there is an explicit characterization of the family \mathcal{P} . Let G be represented in the following way:

$$G(A) = \frac{1}{2} \sup_{Q \in \Gamma} \text{tr}(AQQ^T),$$

for some bounded, closed and convex subset $\Gamma \subset \mathbb{R}^{d \times d}$, and let \mathcal{A}_Γ be the collection of all Γ -valued $\{\mathcal{F}_t^W : t \geq 0\}$ -adapted processes on $[0, \infty)$, where $\{\mathcal{F}_t^W : t \geq 0\}$ is the natural filtration of the coordinate process on Ω . Let \mathcal{P}_0 be the collection of probability laws of the following classical Itô integral processes with respect to the standard Wiener measure:

$$B_t^\gamma := \int_0^t \gamma_s dW_s, \quad t \geq 0, \quad \gamma \in \mathcal{A}_\Gamma.$$

Then $\mathcal{P} = \overline{\mathcal{P}_0}$. For the proof of this result, please refer to [7].

For this particular family \mathcal{P} , define the set function c by

$$c(A) := \sup_{P \in \mathcal{P}} P(A), \quad A \in \mathcal{B}(\Omega).$$

Then we have the following result.

Theorem 3. *The set function c is a Choquet capacity (for an introduction of capacity theory, see [3], [6]). In other words,*

- (1) for any $A \in \mathcal{B}(\Omega)$, $0 \leq c(A) \leq 1$;
- (2) if $A \subset B$, then $c(A) \leq c(B)$;
- (3) if A_n is a sequence in $\mathcal{B}(\Omega)$, then $c(\cup_n A_n) \leq \sum_n c(A_n)$;

(4) if A_n is increasing in $\mathcal{B}(\Omega)$, then $c(\cup A_n) = \lim_{n \rightarrow \infty} c(A_n)$.

For any $\mathcal{B}(\Omega)$ -measurable random variable X such that $\mathbb{E}_P[X]$ is well-defined for all $P \in \mathcal{P}$, define the upper expectation

$$\hat{\mathbb{E}}[X] := \sup_{P \in \mathcal{P}} \mathbb{E}_P[X].$$

Then we can prove that for any $0 \leq T \leq \infty$ and $X \in L_G^1(\Omega_T)$,

$$\mathbb{E}^G[X] = \hat{\mathbb{E}}[X].$$

For a detailed discussion and other related properties, please refer to [7].

The following Markov inequality and Borel-Cantelli lemma under the capacity c are important for us.

Theorem 4. (1) For any $X \in L_G^p(\Omega)$ and $\lambda > 0$, we have

$$c(|X| > \lambda) \leq \frac{\mathbb{E}^G[|X|^p]}{\lambda^p}.$$

(2) Let A_n be a sequence in $\mathcal{B}(\Omega)$ such that

$$\sum_{n=1}^{\infty} c(A_n) < \infty.$$

Then

$$c(\limsup A_n) = 0.$$

Definition 5. A property depending on $\omega \in \Omega$ is said to hold quasi-surely, if it holds outside a $\mathcal{B}(\Omega)$ -measurable subset of zero capacity.

2.2 Rough Path Theory and Rough Differential Equations

Now we introduce some fundamentals in the theory of rough paths and rough differential equations. For a systematic introduction, please refer to [9], [18], [19].

For $n \geq 1$, define

$$T^{(\infty)}(\mathbb{R}^d) := \oplus_{k=0}^{\infty} (\mathbb{R}^d)^{\otimes k}$$

to be the infinite tensor algebra and

$$T^{(n)}(\mathbb{R}^d) := \oplus_{k=0}^n (\mathbb{R}^d)^{\otimes k}$$

to be the truncated tensor algebra of order n , equipped with the Euclidean norm. Let Δ be the triangle region $\{(s, t) : 0 \leq s < t \leq 1\}$. A functional $\mathbf{X} : \Delta \rightarrow T^{(n)}(\mathbb{R}^d)$ of order n is called multiplicative if for any $s < u < t$,

$$\mathbf{X}_{s,t} = \mathbf{X}_{s,u} \otimes \mathbf{X}_{u,t}.$$

Such a multiplicative structure is called the Chen identity. It describes the (nonlinear) additive structure of integrals over different intervals.

A control function ω is a nonnegative continuous function on Δ such that for any $s < u < t$,

$$\omega(s,u) + \omega(u,t) \leq \omega(s,t),$$

and for any $t \in [0,1]$, $\omega(t,t) = 0$. An example of control function $\omega(s,t)$ is the 1-variation norm over $[s,t]$ of a path with bounded variation.

Let $p \geq 1$ be a fixed constant. A continuous and multiplicative functional

$$\mathbf{X}_{s,t} = (1, X_{s,t}^1, \dots, X_{s,t}^n)$$

of order n has finite p -variation if for some control function ω ,

$$|X_{s,t}^i| \leq \omega(s,t)^{\frac{i}{p}}, \quad \forall i = 1, 2, \dots, n, (s,t) \in \Delta. \quad (4)$$

\mathbf{X} has finite p -variation if and only if for any $i = 1, 2, \dots, n$,

$$\sup_D \sum_l |X_{t_{l-1}, t_l}^i|^{\frac{p}{i}} < \infty,$$

where \sup_D runs over all finite partitions of $[0,1]$. We can also introduce the notion of finite p -variation for multiplicative functionals in $T^{(\infty)}(\mathbb{R}^d)$ by allowing $1 \leq i < \infty$ in (4). A continuous and multiplicative functional \mathbf{X} of order $[p]$ with finite p -variation is called a rough path with roughness p . The space of rough paths with roughness p is denoted by $\Omega_p(\mathbb{R}^d)$.

The following Lyons lifting theorem (see [17]) shows that the higher levels of a rough path \mathbf{X} with roughness p are uniquely determined by \mathbf{X} itself.

Theorem 5. *Let \mathbf{X} be a rough path with roughness p . Then \mathbf{X} can be uniquely extended to a continuous and multiplicative functional in $T^{(\infty)}(\mathbb{R}^d)$ with finite p -variation.*

One of the motivation of introducing the concept of rough paths is to develop the theory of differential equations driven by rough signals.

If an \mathbb{R}^d -valued path X has bounded variation, we know that the Picard iteration for the following differential equation converges:

$$dY_t = V(Y_t)dX_t, \quad (5)$$

where $V = (V_1, \dots, V_d)$ is a family of Lipschitz vector fields. Another way to consider (5) is to use the Euler scheme, which can be regarded as the Taylor expansion of functional of paths. Namely, we can write informally that

$$Y_t - Y_s \sim \sum_{n=1}^{\infty} \sum_{\alpha_1, \dots, \alpha_n=1}^d V_{\alpha_1} \cdots V_{\alpha_n} I(Y_s) \int_{s < u_1 < \dots < u_n < t} dX_{u_1}^{\alpha_1} \cdots dX_{u_n}^{\alpha_n}. \quad (6)$$

From (6) we can see that the sequence

$$\mathbf{X}_{s,t} := (1, X_t - X_s, \int_{s < u < v < t} dX_u \otimes dX_v, \dots, \int_{s < u_1 < \dots < u_n < t} dX_{u_1} \otimes \dots \otimes dX_{u_n}, \dots)$$

contains exactly all the information to determine the solution Y . On the other hand, it can be proved that \mathbf{X} is multiplicative and of finite 1-variation. Since X has bounded variation, it follows from Theorem 5 that \mathbf{X} is the unique enhancement of X . This is the fundamental reason why we don't need to see the higher levels when solving equation (5)-all information about \mathbf{X} , which uniquely determines the solution of (5), is incorporated in the first level.

If the driven signal is rougher, the situation becomes different. The same thing is that the information to determine the solution lies in the multiplicative structure in $T^{(\infty)}(\mathbb{R}^d)$, while the difference is that, unlike the case of paths with bounded variation, the classical path itself may not be able to determine the higher levels which are crucial to characterize the solution of a differential equation. In other words, we need to specify higher levels of the classical path in order to make sense of differential equations. According to Theorem 5, we know that the higher levels (levels above $[p]$) of a rough path \mathbf{X} with roughness p are uniquely determined by \mathbf{X} itself. Therefore, to establish differential equations driven by signals rougher than paths of bounded variation, we need to interpret the driven signal as a rough path with certain roughness p , that is, the driving signal should be an element in the space $\Omega_p(\mathbb{R}^d)$.

When the driving signal \mathbf{X} is in some smaller space of $\Omega_p(\mathbb{R}^d)$ in which \mathbf{X} can be approximated by paths of bounded variation in some sense, we are able to use a natural approximation procedure to introduce the notion of differential equations. But first we need to introduce a certain kind of topology.

Define the p -variation distance $d_p(\cdot, \cdot)$ on $\Omega_p(\mathbb{R}^d)$ by

$$d_p(\mathbf{X}, \mathbf{Y}) := \max_{1 \leq i \leq [p]} \sup_D \left(\sum_I |X_{t_{i-1}, t_i}^i - Y_{t_{i-1}, t_i}^i|^{\frac{p}{i}} \right)^{\frac{1}{p}}.$$

Then $(\Omega_p(\mathbb{R}^d), d_p)$ is a complete metric space.

A continuous path $X \in C([0, 1]; \mathbb{R}^d)$ is called smooth if it has bounded variation. Let

$$\Omega_p^\infty(\mathbb{R}^d) := \{\mathbf{X} : \mathbf{X} \text{ is the unique enhancement of } X \text{ in } T^{([p])}(\mathbb{R}^d), \text{ for smooth } X\}$$

be the subspace of enhanced smooth paths of order $[p]$. The closure of $\Omega_p^\infty(\mathbb{R}^d)$ under the p -variation distance d_p , denoted by $G\Omega_p(\mathbb{R}^d)$, is called the space of geometric rough paths with roughness p .

The following theorem, proved by Lyons [17], which is usually known as the universal limit theorem, enables us to introduce the notion of differential equations driven by geometric rough paths.

Theorem 6. Let $V_1, \dots, V_d \in C_b^{[p]+1}(\mathbb{R}^d)$ be given vector fields on \mathbb{R}^N . For a given $y \in \mathbb{R}^d$, define the mapping

$$F(y, \cdot) : \Omega_p^\infty(\mathbb{R}^d) \rightarrow \Omega_p^\infty(\mathbb{R}^N)$$

in the following way. For any $\mathbf{X} \in \Omega_p^\infty(\mathbb{R}^d)$, let X be the smooth path associated with \mathbf{X} starting at the origin (i.e., projection of \mathbf{X} onto the first level), and Y be the unique smooth path which is the solution of the following ODE:

$$dY_t = V_\alpha(Y_t) dX_t^\alpha$$

with $Y_0 = y$. $F(y, \mathbf{X})$ is defined to be the enhancement of Y in $\Omega_p^\infty(\mathbb{R}^d)$. Then the mapping $F(y, \cdot)$ is continuous with respect to the corresponding p -variation distance d_p .

According to Theorem 6, there exists a unique continuous extension of $F(y, \cdot)$ on $G\Omega_p(\mathbb{R}^d)$. The extended mapping

$$F(y, \cdot) : G\Omega_p(\mathbb{R}^d) \rightarrow G\Omega_p(\mathbb{R}^N),$$

is called the Itô-Lyons mapping. Such a mapping defines the (unique) solution in the space $G\Omega_p(\mathbb{R}^d)$ to the following differential equation:

$$dY_t = V(Y_t) dX_t, \tag{7}$$

with initial value y . Equation (7) is called a rough differential equation driven by \mathbf{X} (or simply called an RDE), and the solution \mathbf{Y} is called the full solution of (7). If we are only interested in classical paths, then

$$Y_t := y + \pi_1(\mathbf{Y}), \quad t \in [0, 1],$$

is called the solution of (7) with initial value y .

3 The Euler-Maruyama Approximation for SDEs Driven by G-Brownian Motion

In this section, we are going to establish the Euler-Maruyama approximation for SDEs driven by G -Brownian motion.

This result can be used to establish the Wong-Zakai type approximation which reveals the relation between SDEs (in the sense of $L_G^2(\Omega; \mathbb{R}^N)$ by S. Peng) and RDEs (in the sense of rough paths by Lyons) driven by G -Brownian motion. In Section 5, the study of such relation will be our main focus. However, based on the result in the next section which reveals the rough path nature of G -Brownian motion, we are going to use the rough Taylor expansion in the theory of RDEs instead of developing the Wong-Zakai type approximation to show that the solution of an SDE solves

some associated RDE with a correction term in terms of the cross variation process of multidimensional G -Brownian motion. Such approach reveals the natural of G -Brownian motion and differential equations in the sense of rough paths in a more fundamental way.

We also believe that there will be other interesting applications of the Euler-Maruyama approximation, such as in numerical analysis under G -expectation, and in mathematical finance under uncertainty.

Consider the following N -dimensional SDE driven by the canonical d -dimensional G -Brownian motion over $[0, 1]$ on the sublinear expectation space $(\Omega, L_G^2(\Omega), \mathbb{E}^G)$ which is the L_G^2 -completion of the canonical path space $(\Omega, Lip(\Omega), \mathbb{E}^G)$:

$$dX_t = b(X_t)dt + h_{\alpha\beta}(X_t)d\langle B^\alpha, B^\beta \rangle_t + V_\alpha(X_t)dB_t^\alpha, \quad (8)$$

with initial condition $X_0 = \xi \in \mathbb{R}^N$, where the coefficients $b^i, h_{\alpha\beta}^i, V_\alpha^i$ are bounded and uniformly Lipschitz. The existence and uniqueness of solution is studied by Peng [27].

The Euler-Maruyama approximation of the solution X_t of (8) is defined as follows.

For $n \geq 1$, consider the dyadic partition of the time interval $[0, 1]$, i.e.,

$$t_k^n = \frac{k}{2^n}, \quad k = 0, 1, \dots, 2^n.$$

Define X_t^n to be the approximation of X_t in the following evolutive way:

$$X_0^n = \xi,$$

and for $t \in [t_{k-1}^n, t_k^n]$,

$$(X_t^n)^i = (X_{t_{k-1}^n}^n)^i + V_\alpha^i(X_{t_{k-1}^n}^n)\Delta_k^n B^\alpha + b^i(X_{t_{k-1}^n}^n)\Delta t^n + h_{\alpha\beta}^i(X_{t_{k-1}^n}^n)\Delta_k^n \langle B^\alpha, B^\beta \rangle,$$

where

$$\begin{aligned} X_{t_{k-1}^n}^n &:= X_{t_{k-1}^n}^n, \quad \Delta_k^n B^\alpha := B_{t_k^n}^\alpha - B_{t_{k-1}^n}^\alpha, \quad \Delta t^n := \frac{1}{2^n}, \\ \Delta_k^n \langle B^\alpha, B^\beta \rangle &:= \langle B^\alpha, B^\beta \rangle_{t_k^n} - \langle B^\alpha, B^\beta \rangle_{t_{k-1}^n}. \end{aligned}$$

In this section, we are going to prove that X_t^n converges to the solution X_t of (8) in $L_G^2(\Omega; \mathbb{R}^N)$ with convergence rate 0.5, which coincides with the classical case when B_t reduces to a classical Brownian motion.

First of all, the following lemmas is useful for us.

Lemma 1. *Let η_t be a bounded process in $M_G^2(0, 1)$. Then for any $v \in \mathbb{R}^d$, $0 \leq s < t \leq 1$,*

$$\mathbb{E}^G[(\int_s^t \eta_u d\langle B^v \rangle_u)^2] \leq \bar{\sigma}_v^2(t-s)\mathbb{E}^G[\int_s^t \eta_u^2 d\langle B^v \rangle_u],$$

where $\bar{\sigma}_v^2 := 2G(v \cdot v^T)$ and $B^v := \langle v, B \rangle$, in which $\langle \cdot, \cdot \rangle$ denotes the Euclidean inner product of \mathbb{R}^d .

Proof. By approximation, it suffices to consider

$$\eta_u = \sum_{j=1}^k \zeta_{j-1} 1_{[u_{j-1}, u_j]},$$

where $s = u_0 < u_1 < \dots < u_k = t$ and $\zeta_j \in L_{ip}(\Omega_{u_j})$ are bounded. In this case, by definition

$$\int_s^t \eta_u d\langle B^v \rangle_u = \sum_{j=1}^k \zeta_{j-1} (\langle B^v \rangle_{u_j} - \langle B^v \rangle_{u_{j-1}}),$$

and

$$\int_s^t \eta_u^2 d\langle B^v \rangle_u = \sum_{j=1}^k \zeta_{j-1}^2 (\langle B^v \rangle_{u_j} - \langle B^v \rangle_{u_{j-1}}),$$

which are both defined in the pathwise sense for step functions. Since $\langle B^v \rangle$ is increasing, the Cauchy-Schwarz inequality yields that

$$\left(\int_s^t \eta_u d\langle B^v \rangle_u \right)^2 \leq (\langle B^v \rangle_t - \langle B^v \rangle_s) \cdot \int_s^t \eta_u^2 d\langle B^v \rangle_u.$$

Since ζ_j are bounded, if we use M to denote an upper bound of η_u^2 , it follows that for any $c \geq \bar{\sigma}_v^2$,

$$\begin{aligned} \left(\int_s^t \eta_u d\langle B^v \rangle_u \right)^2 &\leq M(\langle B^v \rangle_t - \langle B^v \rangle_s - c(t-s))^+ (\langle B^v \rangle_t - \langle B^v \rangle_s) \\ &\quad + c(t-s) \int_s^t \eta_u^2 d\langle B^v \rangle_u. \end{aligned}$$

Let $\varphi(x) = (x - c(t-s))^+ x$. Since $\langle B^v \rangle_t - \langle B^v \rangle_s$ is $N([\underline{\sigma}_v^2, \bar{\sigma}_v^2] \times \{0\})$ -distributed, it follows that

$$\begin{aligned} \mathbb{E}^G[\varphi(\langle B^v \rangle_t - \langle B^v \rangle_s)] &= \sup_{\underline{\sigma}_v^2 \leq x \leq \bar{\sigma}_v^2} \varphi(x(t-s)) \\ &= (t-s)^2 \sup_{\underline{\sigma}_v^2 \leq x \leq \bar{\sigma}_v^2} (x-c)^+ x \\ &= 0. \end{aligned}$$

Therefore, by the sub-linearity of G , we have

$$\mathbb{E}^G\left[\left(\int_s^t \eta_u d\langle B^v \rangle_u\right)^2\right] \leq c(t-s) \mathbb{E}^G\left[\int_s^t \eta_u^2 d\langle B^v \rangle_u\right], \quad c \geq \bar{\sigma}_v^2.$$

Now the proof is complete.

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

Theorem 7. *We have the following error estimate for the Euler-Maruyama approximation:*

$$\sup_{t \in [0,1]} \mathbb{E}^G [|X_t^n - X_t|^2] \leq C \Delta t^n,$$

where C is some positive constant only depending on d, N, G and the coefficients of (8). In particular,

$$\lim_{n \rightarrow \infty} \sup_{t \in [0,1]} \mathbb{E}^G [|X_t^n - X_t|^2] = 0.$$

Proof. For $t \in [t_{k-1}^n, t_k^n]$, by construction we have

$$X_t^i - (X_t^n)^i = I_1^i + J_1^i + K_1^i + I_2^i + J_2^i + K_2^i,$$

where

$$\begin{aligned} I_1^i &= \sum_{l=1}^{k-1} \int_{t_{l-1}^n}^{t_l^n} (V_\alpha^i(X_s) - V_\alpha^i(X_s^n)) dB_s^\alpha + \int_{t_{k-1}^n}^t (V_\alpha^i(X_s) - V_\alpha^i(X_s^n)) dB_s^\alpha, \\ J_1^i &= \sum_{l=1}^{k-1} \int_{t_{l-1}^n}^{t_l^n} (b^i(X_s) - b^i(X_s^n)) ds + \int_{t_{k-1}^n}^t (b^i(X_s) - b^i(X_s^n)) ds, \\ K_1^i &= \sum_{l=1}^{k-1} \int_{t_{l-1}^n}^{t_l^n} (h_{\alpha\beta}^i(X_s) - h_{\alpha\beta}^i(X_s^n)) d\langle B^\alpha, B^\beta \rangle_s \\ &\quad + \int_{t_{k-1}^n}^t (h_{\alpha\beta}^i(X_s) - h_{\alpha\beta}^i(X_s^n)) d\langle B^\alpha, B^\beta \rangle_s, \\ I_2^i &= \sum_{l=1}^{k-1} \int_{t_{l-1}^n}^{t_l^n} (V_\alpha^i(X_s^n) - V_\alpha^i(X_{t_{l-1}^n}^n)) dB_s^\alpha + \int_{t_{k-1}^n}^t (V_\alpha^i(X_s) - V_\alpha^i(X_{t_{l-1}^n}^n)) dB_s^\alpha, \\ J_2^i &= \sum_{l=1}^{k-1} \int_{t_{l-1}^n}^{t_l^n} (b^i(X_s^n) - b^i(X_{t_{l-1}^n}^n)) ds + \int_{t_{k-1}^n}^t (b^i(X_s) - b^i(X_{t_{l-1}^n}^n)) ds, \\ K_2^i &= \sum_{l=1}^{k-1} \int_{t_{l-1}^n}^{t_l^n} (h_{\alpha\beta}^i(X_s) - h_{\alpha\beta}^i(X_s^n)) d\langle B^\alpha, B^\beta \rangle_s \\ &\quad + \int_{t_{k-1}^n}^t (h_{\alpha\beta}^i(X_s) - h_{\alpha\beta}^i(X_s^n)) d\langle B^\alpha, B^\beta \rangle_s. \end{aligned}$$

It follows that

$$(X_t^i - (X_t^n)^i)^2 \leq 6((I_1^i)^2 + (J_1^i)^2 + (K_1^i)^2 + (I_2^i)^2 + (J_2^i)^2 + (K_2^i)^2). \quad (9)$$

Throughout the rest of this section, we will always use the same notation C to denote constants only depending on d, N, G and the coefficients of (8), although they may be different from line to line.

Now the following estimates are important for further development.

(1) From the G -Itô isometry, the distribution of $\langle B^\alpha \rangle$ and the Lipschitz property, we have,

$$\mathbb{E}^G[(\int_{t_{l-1}^n}^u (V_\alpha^i(X_s) - V_\alpha^i(X_s^n))dB_s^\alpha)^2] \leq C \int_{t_{l-1}^n}^u \mathbb{E}^G[|X_s - X_s^n|^2]ds, \forall u \in [t_{l-1}^n, t_l^n].$$

(2) Similarly, by Cauchy-Schwarz inequality, we have

$$\mathbb{E}^G[(\int_{t_{l-1}^n}^u (b^i(X_s) - b^i(X_s^n))ds)^2] \leq C(u - t_{l-1}^n) \int_{t_{l-1}^n}^u \mathbb{E}^G[|X_s - X_s^n|^2]ds, \forall u \in [t_{l-1}^n, t_l^n].$$

By the definition of $\langle B^\alpha, B^\beta \rangle$ and Lemma 1, we also have for any $u \in [t_{l-1}^n, t_l^n]$,

$$\mathbb{E}^G[(\int_{t_{l-1}^n}^u (h_{\alpha\beta}^i(X_s) - h_{\alpha\beta}^i(X_s^n))d\langle B^\alpha, B^\beta \rangle_s)^2] \leq C(u - t_{l-1}^n) \int_{t_{l-1}^n}^u \mathbb{E}^G[|X_s - X_s^n|^2]ds.$$

(3) By construction and similar arguments to (1), (2), we have

$$\begin{aligned} \mathbb{E}^G[(\int_{t_{l-1}^n}^u (V_\alpha^i(X_s^n) - V_\alpha^i(X_{t_{l-1}^n}^n))dB_s^\alpha)^2] &\leq C(u - t_{l-1}^n)^2, \\ \mathbb{E}^G[(\int_{t_{l-1}^n}^u (b^i(X_s^n) - b^i(X_{t_{l-1}^n}^n))ds)^2] &\leq C(u - t_{l-1}^n)^3, \\ \mathbb{E}^G[(\int_{t_{l-1}^n}^u (h_{\alpha\beta}^i(X_s^n) - h_{\alpha\beta}^i(X_{t_{l-1}^n}^n))d\langle B^\alpha, B^\beta \rangle_s)^2] &\leq C(u - t_{l-1}^n)^3, \end{aligned}$$

for all $u \in [t_{l-1}^n, t_l^n]$.

(4) By conditioning and from the properties of Itô integral with respect to G -Brownian motion, we know that the G -expectation of each “cross term” in $(I_1^i)^2$ and in $(I_2^i)^2$ is zero.

Combining (1) to (4) and applying the following elementary inequality to $(J_1^i)^2$, $(J_2^i)^2$, $(K_1^i)^2$ and $(K_2^i)^2$:

$$(a_1 + \cdots + a_m)^2 \leq m(a_1^2 + \cdots + a_m^2),$$

it is not hard to obtain that

$$\mathbb{E}^G[\|X_t - X_t^n\|^2] \leq C \int_0^t \mathbb{E}^G[\|X_s - X_s^n\|^2]ds + C(\Delta t^n), \forall t \in [0, 1].$$

By using Gronwall inequality, we arrive at

$$\mathbb{E}^G[\|X_t - X_t^n\|^2] \leq C(\Delta t^n),$$

which completes the proof of the theorem.

4 G -Brownian Motion as Rough Paths and RDEs Driven by G -Brownian Motion

In this section, we are going to study the nature of sample paths of G -Brownian motion under the framework of rough path theory. More precisely, we are going to show that: on the canonical path space, outside a Borel-measurable set of capacity zero, the sample paths of G -Brownian motion can be enhanced to the second level in a canonical way so that they become geometric rough paths with roughness $2 < p < 3$. As pointed out before, such a result will enable us to establish RDEs driven by G -Brownian motion in the space of geometric rough paths.

Recall that $(\Omega, L_{ip}(\Omega), \mathbb{E}^G)$ is the canonical path space associated with the function G , on which the coordinate process

$$B_t(\omega) := \omega_t, \quad t \in [0, 1],$$

is a d -dimensional G -Brownian motion with continuous sample paths.

By the following moment inequality for B_t :

$$\mathbb{E}^G[|B_t - B_s|^{2q}] \leq C_q(t-s)^q, \quad \forall 0 \leq s < t \leq 1, \quad q > 1, \quad (10)$$

and the generalized Kolmogorov criterion (see [27] for details), we know that for quasi-surely, the sample paths of B_t are α -Hölder continuous for any $\alpha \in (0, \frac{1}{2})$. Therefore, if the sample paths of B_t can be regarded as objects in the space of geometric rough paths, the correct roughness should be $2 < p < 3$ (so we should look for the enhancement of B_t to the second level); or in other words, the right topology we should work with is the p -variation topology induced by the p -variation distance d_p on the space of geometric rough paths with roughness $2 < p < 3$. The situation here is the same as the classical Brownian motion, and the fundamental reason behind lies in the distribution of B_t (or more precisely, the moment inequality (10)), which yields the same kind of Hölder continuity for sample paths of B_t as the classical one.

From now on, we will assume that $p \in (2, 3)$ is some fixed constant.

As in the last section, for $n \geq 1$, $k = 0, 1, \dots, 2^n$, let $t_k^n = \frac{k}{2^n}$ be the dyadic partition of $[0, 1]$, and let B_t^n be the piecewise linear approximation of B_t over the partition points $\{t_0^n, t_1^n, \dots, t_{2^n}^n\}$. Since the sample paths of B_t^n are smooth, B_t^n has a unique enhancement

$$B_{s,t}^n = (1, B_{s,t}^{n,1}, B_{s,t}^{n,2}), \quad 0 \leq s < t \leq 1,$$

to the space $G\Omega_p(\mathbb{R}^d)$ of geometric rough paths with roughness p (in fact, for any $p \geq 1$) determined by iterated integrals.

Our goal is to show that for quasi-surely, B^n is a Cauchy sequence under the p -variation distance d_p . It follows that for quasi-surely, the sample paths of B_t can be enhanced to the second level as geometric rough paths with roughness p , which are defined as limits of B^n under d_p . Such an enhancement can be regarded as a canonical lifting by using dyadic approximations.

Throughout the rest of this section, we will use $\|\cdot\|_q$ to denote the L^q -norm under the G -expectation \mathbb{E}^G . Moreover, we will use the same notation C to denote constants only depending on d, G, p , although they may be different from line to line.

The following estimates are crucial for the proof of the main result of this section.

Lemma 2. *Let $m, n \geq 1$, and $k = 1, 2, \dots, 2^n$. Then*

(1)

$$\|B_{t_{k-1}^n, t_k^n}^{m,j}\|_{\frac{p}{j}} \leq \begin{cases} C(\frac{1}{2})^j, & n \leq m; \\ C(\frac{2^{\frac{m}{2}}}{2^n})^j, & n > m, \end{cases}$$

where $j = 1, 2$.

(2)

$$\|B_{t_{k-1}^n, t_k^n}^{m+1,1} - B_{t_{k-1}^n, t_k^n}^{m,1}\|_p \leq \begin{cases} 0, & n \leq m; \\ C2^{\frac{m}{2}}, & n > m, \end{cases}$$

$$\|B_{t_{k-1}^n, t_k^n}^{m+1,2} - B_{t_{k-1}^n, t_k^n}^{m,2}\|_{\frac{p}{2}} \leq \begin{cases} C\frac{1}{2^{\frac{m}{2}}2^{\frac{m}{2}}}, & n \leq m; \\ C\frac{2^{\frac{m}{2}}}{2^n}, & n > m. \end{cases}$$

Here $\|\cdot\|_q$ denotes the L^q -norm under the G -expectation \mathbb{E}^G , and C is some positive constant not depending on m, n, k .

Proof. (1) The first level.

If $n \leq m$, then

$$B_{t_{k-1}^n, t_k^n}^{m,1} = B_{t_k^n} - B_{t_{k-1}^n}.$$

It follows from the moment inequality (10) that

$$\mathbb{E}^G[|B_{t_{k-1}^n, t_k^n}^{m,1}|^p] \leq C\frac{1}{2^{\frac{np}{2}}},$$

and thus

$$\|B_{t_{k-1}^n, t_k^n}^{m,1}\|_p \leq C\frac{1}{2^{\frac{n}{2}}}.$$

Also it is trivial to see that

$$B_{t_{k-1}^n, t_k^n}^{m+1,1} - B_{t_{k-1}^n, t_k^n}^{m,1} = (B_{t_k^n} - B_{t_{k-1}^n}) - (B_{t_k^n} - B_{t_{k-1}^n}) = 0.$$

If $n > m$, then by construction we know that

$$B_{t_{k-1}^n, t_k^n}^{m,1} = \frac{2^m}{2^n}(B_{t_l^m} - B_{t_{l-1}^m}),$$

where l is the unique integer such that $[t_{k-1}^n, t_k^n] \subset [t_{l-1}^m, t_l^m]$. Therefore,

$$\|B_{t_{k-1}^n, t_k^n}^{m,1}\|_p = \frac{2^m}{2^n} \|B_{t_l^m} - B_{t_{l-1}^m}\|_p \leq C\frac{2^{\frac{m}{2}}}{2^n}.$$

On the other hand, if $[t_{k-1}^n, t_k^n] \subset [t_{2l-2}^{m+1}, t_{2l-1}^{m+1}]$, then

$$\begin{aligned} B_{t_{k-1}^n, t_k^n}^{m+1,1} - B_{t_{k-1}^n, t_k^n}^{m,1} &= \frac{2^{m+1}}{2^n} (B_{t_{2l-1}^{m+1}} - B_{t_{2l-2}^{m+1}}) - \frac{2^m}{2^n} (B_{t_l^m} - B_{t_{l-1}^m}) \\ &= \frac{2^m}{2^n} ((B_{\frac{2l-1}{2^{m+1}}} - B_{\frac{2l-2}{2^{m+1}}}) - (B_{\frac{2l}{2^{m+1}}} - B_{\frac{2l-1}{2^{m+1}}})) \end{aligned}$$

It follows that

$$\|B_{t_{k-1}^n, t_k^n}^{m+1,1} - B_{t_{k-1}^n, t_k^n}^{m,1}\|_p \leq C \frac{2^{\frac{m}{2}}}{2^n}.$$

Similarly, if $[t_{k-1}^n, t_k^n] \subset [t_{2l-1}^{m+1}, t_{2l}^{m+1}]$, we will obtain the same estimate.

(2) The second level.

Since $\frac{\beta}{2} < 2$, by monotonicity it suffices to establish the desired estimates under the L^2 -norm.

First consider the term $B_{t_{k-1}^n, t_k^n}^{m+1,2} - B_{t_{k-1}^n, t_k^n}^{m,2}$.

If $n \leq m$, by the construction of $B_{s,t}^{m,2}$, we have

$$\begin{aligned} B_{t_{k-1}^n, t_k^n}^{m,2;\alpha,\beta} &= \int_{t_{k-1}^n < u < v < t_k^n} dB_u^\alpha dB_v^\beta \\ &= \int_{t_{k-1}^n}^{t_k^n} B_{t_{k-1}^n, v}^{m,1;\alpha} dB_v^{m,1;\beta} \\ &= \sum_{l=2^{(m-n)(k-1)+1}}^{2^{m-n}k} \frac{\Delta_l^m B^\beta}{\Delta t^m} \int_{t_{l-1}^m}^{t_l^m} \left(\frac{v-t_{l-1}^m}{\Delta t^m} B_{t_l^m}^\alpha + \frac{t_l^m-v}{\Delta t^m} B_{t_{l-1}^m}^\alpha - B_{t_{k-1}^n}^\alpha \right) dv \\ &= \sum_{l=2^{(m-n)(k-1)+1}}^{2^{m-n}k} \left(\frac{B_{t_{l-1}^m}^\alpha + B_{t_l^m}^\alpha}{2} - B_{t_{k-1}^n}^\alpha \right) \Delta_l^m B^\beta. \end{aligned}$$

Therefore,

$$\begin{aligned} &B_{t_{k-1}^n, t_k^n}^{m+1,2;\alpha,\beta} - B_{t_{k-1}^n, t_k^n}^{m,2;\alpha,\beta} \\ &= \sum_{l=2^{(m+1-n)(k-1)+1}}^{2^{m+1-n}k} \left(\frac{B_{t_{l-1}^{m+1}}^\alpha + B_{t_l^{m+1}}^\alpha}{2} - B_{t_{k-1}^n}^\alpha \right) \Delta_l^m B^\beta \\ &\quad - \sum_{l=2^{(m-n)(k-1)+1}}^{2^{m-n}k} \left(\frac{B_{t_{l-1}^m}^\alpha + B_{t_l^m}^\alpha}{2} - B_{t_{k-1}^n}^\alpha \right) \Delta_l^m B^\beta \\ &= \sum_{l=2^{(m-n)(k-1)+1}}^{2^{(m-n)k}} \left(\left(\frac{B_{t_{2l-2}^{m+1}}^\alpha + B_{t_{2l-1}^{m+1}}^\alpha}{2} - B_{t_{k-1}^n}^\alpha \right) \Delta_{2l-1}^{m+1} B^\beta + \left(\frac{B_{t_{2l-1}^{m+1}}^\alpha + B_{t_{2l}^{m+1}}^\alpha}{2} - B_{t_{k-1}^n}^\alpha \right) \right. \\ &\quad \left. \cdot \Delta_{2l}^{m+1} B^\beta - \left(\frac{B_{t_{2l-2}^m}^\alpha + B_{t_{2l}^m}^\alpha}{2} - B_{t_{k-1}^n}^\alpha \right) (\Delta_{2l-1}^{m+1} B^\beta + \Delta_{2l}^{m+1} B^\beta) \right) \end{aligned}$$

$$= \frac{1}{2} \sum_{l=2^{m-n}(k-1)+1}^{2^{m-n}k} (\Delta_{2l-1}^{m+1} B^\alpha \Delta_{2l}^{m+1} B^\beta - \Delta_{2l}^{m+1} B^\alpha \Delta_{2l-1}^{m+1} B^\beta).$$

By using the notation of tensor products, we have

$$B_{t_{k-1}^n, t_k^n}^{m+1,2} - B_{t_{k-1}^n, t_k^n}^{m,2} = \frac{1}{2} \sum_{l=2^{m-n}(k-1)+1}^{2^{m-n}k} (\Delta_{2l-1}^{m+1} B \otimes \Delta_{2l}^{m+1} B - \Delta_{2l}^{m+1} B \otimes \Delta_{2l-1}^{m+1} B).$$

It follows that

$$\begin{aligned} & \mathbb{E}^G [|B_{t_{k-1}^n, t_k^n}^{m+1,2} - B_{t_{k-1}^n, t_k^n}^{m,2}|^2] \\ &= \frac{1}{4} \mathbb{E}^G [|\sum_{l=2^{m-n}(k-1)+1}^{2^{m-n}k} (\Delta_{2l-1}^{m+1} B \otimes \Delta_{2l}^{m+1} B - \Delta_{2l}^{m+1} B \otimes \Delta_{2l-1}^{m+1} B)|^2] \\ &\leq C \sum_{\substack{\alpha \neq \beta \\ \alpha, \beta = 1, \dots, d}} \mathbb{E}^G [|\sum_l (\Delta_{2l-1}^{m+1} B^\alpha \Delta_{2l}^{m+1} B^\beta - \Delta_{2l}^{m+1} B^\alpha \Delta_{2l-1}^{m+1} B^\beta)|^2] \\ &\leq C \sum_{\alpha \neq \beta} \sum_{l,r} \mathbb{E}^G [(\Delta_{2l-1}^{m+1} B^\alpha \Delta_{2l}^{m+1} B^\beta - \Delta_{2l}^{m+1} B^\alpha \Delta_{2l-1}^{m+1} B^\beta) \\ &\quad \cdot (\Delta_{2r-1}^{m+1} B^\alpha \Delta_{2r}^{m+1} B^\beta - \Delta_{2r}^{m+1} B^\alpha \Delta_{2r-1}^{m+1} B^\beta)] \\ &\leq C \sum_{\alpha \neq \beta} \sum_{l,r} (\mathbb{E}^G [\Delta_{2l-1}^{m+1} B^\alpha \Delta_{2r-1}^{m+1} B^\alpha \Delta_{2l}^{m+1} B^\beta \Delta_{2r}^{m+1} B^\beta] \\ &\quad + \mathbb{E}^G [\Delta_{2l}^{m+1} B^\alpha \Delta_{2r}^{m+1} B^\alpha \Delta_{2l-1}^{m+1} B^\beta \Delta_{2r-1}^{m+1} B^\beta] \\ &\quad + \mathbb{E}^G [-\Delta_{2l-1}^{m+1} B^\alpha \Delta_{2r}^{m+1} B^\alpha \Delta_{2r-1}^{m+1} B^\beta \Delta_{2l}^{m+1} B^\beta] \\ &\quad + \mathbb{E}^G [-\Delta_{2r-1}^{m+1} B^\alpha \Delta_{2l}^{m+1} B^\alpha \Delta_{2l-1}^{m+1} B^\beta \Delta_{2r}^{m+1} B^\beta]), \end{aligned}$$

where the summation over l and r is taken from $2^{m-n}(k-1)+1$ to $2^{m-n}k$. Here we have used the sublinearity of \mathbb{E} . Now we study every term separately. If $l < r$, by the properties of conditional G -expectation and the distribution of B_t , we have

$$\begin{aligned} & \mathbb{E}^G [\Delta_{2l-1}^{m+1} B^\alpha \Delta_{2l}^{m+1} B^\beta \Delta_{2r-1}^{m+1} B^\alpha \Delta_{2r}^{m+1} B^\beta] \\ &= \mathbb{E}^G [\mathbb{E} [\Delta_{2l-1}^{m+1} B^\alpha \Delta_{2l}^{m+1} B^\beta \Delta_{2r-1}^{m+1} B^\alpha \Delta_{2r}^{m+1} B^\beta | \Omega_{t_{2r-1}^{m+1}}]] \\ &= \mathbb{E}^G [\eta^+ \mathbb{E} [\Delta_{2r}^{m+1} B^\beta | \Omega_{t_{2r-1}^{m+1}}] + \eta^- \mathbb{E} [-\Delta_{2r}^{m+1} B^\beta | \Omega_{t_{2r-1}^{m+1}}]] \\ &= 0, \end{aligned}$$

where $\eta = \Delta_{2l-1}^{m+1} B^\alpha \Delta_{2l}^{m+1} B^\beta \Delta_{2r-1}^{m+1} B^\alpha$. Similarly, we can prove that for any $l \neq r$,

$$\begin{aligned} & \mathbb{E}^G [\Delta_{2l-1}^{m+1} B^\alpha \Delta_{2r-1}^{m+1} B^\beta \Delta_{2l}^{m+1} B^\beta \Delta_{2r}^{m+1} B^\beta] \\ &= \mathbb{E}^G [(\Delta_{2l}^{m+1} B^\alpha \Delta_{2r}^{m+1} B^\alpha \Delta_{2l-1}^{m+1} B^\beta \Delta_{2r-1}^{m+1} B^\beta)] \end{aligned}$$

$$\begin{aligned}
&= \mathbb{E}^G [(-\Delta_{2l-1}^{m+1} B^\alpha \Delta_{2r}^{m+1} B^\alpha \Delta_{2r-1}^{m+1} B^\beta \Delta_{2l}^{m+1} B^\beta)] \\
&= \mathbb{E}^G [(-\Delta_{2r-1}^{m+1} B^\alpha \Delta_{2l}^{m+1} B^\alpha \Delta_{2l-1}^{m+1} B^\beta \Delta_{2r}^{m+1} B^\beta)] \\
&= 0.
\end{aligned}$$

On the other hand, if $l = r$, it is straight forward that

$$\mathbb{E}^G [(\Delta_{2l-1}^{m+1} B^\alpha)^2 (\Delta_{2l}^{m+1} B^\beta)^2] \leq \frac{1}{2} (\mathbb{E}^G [(\Delta_{2l-1}^{m+1} B^\alpha)^4] + \mathbb{E}^G [(\Delta_{2l}^{m+1} B^\beta)^4]) \leq C \frac{1}{2^{2m}},$$

and similarly,

$$\begin{aligned}
\mathbb{E}^G (-\Delta_{2l-1}^{m+1} B^\alpha \Delta_{2l-1}^{m+1} B^\beta \Delta_{2l}^{m+1} B^\alpha \Delta_{2l}^{m+1} B^\beta) &\leq \frac{1}{4} (\mathbb{E}^G [(\Delta_{2l-1}^{m+1} B^\alpha)^4] + \mathbb{E}^G [(\Delta_{2l-1}^{m+1} B^\beta)^4] \\
&\quad + \mathbb{E}^G [(\Delta_{2l}^{m+1} B^\alpha)^4] + \mathbb{E}^G [(\Delta_{2l}^{m+1} B^\beta)^4]) \\
&\leq C \frac{1}{2^{2m}}.
\end{aligned}$$

Combining all the estimates above, we arrive at

$$\mathbb{E}^G [|B_{t_{k-1}^n, t_k^n}^{m+1,2} - B_{t_{k-1}^n, t_k^n}^{m,2}|^2] \leq C \sum_{\alpha \neq \beta} \sum_{l=2^{m-n}(k-1)+1}^{2^{m-n}k} \frac{1}{2^{2m}} \leq C \frac{1}{2^m 2^n},$$

and hence

$$\|B_{t_{k-1}^n, t_k^n}^{m+1,2} - B_{t_{k-1}^n, t_k^n}^{m,2}\|_2 \leq C \frac{1}{2^{\frac{m}{2}} 2^{\frac{n}{2}}}.$$

If $n > m$, by construction we have

$$\begin{aligned}
B_{t_{k-1}^n, t_k^n}^{m,2;\alpha,\beta} &= \int_{t_{k-1}^n < u < v < t_k^n} d(B^m)_u^\alpha d(B^m)_v^\beta \\
&= \int_{t_{k-1}^n}^{t_k^n} B_{t_{k-1}^n, v}^{m,1;\alpha} d(B^m)_v^\beta \\
&= \frac{\Delta_l^m B^\alpha \Delta_l^m B^\beta}{(\Delta t^m)^2} \int_{t_{k-1}^n}^{t_k^n} (v - t_{k-1}^n) dv \\
&= \frac{1}{2} 2^{2(m-n)} \Delta_l^m B^\alpha \Delta_l^m B^\beta,
\end{aligned}$$

where l is the unique integer such that $[t_{k-1}^n, t_k^n] \subset [t_{l-1}^m, t_l^m]$. In other words, we have

$$B_{t_{k-1}^n, t_k^n}^{m,2} = \frac{1}{2} 2^{2(m-n)} (\Delta_l^m B)^{\otimes 2},$$

It follows that

$$\begin{aligned}
& B_{t_{k-1}^n, t_k^n}^{m+1,2} - B_{t_{k-1}^n, t_k^n}^{m,2} \\
&= \begin{cases} 2^{2(m-n)+1}(\Delta_{2^{l-1}}^{m+1}B)^{\otimes 2} - 2^{2(m-n)-1}(\Delta_l^m B)^{\otimes 2}, & [t_{k-1}^n, t_k^n] \subset [t_{2^{l-2}}^{m+1}, t_{2^{l-1}}^{m+1}]; \\ 2^{2(m-n)+1}(\Delta_{2^l}^{m+1}B)^{\otimes 2} - 2^{2(m-n)-1}(\Delta_l^m B)^{\otimes 2}, & [t_{k-1}^n, t_k^n] \subset [t_{2^{l-1}}^{m+1}, t_{2^l}^{m+1}]. \end{cases}
\end{aligned}$$

By using the Minkowski inequality, the Cauchy-Schwarz inequality and the sublinearity of \mathbb{E} , it is easy to obtain that

$$\|B_{t_{k-1}^n, t_k^n}^{m+1,2} - B_{t_{k-1}^n, t_k^n}^{m,2}\|_2 \leq C \frac{2^m}{2^{2n}}.$$

Now consider the term $B_{t_{k-1}^n, t_k^n}^{m,2}$.

If $n \geq m$, by using

$$B_{t_{k-1}^n, t_k^n}^{m,2} = 2^{2(m-n)-1}(\Delta_l^m B)^{\otimes 2},$$

we can proceed in the same way as before to obtain that

$$\|B_{t_{k-1}^n, t_k^n}^{m,2}\|_2 \leq C \frac{2^m}{2^{2n}}.$$

If $n < m$, then

$$B_{t_{k-1}^n, t_k^n}^{m,2} = \sum_{l=n+1}^m (B_{t_{k-1}^n, t_k^n}^{l,2} - B_{t_{k-1}^n, t_k^n}^{l-1,2}) + B_{t_{k-1}^n, t_k^n}^{n,2}.$$

It follows that

$$\begin{aligned}
\|B_{t_{k-1}^n, t_k^n}^{m,2}\|_2 &\leq \sum_{l=n+1}^m \|B_{t_{k-1}^n, t_k^n}^{l,2} - B_{t_{k-1}^n, t_k^n}^{l-1,2}\|_2 + \|B_{t_{k-1}^n, t_k^n}^{n,2}\|_2 \\
&\leq C \left(\frac{1}{2^{\frac{n}{2}}} \sum_{l=n+1}^{\infty} \frac{1}{2^{\frac{l}{2}}} + \frac{1}{2^n} \right) \\
&\leq C \frac{1}{2^n}.
\end{aligned}$$

Now the proof is complete.

In order to study the behavior of B^m in the space $G\Omega_p(\mathbb{R}^d)$, we may need to control the p -variation distance d_p in a suitable way. For $w, \tilde{w} \in G\Omega(\mathbb{R}^d)$, define

$$\rho_j(w, \tilde{w}) := \left(\sum_{n=1}^{\infty} n^\gamma \sum_{k=1}^{2^n} |w_{t_{k-1}^n, t_k^n}^j - \tilde{w}_{t_{k-1}^n, t_k^n}^j|^{\frac{p}{j}} \right)^{\frac{j}{p}}, \quad j = 1, 2, \quad (11)$$

where $\gamma > p - 1$ is some fixed universal constant. The functional ρ_j was initially introduced by Hambly and Lyons [11] to construct the stochastic area process associated with the Brownian motion on the Sierpinski gasket. We use $\rho_j(w)$ to denote $\rho_j(w, \tilde{w})$ with $\tilde{w} = (1, 0, 0)$.

The following result, which is important for us, is proved in [19].

Proposition 4. *There exists some positive constant $R = R(p, \gamma)$, such that for any $w, \tilde{w} \in G\Omega(\mathbb{R}^d)$,*

$$d_p(w, \tilde{w}) \leq R \max\{\rho_1(w, \tilde{w}), \rho_1(w, \tilde{w})(\rho_1(w) + \rho_1(\tilde{w})), \rho_2(w, \tilde{w})\}.$$

Now let

$$I(w, \tilde{w}) := \max\{\rho_1(w, \tilde{w}), \rho_1(w, \tilde{w})(\rho_1(w) + \rho_1(\tilde{w})), \rho_2(w, \tilde{w})\}, \quad (12)$$

and observe that

$$\begin{aligned} \{\omega : B^m \text{ is not Cauchy under } d_p\} &\subset \{\omega : \sum_{m=1}^{\infty} d_p(B^m, B^{m+1}) = \infty\} \\ &\subset \limsup_{m \rightarrow \infty} \{\omega : d_p(B^m, B^{m+1}) > \frac{R}{2^{m\beta}}\} \\ &\subset \limsup_{m \rightarrow \infty} \{\omega : I(B^m, B^{m+1}) > \frac{1}{2^{m\beta}}\}. \end{aligned} \quad (13)$$

where β is some positive constant to be chosen. Notice that the R.H.S. of (13) is $\mathcal{B}(\Omega)$ -measurable so its capacity is well-defined. Therefore, in order to prove that for quasi-surely, B^m is a Cauchy sequence under d_p , it suffices to show that the R.H.S. of (13) has capacity zero. This can be shown by using the Borel-Cantelli lemma.

According to (12), we may first need to establish estimates for

$$c(\rho_j(B^m, B^{m+1}) > \lambda), \quad j = 1, 2,$$

and

$$c(\rho_1(B^m) > \lambda),$$

where $m \geq 1$ and $\lambda > 0$. They are contained in the following lemma.

Lemma 3. *For $m \geq 1$, $\lambda > 0$, we have the following estimates.*

(1)

$$c(\rho_1(B^m) > \lambda) \leq C\lambda^{-p}.$$

(2) Let $\theta \in (0, \frac{p}{2} - 1)$ be some constant such that

$$n^{\gamma+1} \leq C \frac{2^{n(p-1)}}{2^{n(p-\theta-1)}}, \quad \forall n \geq 1.$$

Then we have

$$c(\rho_j(B^m, B^{m+1}) > \lambda) \leq C\lambda^{-\frac{p}{j}} \frac{1}{2^{m(\frac{p}{2}-\theta-1)}}, \quad j = 1, 2.$$

Proof. First consider

$$c(\rho_1(B^m) > \lambda) = c\left(\sum_{n=1}^{\infty} n^\gamma \sum_{k=1}^{2^n} |B_{t_{k-1}^n, t_k^n}^{m,1}|^p > \lambda^p\right).$$

Define

$$A_N = \left\{ \omega : \sum_{n=1}^N n^\gamma \sum_{k=1}^{2^n} |B_{t_{k-1}^n, t_k^n}^{m,1}|^p > \lambda^p \right\} \in \mathcal{B}(\Omega),$$

and

$$A = \left\{ \omega : \sum_{n=1}^{\infty} n^\gamma \sum_{k=1}^{2^n} |B_{t_{k-1}^n, t_k^n}^{m,1}|^p > \lambda^p \right\} \in \mathcal{B}(\Omega).$$

It is obvious that $A_N \uparrow A$. By the properties of the capacity c , we have

$$c(A) = \lim_{N \rightarrow \infty} c(A_N).$$

On the other hand, by the sublinearity of \mathbb{E}^G , the Chebyshev inequality for the capacity c and Lemma 10, we have

$$\begin{aligned} c(A_N) &\leq \lambda^{-p} \sum_{n=1}^N n^\gamma \sum_{k=1}^{2^n} \mathbb{E}[|B_{t_{k-1}^n, t_k^n}^{m,1}|^p] \\ &\leq C\lambda^{-p} \left[\sum_{n=1}^m n^\gamma 2^n \frac{1}{2^{\frac{np}{2}}} + \sum_{n=m+1}^{\infty} n^\gamma 2^n \frac{2^{\frac{mp}{2}}}{2^{np}} \right] \\ &= C\lambda^{-p} \left[\sum_{n=1}^m n^\gamma \frac{1}{2^{n(\frac{p}{2}-1)}} + 2^{\frac{mp}{2}} \sum_{n=m+1}^{\infty} n^\gamma \frac{1}{2^{n(p-1)}} \right] \\ &\leq C\lambda^{-p}. \end{aligned}$$

It follows that

$$c(\rho_1(B^m) > \lambda) = c(A) \leq C\lambda^{-p}.$$

Now consider

$$c(\rho_1(B^m, B^{m+1}) > \lambda) = c\left(\sum_{n=1}^{\infty} n^\gamma \sum_{k=1}^{2^n} |B_{\frac{k}{2^n}, \frac{k}{2^n}}^{(m+1),1} - B_{\frac{k}{2^n}, \frac{k}{2^n}}^{(m),1}|^p > \lambda^p\right).$$

By similar reasons we will have

$$\begin{aligned} c(\rho_1(B^m, B^{m+1}) > \lambda) &\leq \lambda^{-p} \sum_{n=1}^{\infty} n^\gamma \sum_{k=1}^{2^n} \mathbb{E}[|B_{t_{k-1}^n, t_k^n}^{m+1,1} - B_{t_{k-1}^n, t_k^n}^{m,1}|^p] \\ &\leq C\lambda^{-p} \left(\sum_{n=m+1}^{\infty} n^\gamma 2^n \frac{2^{\frac{mp}{2}}}{2^{np}} \right) \\ &= C\lambda^{-p} 2^{\frac{mp}{2}} \sum_{n=m+1}^{\infty} n^\gamma \frac{1}{2^{n(p-1)}}. \end{aligned}$$

Since $\theta \in (0, \frac{p}{2} - 1)$ is such that

$$n^{\gamma+1} \leq C \frac{2^{n(p-1)}}{2^{n(p-\theta-1)}}, \quad \forall n \geq 1,$$

we arrive at

$$c(\rho_1(B^m, B^{m+1}) > \lambda) \leq C \lambda^{-p} \frac{1}{2^{m(\frac{p}{2}-\theta-1)}}.$$

Finally, consider the second level part. By similar reasons, we have

$$\begin{aligned} c(\rho_2(B^m, B^{m+1}) > \lambda) &\leq C \lambda^{-\frac{p}{2}} \left[\sum_{n=1}^m n^\gamma 2^n \frac{1}{2^{\frac{mp}{4}} 2^{\frac{np}{4}}} + 2^{\frac{mp}{2}} \sum_{n=m+1}^{\infty} n^\gamma 2^n \frac{1}{2^{np}} \right] \\ &= C \lambda^{-\frac{p}{2}} \left[\frac{1}{2^{\frac{mp}{4}}} \sum_{n=1}^m n^\gamma 2^{n(1-\frac{p}{4})} + 2^{\frac{mp}{2}} \sum_{n=m+1}^{\infty} n^\gamma \frac{1}{2^{n(p-1)}} \right] \\ &\leq C \lambda^{-\frac{p}{2}} \left[\frac{1}{2^{\frac{mp}{4}}} m^{\gamma+1} 2^{m(1-\frac{p}{4})} + 2^{\frac{mp}{2}} \frac{1}{2^{m(p-\theta-1)}} \right] \\ &\leq C \lambda^{-\frac{p}{2}} \frac{1}{2^{m(\frac{p}{2}-\theta-1)}}. \end{aligned}$$

Now we are in position to prove the main result of this section.

Theorem 8. *Outside a $\mathcal{B}(\Omega)$ -measurable set of capacity zero, B^m is a Cauchy sequence under the p -variation distance d_p . In particular, for quasi-surely, the sample paths of B_t can be enhanced to be geometric rough paths*

$$B_{s,t} = (1, B_{s,t}^1, B_{s,t}^2), \quad 0 \leq s < t \leq 1,$$

with roughness p , which are defined as the limit of sample (geometric rough) paths of B^m in $G\Omega_p(\mathbb{R}^d)$ under the p -variation distance d_p .

Proof. By Lemma 3, we have

$$\begin{aligned} c(I(B^m, B^{m+1}) > \frac{1}{2^{m\beta}}) &\leq \sum_{j=1}^2 c(\rho_j(B^m, B^{m+1}) > \frac{1}{2^{m\beta}}) \\ &\quad + c(\rho_1(B^m, B^{m+1})(\rho_1(B^m) + \rho_1(B^{m+1})) > \frac{1}{2^{m\beta}}) \\ &\leq 2c(\rho_1(B^m, B^{m+1}) > \frac{1}{2^{2m\beta}}) + c(\rho_2(B^m, B^{m+1}) > \frac{1}{2^{m\beta}}) \\ &\quad + c(\rho_1(B^m) > \frac{2^{m\beta}}{2}) + c(\rho_1(B^{m+1}) > \frac{2^{m\beta}}{2}) \\ &\leq C \left[\frac{1}{2^{m\beta p}} + \frac{1}{2^{m(\frac{p}{2}-\theta-2\beta p-1)}} + \frac{1}{2^{m(\frac{p}{2}-\theta-\frac{\beta p}{2}-1)}} \right], \end{aligned}$$

where $\theta \in (0, \frac{p}{2} - 1)$ is some fixed constant.

If we choose β such that

$$0 < \beta < \frac{p-2\theta-2}{4p},$$

then

$$\sum_{m=1}^{\infty} c(I(B^m, B^{m+1})) > \frac{1}{2^{m\beta}} < \infty.$$

By the Borel-Cantelli lemma, we have

$$c(\limsup_{m \rightarrow \infty} \{\omega : I(B^m, B^{m+1}) > \frac{1}{2^{m\beta}}\}) = 0,$$

and the result follows from the inclusion (13).

With the help of Theorem 8 and the smoothness of $\langle B^\alpha, B^\beta \rangle_t$ (by definition the sample paths of $\langle B^\alpha, B^\beta \rangle_t$ are smooth), we are able to apply the universal limit theorem in rough path theory to define RDEs driven by G -Brownian motion in the pathwise sense. More precisely, consider the following N -dimensional RDE in the sense of rough paths:

$$dY_t = \tilde{b}(Y_t)dt + \tilde{h}_{\alpha\beta}(Y_t)d\langle B^\alpha, B^\beta \rangle_t + V_\alpha(Y_t)dB_t^\alpha, \quad (14)$$

with initial condition $Y_0 = x$, where $\tilde{b}, \tilde{h}_{\alpha\beta}, V_\alpha$ are C_b^3 -vector fields on \mathbb{R}^N . Then outside a $\mathcal{B}(\Omega)$ -measurable set of capacity zero, (14) has a unique full solution Y in $G\Omega_p(\mathbb{R}^N)$. Y is constructed as the limit of the enhancement of Y_t^n in $G\Omega_p(\mathbb{R}^N)$ under the p -variation distance, where Y_t^n is the unique classical solution of the following ordinary differential equation:

$$dY_t^n = \tilde{b}(Y_t^n)dt + \tilde{h}_{\alpha\beta}(Y_t^n)d\langle B^\alpha, B^\beta \rangle_t + V_\alpha(Y_t^n)d(B^n)_t^\alpha, \quad (15)$$

with $Y_0^n = x$, in which B_t^n is the dyadic piecewise linear approximation of B_t .

If we only consider solutions instead of full solutions (i.e., only consider the first level), then for quasi-surely, (14) has a unique solution $Y_t \in C([0, 1]; \mathbb{R}^N)$, which is constructed as the uniform limit of the solution of (15) with initial condition $Y_0^n = x$.

Before the end of this section, we are going to give an explicit description of the second level $B_{s,t}^2$ of B_t defined in Theorem 8, which reveals the nature of $B_{s,t}^2$ itself. Such result is fundamental to understand the relation between SDEs and RDEs driven by G -Brownian motion.

Lemma 4. *Assume that X_n converges to X in $L_G^2(\Omega)$ and converges to Y quasi-surely. Then for quasi-surely, $X = Y$.*

Proof. By the Chebyshev inequality for the capacity, we have

$$c(|X_n - X| > \varepsilon) \leq \frac{1}{\varepsilon^2} \mathbb{E}^G[|X_n - X|^2], \quad \forall \varepsilon > 0.$$

Since

$$X_n \rightarrow X \quad \text{in } L_G^2(\Omega),$$

we can extract a subsequence X_{n_k} , such that for any $k \geq 1$,

$$\mathbb{E}^G[|X_{n_k} - X|^2] \leq \frac{1}{k^4}.$$

It follows that

$$c(|X_{n_k} - X| > \frac{1}{k}) \leq \frac{1}{k^2}, \quad \forall k \geq 1,$$

and

$$\sum_{k=1}^{\infty} c(|X_{n_k} - X| > \frac{1}{k}) < \infty.$$

By the Borel-Cantelli lemma for the capacity, we arrive at for quasi-surely, X_{n_k} converges to X . By assumption it follows that for quasi-surely, $X = Y$.

The following result shows the nature of the second level of B_t . In the case when B_t reduces to the classical Brownian motion, it is essentially the relation between Stratonovich and Itô integrals.

Proposition 5. *Let $B_{s,t} = (1, B_{s,t}^1, B_{s,t}^2)$ be the quasi-surely defined enhancement of B_t in Theorem 8. Then for any $0 \leq s < t \leq 1$, for quasi-surely, we have*

$$B_{s,t}^{2;\alpha,\beta} = \int_s^t B_{s,u}^\alpha dB_u^\beta + \frac{1}{2} \langle B^\alpha, B^\beta \rangle_{s,t}, \quad (16)$$

where the integral on the R.H.S. of (16) is the Itô integral.

Proof. We know from Theorem 8 that for quasi-surely,

$$\lim_{n \rightarrow \infty} d_p(B^n, B) = 0.$$

From the definition of d_p , it is straight forward that for quasi-surely, $B_{s,t}^{n,2}$ converges uniformly to $B_{s,t}^2$.

Without loss of generality, we assume that s, t are both dyadic points in $[0, 1]$. It follows that when n is large enough,

$$\begin{aligned} B_{s,t}^{n,2;\alpha,\beta} &= \int_{s < u < v < t} d(B^n)_u^\alpha d(B^n)_v^\beta \\ &= \int_s^t (B^n)_{s,v}^\alpha d(B^n)_v^\beta \\ &= \sum_{k: [t_{k-1}^n, t_k^n] \subset [s,t]} \frac{\Delta_k^n B^\beta}{\Delta t^n} \int_{t_{k-1}^n}^{t_k^n} \left(\frac{v - t_{k-1}^n}{\Delta t^n} B_k^\alpha + \frac{t_k^n - v}{\Delta t^n} B_{k-1}^\alpha - B_s^\alpha \right) dv \\ &= \sum_{k: [t_{k-1}^n, t_k^n] \subset [s,t]} \left(\frac{B_{k-1}^\alpha + B_k^\alpha}{2} - B_s^\alpha \right) \Delta_k^n B^\beta \\ &= \sum_{k: [t_{k-1}^n, t_k^n] \subset [s,t]} (B_{k-1}^\alpha - B_s^\alpha) \Delta_k^n B^\beta + \frac{1}{2} \sum_k \Delta_k^n B^\alpha \Delta_k^n B^\beta. \end{aligned}$$

From properties of Itô integral and the cross-variation $\langle B^\alpha, B^\beta \rangle_t$, we know that the R.H.S. of the above equality converges to $\int_s^t B_{s,u}^\alpha dB_u^\beta + \frac{1}{2} \langle B^\alpha, B^\beta \rangle_{s,t}$ in $L_G^2(\Omega)$.

Consequently, by Lemma 4 $B_{s,t}^2$ must coincide with $\int_s^t B_{s,u}^\alpha dB_u^\beta + \frac{1}{2} \langle B^\alpha, B^\beta \rangle_{s,t}$ quasi-surely.

5 The Fundamental Relation between SDEs and RDEs Driven by G -Brownian Motion

So far we already know that there are two types of well-defined differential equations driven by G -Brownian motion: SDEs which are defined in the L_G^2 -sense with respect to the G -expectation \mathbb{E}^G , and RDEs which are quasi-surely defined in the pathwise sense. This section is devoted to the study of the fundamental relation between these two types of differential equations.

Consider the following N -dimensional SDE driven by G -Brownian motion on $(\Omega, L_G^2(\Omega), \mathbb{E})$:

$$dX_t = b(X_t)dt + h_{\alpha\beta}(X_t)d\langle B^\alpha, B^\beta \rangle_t + V_\alpha(X_t)dB_t^\alpha, \quad (17)$$

with initial condition $X_0 = x \in \mathbb{R}^N$. Here we assume that $b, h_{\alpha\beta}, V_\alpha$ are C_b^3 -vector fields on \mathbb{R}^N .

Our aim is to find the correct RDE of the form (14) whose strong solution coincides with X_t quasi-surely in the pathwise sense.

Let's first illustrate the idea in an informal way. We are going to use the rough Taylor expansion in the theory of RDEs (see Corollary 12.8 in [9]) and Proposition 5 to find the correct form of the RDE we are looking for.

Consider the following general RDE:

$$dY_t = \tilde{b}(Y_t)dt + \tilde{h}_{\alpha\beta}(Y_t)d\langle B^\alpha, B^\beta \rangle_t + \tilde{V}_\alpha(Y_t)dB_t^\alpha, \quad (18)$$

with initial condition $Y_0 = x$, where $\tilde{b}, \tilde{h}_{\alpha\beta}, \tilde{V}_\alpha$ are C_b^3 -vector fields on \mathbb{R}^N . By the smoothness of the cross variation process $\langle B^\alpha, B^\beta \rangle$, and the roughness of B_t studied in the last section, we know from the rough Taylor expansion theorem that for quasi-surely, for some control function $\omega(s, t)$, the solution Y_t of (18) satisfies, when $\omega(s, t) \leq 1$,

$$\begin{aligned} & |Y_{s,t} - \tilde{b}(Y_s)(t-s) - \tilde{h}_{\alpha\beta}(Y_s)\langle B^\alpha, B^\beta \rangle_{s,t} - \tilde{V}_\alpha(Y_s)B_{s,t}^{1;\alpha} - D\tilde{V}_\beta(Y_s) \cdot \tilde{V}_\alpha(Y_s)B_{s,t}^{2;\alpha;\beta}| \\ & \leq C\omega(s, t)^\theta, \end{aligned} \quad (19)$$

where $\omega(s, t)$ C and $\theta > 1$ are two constants not depending on s, t . Note that inequality (19) reveals the local behavior of the solution Y_t . It follows from Proposition 5 that for quasi-surely,

$$|Y_{s,t} - \tilde{I}_{s,t}| \leq C\omega(s, t)^\theta,$$

where

$$\begin{aligned} \tilde{I}_{s,t} := & \tilde{b}(Y_s)(t-s) + (\tilde{h}_{\alpha\beta}(Y_s) + \frac{1}{2}D\tilde{V}_\beta(Y_s) \cdot \tilde{V}_\alpha(Y_s))d\langle B^\alpha, B^\beta \rangle_t + \tilde{V}_\alpha(Y_s)B_{s,t}^{1;\alpha} \\ & + D\tilde{V}_\beta(Y_s) \cdot \tilde{V}_\alpha(Y_s) \int_s^t B_{s,u}^\alpha dB_u^\beta. \end{aligned} \quad (20)$$

Now if we consider the global behavior of Y_t , we may sum up inequality (20) over dyadic intervals $[t_{k-1}^n, t_k^n]$ and then take limit (in $L_G^2(\Omega; \mathbb{R}^N)$) to obtain that for quasi-surely,

$$\begin{aligned} Y_{s,t} = & \int_s^t \tilde{b}(Y_u)du + \int_s^t (\tilde{h}_{\alpha\beta}(Y_u) + \frac{1}{2}D\tilde{V}_\beta(Y_u) \cdot \tilde{V}_\alpha(Y_u))d\langle B^\alpha, B^\beta \rangle_u + \int_s^t \tilde{V}_\alpha(Y_u)dB_u^\alpha \\ & + (L_G^2-) \lim_{n \rightarrow \infty} \sum_{k: [t_{k-1}^n, t_k^n] \subset [s,t]} D\tilde{V}_\alpha(Y_{t_{k-1}^n}) \cdot \tilde{V}_\beta(Y_{t_{k-1}^n}) \int_{t_{k-1}^n}^{t_k^n} B_{t_{k-1}^n, u}^\alpha dB_u^\beta, \end{aligned} \quad (21)$$

where the integrals with respect to B_t are interpreted as Itô integrals. On the other hand, by the distribution of B_t and properties of G -Itô integral, it is not hard to prove that the L_G^2 -limit in the last term of the above identity is zero. Therefore, we know that Y_t solves the SDE

$$dX_t = \tilde{b}(X_t)dt + (\tilde{h}_{\alpha\beta}(X_t) + \frac{1}{2}D\tilde{V}_\beta(X_t) \cdot \tilde{V}_\alpha(X_t))d\langle B^\alpha, B^\beta \rangle_t + \tilde{V}_\alpha(X_t)dB_t^\alpha.$$

In other words, if X_t is the solution of the SDE (17), it is natural to expect that for quasi-surely, X_t is the solution of the following RDE:

$$dY_t = b(Y_t)dt + (h_{\alpha\beta}(Y_t) - \frac{1}{2}DV_\beta(Y_t) \cdot V_\alpha(Y_t))d\langle B^\alpha, B^\beta \rangle + V_\alpha(Y_t)dB_t^\alpha, \quad (22)$$

with the same initial condition.

In the remaining of this section, we are going to prove this claim in a rigorous way.

From now on, assume that X_t is the solution of the SDE (17) and Y_t is the solution of the RDE (22) with the same initial condition $x \in \mathbb{R}^N$, where the coefficients $b, h_{\alpha\beta}, V_\alpha$ are C_b^3 -vector fields on \mathbb{R}^N . For simplicity we will also use the same notation to denote constants only depending on d, N, G, p and the coefficients of (17), although they may be different from line to line.

The following lemma enables us to show that the L_G^2 -limit in the last term of the identity (21) is zero.

Lemma 5. *Let $f \in C_b(\mathbb{R}^N)$, and $s < t$ be two dyadic points in $[0, 1]$ (i.e., $s = t_k^m$ and $t = t_l^m$ for some m and $k < l$). Then for any $\alpha, \beta = 1, 2, \dots, d$,*

$$\lim_{n \rightarrow \infty} \mathbb{E}^G [(\sum_{k: [t_{k-1}^n, t_k^n] \subset [s,t]} f(Y_{t_{k-1}^n}) \int_{t_{k-1}^n}^{t_k^n} B_{t_{k-1}^n, u}^\alpha dB_u^\beta)^2] = 0.$$

Proof. From direct calculation, we have

$$\begin{aligned}
& \mathbb{E}^G \left[\left(\sum_{k: [t_{k-1}^n, t_k^n] \subset [s, t]} f(Y_{t_{k-1}^n}) \int_{t_{k-1}^n}^{t_k^n} B_{t_{k-1}^n, u}^\alpha dB_u^\beta \right)^2 \right] \\
& \leq \|f\|_\infty^2 \sum_{k: [t_{k-1}^n, t_k^n] \subset [s, t]} \mathbb{E}^G \left[\left(\int_{t_{k-1}^n}^{t_k^n} B_{t_{k-1}^n, u}^\alpha dB_u^\beta \right)^2 \right] \\
& + 2 \sum_{\substack{k < l \\ [t_{k-1}^n, t_k^n], [t_{l-1}^n, t_l^n] \subset [s, t]}} \mathbb{E}^G [f(Y_{t_{k-1}^n}) \left(\int_{t_{k-1}^n}^{t_k^n} B_{t_{k-1}^n, u}^\alpha dB_u^\beta \right) f(Y_{t_{l-1}^n}) \left(\int_{t_{l-1}^n}^{t_l^n} B_{t_{l-1}^n, u}^\alpha dB_u^\beta \right)] \\
& \leq C \|f\|_\infty^2 \sum_{k: [t_{k-1}^n, t_k^n] \subset [s, t]} (\Delta t^n)^2 \\
& + 2 \sum_{\substack{k < l \\ [t_{k-1}^n, t_k^n], [t_{l-1}^n, t_l^n] \subset [s, t]}} \left(\mathbb{E}^G [(f(Y_{t_{k-1}^n}) \left(\int_{t_{k-1}^n}^{t_k^n} B_{t_{k-1}^n, u}^\alpha dB_u^\beta \right) f(Y_{t_{l-1}^n}))^+ \right. \right. \\
& \cdot \mathbb{E}^G \left[\int_{t_{l-1}^n}^{t_l^n} B_{t_{l-1}^n, u}^\alpha dB_u^\beta \mid \Omega_{t_{l-1}^n} \right] + \mathbb{E}^G [(f(Y_{t_{k-1}^n}) \left(\int_{t_{k-1}^n}^{t_k^n} B_{t_{k-1}^n, u}^\alpha dB_u^\beta \right) f(Y_{t_{l-1}^n}))^- \\
& \cdot \mathbb{E}^G \left[- \int_{t_{l-1}^n}^{t_l^n} B_{t_{l-1}^n, u}^\alpha dB_u^\beta \mid \Omega_{t_{l-1}^n} \right]] \Big) \\
& \leq C \|f\|_\infty^2 \Delta t^n,
\end{aligned}$$

and the result follows easily.

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

Theorem 9. *For quasi-surely,*

$$X_t = Y_t, \quad \forall t \in [0, 1].$$

Proof. Since the coefficients of the RDE (22) are in $C_b^3(\mathbb{R}^N)$, for quasi-surely define the following pathwise control: for $0 \leq s < t \leq 1$,

$$\begin{aligned}
\omega(s, t) & := (\|V\|_{2, \infty} \|B\|_{p\text{-var}; [s, t]})^p + \|b\|_{1, \infty} (t - s) \\
& + \|h - \frac{1}{2} DV \cdot V\|_{1, \infty} \| \langle B, B \rangle \|_{1\text{-var}; [s, t]},
\end{aligned}$$

where $\|\cdot\|_{m, \infty}$ denotes the maximum of uniform norms of derivatives up to order m . It follows from the rough Taylor expansion (Corollary 12.8 [9]) that for quasi-surely, there exists some positive constant $\theta > 1$, such that for $0 \leq s < t \leq 1$, when $\omega(s, t) \leq 1$, we have

$$|Y_{s, t} - I_{s, t}| \leq C \omega(s, t)^\theta,$$

where

$$\begin{aligned}
I_{s,t} &= b(Y_s)(t-s) + (h_{\alpha\beta}(Y_s) - \frac{1}{2}DV_{\beta}(Y_s) \cdot V_{\alpha}(Y_s)) \langle B^{\alpha}, B^{\beta} \rangle_{s,t} + V_{\alpha}(Y_s) B_{s,t}^{1;\alpha} \\
&\quad + DV_{\beta}(Y_s) \cdot V_{\alpha}(Y_s) B_{s,t}^{2;\alpha,\beta}
\end{aligned}$$

By Proposition 5, we have for quasi-surely,

$$\begin{aligned}
&|Y_{s,t} - b(Y_s)(t-s) - h_{\alpha\beta}(Y_s) \langle B^{\alpha}, B^{\beta} \rangle_{s,t} - V_{\alpha}(Y_s) B_{s,t}^{1;\alpha} \\
&\quad - DV_{\beta}(Y_s) \cdot V_{\alpha}(Y_s) \int_s^t B_{s,u}^{\alpha} dB_u^{\beta}| \leq C\omega(s,t)^{\theta}. \tag{23}
\end{aligned}$$

Now consider fixed $s < t$ being two dyadic points in $[0, 1]$. When n is large enough, by applying inequality (23) on each small dyadic interval $[t_{k-1}^n, t_k^n] \subset [s, t]$ and summing up through the triangle inequality, we obtain that for quasi-surely,

$$\begin{aligned}
|Y_{s,t} - I_{s,t}^n| &\leq C \sum \omega(t_{k-1}^n, t_k^n)^{\theta} \\
&\leq C\omega(s,t) \max\{\omega(t_{k-1}^n, t_k^n)^{\theta-1} : [t_{k-1}^n, t_k^n] \subset [s, t]\},
\end{aligned}$$

where

$$\begin{aligned}
I_{s,t}^n &= \sum b(Y_{t_{k-1}^n}) \Delta t^n + \sum h_{\alpha\beta}(Y_{t_{k-1}^n}) \Delta_k^n \langle B^{\alpha}, B^{\beta} \rangle + \sum V_{\alpha}(Y_{t_{k-1}^n}) \Delta_k^n B^{\alpha} \\
&\quad + \sum DV_{\beta}(Y_{t_{k-1}^n}) \cdot V_{\alpha}(Y_{t_{k-1}^n}) \int_{t_{k-1}^n}^{t_k^n} B_{t_{k-1}^n, u}^{\alpha} dB_u^{\beta},
\end{aligned}$$

and each sum is over all k such that $[t_{k-1}^n, t_k^n] \subset [s, t]$. It follows that for quasi-surely,

$$I_{s,t}^n \rightarrow Y_{s,t}, \quad n \rightarrow \infty.$$

On the other hand, the following convergence in $L_G^2(\Omega; \mathbb{R}^N)$ holds:

$$\begin{aligned}
\sum b(Y_{t_{k-1}^n}) \Delta t^n &\rightarrow \int_s^t b(Y_u) du, \\
\sum h_{\alpha\beta}(Y_{t_{k-1}^n}) \Delta_k^n \langle B^{\alpha}, B^{\beta} \rangle &\rightarrow \int_s^t h_{\alpha\beta}(Y_u) d \langle B^{\alpha}, B^{\beta} \rangle_u, \\
\sum V_{\alpha}(Y_{t_{k-1}^n}) \Delta_k^n B^{\alpha} &\rightarrow \int_s^t V_{\alpha}(Y_u) dB_u^{\alpha},
\end{aligned}$$

as $n \rightarrow \infty$.

The reason is the following. For simplicity we only consider the third one, as the first two are similar (and in fact easier). It is straight forward that

$$\int_0^1 |V_{\alpha}(Y_t) - \sum_{k=1}^{2^n} V_{\alpha}(Y_{t_{k-1}^n}) \mathbf{1}_{[t_{k-1}^n, t_k^n)}(t)|^2 dt$$

$$\begin{aligned}
&= \sum_{k=1}^{2^n} \int_{t_{k-1}^n}^{t_k^n} |V_\alpha(Y_t) - V_\alpha(Y_{t_{k-1}^n})|^2 dt \\
&\leq C \sum_{k=1}^{2^n} \int_{t_{k-1}^n}^{t_k^n} |Y_t - Y_{t_{k-1}^n}|^2 dt \\
&\leq C \sum_{k=1}^{2^n} \|Y\|_{p\text{-var}; [t_{k-1}^n, t_k^n]}^2 \Delta t^n \\
&\leq C \left(\sum_{k=1}^{2^n} \|Y\|_{p\text{-var}; [t_{k-1}^n, t_k^n]}^p \Delta t^n \right)^{\frac{2}{p}} \\
&\leq C (\Delta t^n)^{\frac{2}{p}} \|Y\|_{p\text{-var}; [0,1]}^2,
\end{aligned}$$

where C depends only on V_α . Therefore, it suffices to show that $\|Y\|_{p\text{-var}; [0,1]} \in L_G^2(\Omega)$, as it will imply the G -Itô integrability of $V_\alpha(Y_t)$ and the desired convergence in $L_G^2(\Omega; \mathbb{R}^N)$ will hold. For simplicity we assume that Y_t is the solution of the following RDE

$$dY_t = V_\alpha(Y_t) dB_t^\alpha$$

with $Y_0 = \xi$ (there is no substantial difference because dt and $d\langle B^\alpha, B^\beta \rangle_t$ are more regular than dB_t), then by Theorem 10.14 in [9], we know that

$$\|Y\|_{p\text{-var}; [0,1]} \leq C \|B\|_{p\text{-var}; [0,1]} \vee \|B\|_{p\text{-var}; [0,1]}^p.$$

Therefore, we only need to show that $\|B\|_{p\text{-var}; [0,1]}^p \in L_G^2(\Omega)$. For this purpose, we use Proposition 4 to control the p -variation norm by the functions ρ_1, ρ_2 defined in (11). It follows that

$$\|B\|_{p\text{-var}} \leq C(1 + \rho_1(B)^2 + \rho_2(B)).$$

Therefore, it remains to show that $\rho_1(B)^{2p}, \rho_2(B)^p \in L_G^1(\Omega)$. First consider level one. By the distribution of B_t , we have

$$\begin{aligned}
\left\| \sum_{n=1}^{\infty} n^\gamma \sum_{k=1}^{2^n} |B_{t_{k-1}^n, t_k^n}^1|^p \right\|_2 &\leq \sum_{n=1}^{\infty} n^\gamma \sum_{k=1}^{2^n} \| |B_{t_{k-1}^n, t_k^n}^1|^p \|_2 \\
&\leq \sum_{n=1}^{\infty} n^\gamma (\Delta t^n)^{\frac{p}{2}-1} \\
&< \infty,
\end{aligned}$$

and we know that $\rho_1(B)^{2p} \in L_G^1(\Omega)$. Now consider level two. By Proposition 5 and the distribution of B_t and $\langle B, B \rangle_t$, we have

$$\left\| \sum_{n=1}^{\infty} n^\gamma \sum_{k=1}^{2^n} |B_{t_{k-1}^n, t_k^n}^2|^{\frac{p}{2}} \right\|_2 = \left\| \sum_{n=1}^{\infty} n^\gamma \sum_{k=1}^{2^n} \left| \int_{t_{k-1}^n}^{t_k^n} B_{t_{k-1}^n, u} \otimes dB_u + \frac{1}{2} \langle B, B \rangle_{t_{k-1}^n, t_k^n} \right|^{\frac{p}{2}} \right\|_2$$

$$\begin{aligned}
&\leq \sum_{n=1}^{\infty} n^{\gamma} \sum_{k=1}^{2^n} \left\| \int_{t_{k-1}^n}^{t_k^n} B_{t_{k-1},u}^n \otimes dB_u + \frac{1}{2} \langle B, B \rangle_{t_{k-1}^n, t_k^n} \right\|_2 \\
&\leq C \sum_{n=1}^{\infty} n^{\gamma} (\Delta t^n)^{\frac{\beta}{2}-1} \\
&< \infty.
\end{aligned}$$

It follows that $\rho_2(B)^p \in L_G^1(\Omega)$. Therefore, the desired L_G^2 -convergence holds.

In addition, by Lemma 5 we also have the following L_G^2 -convergence:

$$\sum DV_{\beta}(Y_{t_{k-1}^n}) \cdot V_{\alpha}(Y_{t_{k-1}^n}) \int_{t_{k-1}^n}^{t_k^n} B_{t_{k-1},u}^{\alpha} dB_u^{\beta} \rightarrow 0, \quad n \rightarrow \infty.$$

Consequently, in $L_G^2(\Omega; \mathbb{R}^N)$,

$$Y_{s,t}^n \rightarrow \int_s^t b(Y_u) du + \int_s^t h_{\alpha\beta}(Y_u) d\langle B^{\alpha}, B^{\beta} \rangle_u + \int_s^t V_{\alpha}(Y_u) dB_u^{\alpha},$$

as $n \rightarrow \infty$.

From Lemma 4, we conclude that for quasi-surely,

$$Y_{s,t} = \int_s^t b(Y_u) du + \int_s^t h_{\alpha\beta}(Y_u) d\langle B^{\alpha}, B^{\beta} \rangle_u + \int_s^t V_{\alpha}(Y_u) dB_u^{\alpha}.$$

Since X_t and Y_t are both quasi-surely continuous, it follows that X coincides with Y quasi-surely.

Remark 2. As we mentioned at the beginning of Section 2, it is possible to prove Theorem 9 by establishing the Wong-Zakai type approximation. More precisely, if we let X_t^n to be the Euler-Maruyama approximation of the SDE (17) and let Y_t^n to be the unique classical solution of the following ODE:

$$dY_t^n = b(Y_t^n) dt + (h_{\alpha\beta}(Y_t^n) - \frac{1}{2} DV_{\beta}(Y_t^n) \cdot V_{\alpha}(Y_t^n)) d\langle B^{\alpha}, B^{\beta} \rangle_t + V_{\alpha}(Y_t^n) d(B_t^n)^{\alpha}$$

with $X_0^n = Y_0^n = \xi$, where B_t^n is the dyadic piecewise linear approximation of B_t , then by using our main result in Section 2 and establishing related L_G^2 -estimates, we can prove that

$$\sup_{t \in [0,1]} \mathbb{E}^G[|X_t^n - Y_t^n|^2] \leq C \sqrt{1 + \xi^2} (\Delta t^n)^{\frac{1}{2}}.$$

In other words, Y_t^n converges to the solution X_t of the SDE (17) in the L_G^2 -sense. However, we know that for quasi-surely, Y_t^n converges uniformly to the solution Y_t of the RDE(22). Again by Lemma 4 and continuity, we conclude that for quasi-sure, X coincides with Y .

From the above discussion, if we forget about the RDE (22) and only consider the L_G^2 -limit of Y_t^n , it seems that there is nothing to do with rough paths at all as everything is well-defined in the classical sense. However, the fundamental point of

understanding the convergence of Y_t^n in the pathwise sense lies in the crucial fact that B_t can be regarded as geometric rough paths (i.e., the enhancement defined in Section 3) with approximating sequence in $G\Omega_p(\mathbb{R}^d)$ being the enhancement of the natural dyadic piecewise linear approximation B_t^n . This is exactly what the universal limit theorem tells us.

Remark 3. From the RDE point of view, it is possible to reduce the regularity assumptions on the coefficients. In particular, since the regularity of t and $\langle B^\alpha, B^\beta \rangle_t$ are both “better” than B_t , the regularity assumptions on the coefficients of dt and $d\langle B^\alpha, B^\beta \rangle_t$ can be weaker than the one imposed on the coefficient of dB_t . However, we are not going to present the results under such generality. Please refer to [9] for general existence and uniqueness results of RDEs.

6 SDEs on a Differentiable Manifold Driven by G-Brownian Motion

Our main result in Section 5 can be used to establish SDEs on a differentiable manifold driven by G -Brownian motion, which will be the main focus of this section. The development is based on the idea in the classical case, for which one may refer to [8], [12], [13]. This part is the foundation of developing G -Brownian motion on a Riemannian manifold in the next section.

In classical stochastic analysis, SDEs on a manifold is established under the Stratonovich type formulation, which can be regarded as a pathwise approach. The reason of using Stratonovich type formulation instead of the Itô type one is the following. First of all, the notion of SDE can be introduced by using test functions on the manifold from an intrinsic point of view, which is consistent with ordinary differential calculus and invariant under diffeomorphisms. Moreover, when we construct solutions extrinsically, we can prove that for almost surely, the solution of the extended SDE which starts from the manifold will always live on it. This reveals the intrinsic nature of ordinary differential equations.

In the setting of G -expectation, we will adopt the same idea for the development. However, there is a major difficulty here. The method of constructing solutions in the classical case from the extrinsic point of view depends heavily on the localization technique, which is not available in the setting of G -expectation, mainly due to the reason that concepts of information flows and stopping times are not well understood. To get around with this difficulty, we will use our main result in Section 5 to obtain a pathwise construction. The advantage of such approach is that we can still use localization arguments but don't need to care about measurability and integrability under G -expectation.

Now assume that M is a differentiable manifold. For technical reasons we further assume that M is compact (it is not necessary if we impose more restrictive regularity assumptions on the generating vector fields). Let $\{b, h_{\alpha\beta}, V_\alpha : \alpha, \beta = 1, 2, \dots, d\}$ be a family of C^3 -vector fields on M , and let B_t be the canonical d -dimensional G -

Brownian motion on the path space $(\Omega, L_G^2(\Omega), \mathbb{E}^G)$, where G is a function given by (2).

Consider the following symbolic Stratonovich type SDE over $[0, 1]$:

$$\begin{cases} dX_t &= b(X_t)dt + h_{\alpha\beta}(X_t)d\langle B^\alpha, B^\beta \rangle_t + V_\alpha(X_t) \circ dB_t^\alpha, \\ X_0 &= \xi \in M, \end{cases} \quad (24)$$

on M .

Definition 6. A solution X_t of the SDE (24) is an M -valued continuous stochastic process such that for any $f \in C^\infty(M)$, and any $\alpha, \beta = 1, 2, \dots, d$,

$$\{h_{\alpha\beta}f(X_t) : t \in [0, 1]\} \in M_G^1(0, 1), \{V_\alpha f(X_t) : t \in [0, 1]\} \in M_G^2(0, 1),$$

and the following equality holds on $[0, 1]$:

$$f(X_t) = f(\xi) + \int_0^t b f(X_s) ds + \int_0^t h_{\alpha\beta} f(X_s) d\langle B^\alpha, B^\beta \rangle_s + \int_0^t V_\alpha f(X_s) \circ dB_s^\alpha, \quad (25)$$

where the last term is defined as

$$\int_0^t V_\alpha f(X_s) \circ dB_s^\alpha := \int_0^t V_\alpha f(X_s) dB_s^\alpha + \frac{1}{2} \int_0^t V_\beta V_\alpha f(X_s) d\langle B^\alpha, B^\beta \rangle_s.$$

Remark 4. Definition 6 is intrinsic. It is easy to see that Definition 6 is consistent with the Euclidean case.

Now we are going to construct the solution of (24) from the extrinsic point of view.

According to the Whitney embedding theorem (see [5]), M can be embedded into some ambient Euclidean space \mathbb{R}^N as a submanifold such that the image $i(M)$ of M is closed in \mathbb{R}^N . We simply regard M as a subset of \mathbb{R}^N .

Let $F^1, \dots, F^N \in C^\infty(M)$ be the coordinate functions on M . The following result is easy to prove. It is similar to the classical case.

Proposition 6. X_t is a solution of (24) if and only if for any $i = 1, 2, \dots, N$, $\alpha, \beta = 1, 2, \dots, d$,

$$\{h_{\alpha\beta}F^i(X_t) : t \in [0, 1]\} \in M_G^1(0, 1), \{V_\alpha F^i(X_t) : t \in [0, 1]\} \in M_G^2(0, 1),$$

and for any $t \in [0, 1]$,

$$F^i(X_t) = F^i(\xi) + \int_0^t b F^i(X_s) ds + \int_0^t h_{\alpha\beta} F^i(X_s) d\langle B^\alpha, B^\beta \rangle_s + \int_0^t V_\alpha F^i(X_s) \circ dB_s^\alpha. \quad (26)$$

Proof. Necessity is obvious since $F^i \in C^\infty(M)$ for any $i = 1, 2, \dots, N$.

Now consider sufficiency. Let $f \in C^\infty(M)$, and choose a C^∞ -extension \tilde{f} of f with compact support in \mathbb{R}^N (it is possible since M is compact). Then for any $x \in M$,

$$f(x) = \tilde{f}(F^1(x), \dots, F^N(x)),$$

and thus

$$f(X_t) = \tilde{f}(F^1(X_t), \dots, F^N(X_t)), \forall t \in [0, 1].$$

Since M is compact and \tilde{f} is smooth with compact support, it follows from the G -Itô formula that for $t \in [0, 1]$,

$$\begin{aligned} \tilde{f}(F^1(X_t), \dots, F^N(X_t)) &= f(\xi) + \int_0^t \frac{\partial \tilde{f}}{\partial y^i} (bF^i(X_s) ds + h_{\alpha\beta} F^i(X_s) d\langle B^\alpha, B^\beta \rangle_s \\ &\quad + V_\alpha F^i(X_s) \circ dB_s^\alpha) \\ &= f(\xi) + \int_0^t (bf(X_s) ds + h_{\alpha\beta} f(X_s) d\langle B^\alpha, B^\beta \rangle_s \\ &\quad + V_\alpha f(X_s) \circ dB_s^\alpha), \end{aligned}$$

where we have used the simple fact that for any C^1 -vector field V on M ,

$$Vf = \sum_{i=1}^N \frac{\partial f}{\partial y^i} VF^i.$$

By Definition 6, we know that X_t is a solution of the SDE (24).

Now we are going to prove the existence and uniqueness of (24) by using the main result of Section 5, namely, a pathwise approach based on the associated RDE.

Let $\tilde{b}, \tilde{h}_{\alpha\beta}, \tilde{V}_\alpha$ be C_b^3 -extensions (not unique) of the vector fields $b, h_{\alpha\beta}, V_\alpha$. Consider the following Stratonovich type SDE in the ambient space \mathbb{R}^N :

$$dX_t = \tilde{b}(X_t)dt + \tilde{h}_{\alpha\beta}(X_t)d\langle B^\alpha, B^\beta \rangle_t + \tilde{V}_\alpha(X_t) \circ dB_t^\alpha \quad (27)$$

with $X_0 = x \in \mathbb{R}^N$, which is interpreted as the following Itô type SDE:

$$dX_t = \tilde{b}(X_t)dt + (\tilde{h}_{\alpha\beta}(X_t) + \frac{1}{2}D\tilde{V}_\alpha(X_t) \cdot \tilde{V}_\beta(X_t))d\langle B^\alpha, B^\beta \rangle_t + \tilde{V}_\alpha(X_t)dB_t^\alpha.$$

According to Section 5, we can alternatively interpret (27) as an RDE which is pathwisely defined. Both the SDE and the RDE has a unique solution, and according to Theorem 9 they coincide quasi-surely. Our aim is to show that for quasi-surely, the solution X_t of (27) never leaves M and it is the unique solution of (24).

The following result is important to prove the existence and uniqueness of the SDE (24) on the manifold M .

Proposition 7. *Let x_t be a path of bounded variation in \mathbb{R}^d . Let W_1, \dots, W_d be a family of C^1 -vector fields on M and $\tilde{W}_1, \dots, \tilde{W}_d$ be their C_b^1 -extensions to \mathbb{R}^N . Consider the following ODE in the ambient space \mathbb{R}^N over $[0, 1]$:*

$$dy_t = \tilde{W}_\alpha(y_t)dx_t^\alpha \quad (28)$$

with $y_0 = x \in M$. Then the solution $y_t \in M$ for all $t \in [0, 1]$. Moreover, y_t does not depend on extensions of the vector fields.

Proof. Let $F(x) := d(x, M)^2$ be the squared distance function to the submanifold M . It follows that F is smooth in an open neighborhood of M . By using the cut-off function we may assume that $F \in C_b^\infty(M)$. Now we are able to choose an open neighborhood U of M , such that for any $x \in U$, $F(x) = 0$ if and only if $x \in M$. Moreover, since \tilde{W}_α ($\alpha = 1, 2, \dots, d$) are tangent vector fields of M when restricted on M , U can be chosen such that for any $x \in U$ and $\alpha = 1, 2, \dots, d$,

$$|\tilde{W}_\alpha F(x)| \leq CF(x), \quad (29)$$

for some positive constant C depending on U . The function $F(x)$ was used in [12] to construct SDEs on M driven by classical Brownian motion.

Since x_t is a path of bounded variation and $y_0 = \xi \in M$, by the change of variables formula in ordinary calculus, we have

$$F(y_t) = \int_0^t \tilde{W}_\alpha F(y_s) dx_s^\alpha, \quad \forall t \in [0, 1].$$

Define $\tau := \inf\{t \in [0, 1] : y_t \notin U\}$. It follows from (29) that

$$F(y_t) \leq C \int_0^t F(y_s) d|x|_s, \quad \forall t \in [0, \tau],$$

where $|x|_t$ is the total variation of the path x_t .

By iteration and Fubini theorem, on $[0, \tau]$ we have

$$\begin{aligned} F(y_t) &\leq C^2 \int_0^t \left(\int_0^s F(y_u) d|x|_u \right) d|x|_s \\ &= C^2 \int_0^t (|x|_t - |x|_s) F(y_s) d|x|_s. \end{aligned}$$

By induction, it is easy to see that for any $k \geq 1$,

$$F(y_t) \leq C^k \int_0^t \frac{(|x|_t - |x|_s)^{k-1}}{(k-1)!} F(y_s) d|x|_s, \quad \forall t \in [0, \tau].$$

Since F is bounded, we obtain further that for any $k \geq 1$,

$$F(y_t) \leq \|F\|_\infty \frac{C^k (|x|_t - |x|_0)^k}{k!}, \quad \forall t \in [0, \tau].$$

By letting $k \rightarrow \infty$, it follows that $F(y_t) \equiv 0$ on $[0, \tau]$, which implies that $y_t \in M$ for any $t \in [0, \tau]$. Since y_t is continuous, the only possibility is that y_t never leaves M on $[0, 1]$.

If we rewrite the ODE (28) in its integral form:

$$y_t = \xi + \int_0^t \tilde{W}_\alpha(y_s) dx_s^\alpha, \quad t \in [0, 1], \quad (30)$$

we know from previous discussion that equation (30) depends only on the values of \tilde{W}_α on M , that is, of W_α ($\alpha = 1, 2, \dots, d$). In other words, if \hat{W}_α is another extension of W_α and \hat{y}_t is the solution of the corresponding ODE with the same initial condition, \hat{y}_t is also a solution of (28). By uniqueness, we have $y = \hat{y}$. Therefore, y_t does not depend on extensions of the vector fields.

With the help of Proposition 7, we can prove the following existence and uniqueness result.

Theorem 10. *Let $b, h_{\alpha\beta}, V_\alpha$ be C^3 -vector fields on M . Then the Stratonovich type SDE (24) has a solution X_t which is unique quasi-surely.*

Proof. Fix C_b^3 -extensions $\tilde{b}, \tilde{h}_{\alpha\beta}, \tilde{V}_\alpha$ of $b, h_{\alpha\beta}, V_\alpha$, and let X_t be the solution of the Stratonovich type SDE (27) in \mathbb{R}^N over $[0, 1]$. By Theorem 9, for quasi-surely X_t coincides with the solution of (27) when it is interpreted as an RDE. Since M is closed in \mathbb{R}^N , it follows from Proposition 7 and Theorem 6 (the universal limit theorem) that for quasi-surely, X_t never leaves M over $[0, 1]$. In this case, (27) is equivalent to (26), which implies from Proposition 6 that X_t is a solution of (24). On the other hand, if Y_t is another solution of (24), then it is a solution of (27) (interpreted as an SDE or an RDE). By the uniqueness of RDEs, we know that $X = Y$ quasi-surely.

Remark 5. It is possible to formulate uniqueness in the L_G^2 -sense when M is regarded as a closed submanifold of \mathbb{R}^N . However, we use the quasi-sure formulation because the notion itself is intrinsic although the proof is developed from the extrinsic point of view.

7 G-Brownian Motion on a Compact Riemannian Manifold and the Generating Nonlinear Heat Equation

In this section, we are going to introduce the notion of G -Brownian motion on a Riemannian manifold for a wide and interesting class of G -functions, based on Eells-Elworthy-Malliavin's horizontal lifting construction (see [8], [12], [13] for the construction of Brownian motion on a Riemannian manifold and related topics). Roughly speaking, we will "roll" an Euclidean G -Brownian motion up to a Riemannian manifold "without slipping" via a proper frame bundle (for the class of G -functions we are interested in, such bundle is the orthonormal frame bundle).

In the classical case, we know that the law of a d -dimensional Brownian motion B_t is invariant under orthogonal transformations on \mathbb{R}^d . This is a crucial point to obtain a linear parabolic PDE (in fact, the standard heat equation associated with the Bochner horizontal Laplacian $\Delta_{\mathcal{O}(M)}$) on the orthonormal frame bundle $\mathcal{O}(M)$ over a Riemannian manifold M governing the law of the horizontal lifting ξ_t of B_t to

$\mathcal{O}(M)$, which is invariant under orthogonal transformations along fibers. It is such an invariance that enables us to “project” the PDE onto the base manifold M and obtain the standard heat equation associated with the Laplace-Beltrami operator Δ_M on M . This heat equation governs the law of the development $X_t = \pi(\xi_t)$ of B_t to the Riemannian manifold M via the horizontal lifting ξ_t . As a stochastic process on M , although X_t depends on the initial orthonormal frame ξ at x as well as the initial position $x \in M$, the law of X_t depends only on the initial position x , and it is characterized by the Laplace-Beltrami operator Δ_M via the heat equation. Equivalently, it can be shown that the law of X_t is the unique solution of the martingale problem on M associated with Δ_M starting at x . X_t is called the Brownian motion on M starting at x in the sense of Eells-Elworthy-Malliavin.

It is quite natural to expect that the Brownian sample paths X_t on M will depend on the initial orthonormal frame ξ at x if we look back into the Euclidean case, in which we actually fix the standard orthonormal basis in advance and define Brownian motion in the corresponding coordinate system. If we use another orthonormal basis, we obtain a process (still a Brownian motion) which is an orthogonal transformation of the original Brownian motion. Therefore, it is the law, which is characterized by the Laplace operator on \mathbb{R}^d , rather than the sample paths that captures the intrinsic nature of the Brownian motion, and such nature can be developed in a Riemannian geometric setting.

It should be remarked that in a pathwise manner, we can lift B_t horizontally to the total frame bundle $\mathcal{F}(M)$ instead of $\mathcal{O}(M)$ by solving the same SDE generating by the horizontal vector fields but using a general frame instead of an orthonormal one as initial condition. Moreover, we can write down the generating heat equation on $\mathcal{F}(M)$ which takes the same form of the one on $\mathcal{O}(M)$. The key difference here is that although the horizontal lifting of B_t can be projected onto M , the heat equation on $\mathcal{F}(M)$ cannot. In other words, the heat equation is not invariant under nondegenerate linear transformations along fibers. This becomes uninteresting to us, as we are not able to obtain an intrinsic law of the development of B_t on M which is independent of initial frames. The fundamental reason of using the orthonormal frame bundle is that the Laplace operator on \mathbb{R}^d is invariant exactly under orthogonal transformations.

The case of G -Brownian motion can be understood in a similar manner. From the last section we are able to solve SDEs on a differentiable manifold (in particular, on $\mathcal{F}(M)$) driven by an Euclidean G -Brownian motion B_t . By projection we obtain the development X_t of B_t to M . As we’ve pointed out before, such development is of no interest unless we are able to prove that the law of X_t depends only on the initial position x rather than the initial frame. In fact, if the law of X_t depends on the initial frame, we might not be able to write down the generating PDE of X_t intrinsically on M although it is possible on $\mathcal{F}(M)$. Therefore, for a given G -function, it is crucial to identify a proper frame bundle over M with a specific structure group such that parallel transport preserves fibers and the generating PDE (associated with G) of the horizontal lifting ξ_t of B_t to such frame bundle is invariant under actions by the structure group along fibers. From this, the law of X_t will be independent of initial frames in the fibre over x (x is the starting point of X_t) and we might be able to obtain

the generating PDE of X_t , which is associated with G and intrinsically defined on M .

As we shall see, such idea depends on a crucial algebraic quantity associated with the G -function called the invariant group $I(G)$ of G , which will be defined later on. In this article, we are interested in the case when $I(G)$ is the orthogonal group. We will see that it contains a wide class of G -functions. In particular, one example is the generalization of the one-dimensional Barenblatt equation to higher dimensions.

The concept of the invariant group of G is motivated from the study of infinitesimal diffusive nature of SDEs driven by G -Brownian motion and their generating PDEs, which will be discussed below.

We first consider the Euclidean case.

From now on, we always assume that $G : S(d) \rightarrow \mathbb{R}$ is a given continuous, sub-linear and monotonic function. Equivalently, from Section 2 we know that G is represented by

$$G(A) = \frac{1}{2} \sup_{B \in \Sigma} \text{tr}(AB), \quad \forall A \in S(d), \quad (31)$$

where Σ is some bounded, closed and convex subset of $S_+(d)$. Let B_t be the standard d -dimensional G -Brownian motion on the path space.

Assume that V_1, \dots, V_d are C_b^3 -vector fields on \mathbb{R}^N . Consider the following N -dimensional Stratonovich type SDE over $[0, 1]$:

$$\begin{cases} dX_{t,x} = V_\alpha(X_{t,x}) \circ dB_t^\alpha, \\ X_{0,x} = x, \end{cases} \quad (32)$$

which is either interpreted as an RDE or the associated Itô type SDE

$$\begin{cases} dX_{t,x} = V_\alpha(X_{t,x}) dB_t^\alpha + \frac{1}{2} DV_\alpha(X_{t,x}) V_\beta(X_{t,x}) \langle B^\alpha, B^\beta \rangle_t, \\ X_{t,x} = x, \end{cases}$$

according to the main result of Section 5.

The following result characterizes the generator of the SDE (32) in terms of G . It describes the infinitesimal diffusive nature of (32). One might compare it with the case of linear diffusion processes.

Proposition 8. *For any $p \in \mathbb{R}^N$, $A \in S(N)$,*

$$\begin{aligned} & \lim_{\delta \rightarrow 0^+} \frac{1}{\delta} \mathbb{E}^G [\langle p, X_{\delta,x} - x \rangle + \frac{1}{2} \langle A(X_{\delta,x} - x), X_{\delta,x} - x \rangle] \\ & = G \left(\left(\frac{1}{2} \langle p, DV_\alpha(x) V_\beta(x) + DV_\beta(x) V_\alpha(x) \rangle + \langle AV_\alpha(x), V_\beta(x) \rangle \right)_{1 \leq \alpha, \beta \leq d} \right). \end{aligned} \quad (33)$$

Proof. From the distribution of B_t we know that

$$G(A) = \frac{1}{2} \mathbb{E}^G [\langle AB_1, B_1 \rangle] = \frac{1}{2t} \mathbb{E}^G [\langle AB_t, B_t \rangle], \quad \forall t > 0.$$

Therefore, the R.H.S. of (33) is equal to

$$I_\delta = \frac{1}{2\delta} \mathbb{E}^G[\langle p, DV_\alpha(x)V_\beta(x) \rangle + \langle AV_\alpha(x), V_\beta(x) \rangle B_\delta^\alpha B_\delta^\beta],$$

for any $\delta > 0$.

Since

$$X_{\delta,x} - x = \int_0^\delta V_\alpha(X_{s,x}) dB_s^\alpha + \frac{1}{2} \int_0^\delta DV_\alpha(X_{s,x}) V_\beta(X_{s,x}) d\langle B^\alpha, B^\beta \rangle_s,$$

by the properties of \mathbb{E}^G and the distribution of B_t , we have

$$\begin{aligned} & \left| \frac{1}{\delta} \mathbb{E}^G[\langle p, X_{\delta,x} - x \rangle + \frac{1}{2} \langle A(X_{\delta,x} - x), X_{\delta,x} - x \rangle] - I_\delta \right| \\ & \leq \left| \frac{1}{2\delta} \mathbb{E}^G \left[\int_0^\delta \langle p, DV_\alpha(X_{s,x}) \cdot V_\beta(X_{s,x}) \rangle d\langle B^\alpha, B^\beta \rangle_s \right. \right. \\ & \quad \left. \left. + \langle A \int_0^\delta V_\alpha(X_{s,x}) dB_s^\alpha, \int_0^\delta V_\beta(X_{s,x}) dB_s^\beta \rangle \right] - \frac{1}{2\delta} \mathbb{E}^G[\langle p, DV_\alpha(x) \cdot V_\beta(x) \rangle \langle B^\alpha, B^\beta \rangle_\delta \right. \\ & \quad \left. + \langle AV_\alpha(x) B_\delta^\alpha, V_\beta(x) B_\delta^\beta \rangle \right] + C\delta^{\frac{1}{2}} + C\delta \\ & \leq \frac{1}{2\delta} \left(C \int_0^\delta \sqrt{\mathbb{E}^G[|X_{s,x} - x|^2]} ds + C \int_0^\delta \mathbb{E}^G[|X_{s,x} - x|^2] ds + \right. \\ & \quad \left. C\delta^{\frac{1}{2}} \sqrt{\int_0^\delta \mathbb{E}^G[|X_{s,x} - x|^2] ds} + C\delta^{\frac{1}{2}} + C\delta \right), \end{aligned}$$

where we've also used the fact that G -Itô integrals and $B_\delta^\alpha B_\delta^\beta - \langle B^\alpha, B^\beta \rangle_\delta$ have zero mean uncertainty. Here C always denotes positive constants independent of δ .

Now the result follows easily from the fact that

$$\mathbb{E}^G[|X_{t,x} - x|^2] \leq Ct, \quad \forall t \in [0, 1].$$

The infinitesimal diffusive nature of (32) characterized by Proposition 8 enables us to establish the generating PDE of (32) in terms of viscosity solutions. The understanding of this PDE, especially its intrinsic nature, is essential for the development in a geometric setting.

Theorem 11. *Let $\varphi \in C_b^\infty(\mathbb{R}^N)$, and define*

$$u(t, x) = \mathbb{E}^G[\varphi(X_{t,x})], \quad (t, x) \in [0, 1] \times \mathbb{R}^N.$$

Then $u(t, x)$ is the unique viscosity solution of the following nonlinear parabolic PDE:

$$\begin{cases} \frac{\partial u}{\partial t} - G(\widehat{V_\alpha V_\beta} u)_{1 \leq \alpha, \beta \leq d} = 0, \\ u(0, x) = \varphi(x), \end{cases} \quad (34)$$

where $\widehat{V_\alpha V_\beta}$ denotes the symmetrization of the second order differential operator $V_\alpha V_\beta$, that is,

$$\widehat{V_\alpha V_\beta} = \frac{1}{2}(V_\alpha V_\beta + V_\beta V_\alpha).$$

Proof. The continuity of u in t and x can be shown in a standard way by using the Lipschitz continuity of φ (in fact, u is Lipschitz in x and $\frac{1}{2}$ -Hölder continuous in t). Here the proof is omitted.

Fix $(t_0, x_0) \in (0, 1) \times \mathbb{R}^N$. Let $v(t, x) \in C_b^{2,3}([0, 1] \times \mathbb{R}^N)$ be a test function such that

$$u(t_0, x_0) = v(t_0, x_0)$$

and

$$u(t, x) \leq v(t, x), \quad \forall (t, x) \in [0, 1] \times \mathbb{R}^N.$$

For $0 < \delta < t_0$, by the uniqueness of the SDE (32) and the fact that B_t and $\langle B^\alpha, B^\beta \rangle_t$ have independent and identically distributed increments, we know that

$$\begin{aligned} \mathbb{E}^G[\varphi(X_{t_0, x_0}) | \Omega_\delta] &= \mathbb{E}^G[\varphi(X_{\delta, x_0} + \int_\delta^{t_0} V_\alpha(X_{s, x_0}) dB_s^\alpha \\ &\quad + \frac{1}{2} \int_\delta^{t_0} DV_\alpha(X_{s, x_0}) \cdot V_\beta(X_{s, x_0}) d\langle B^\alpha, B^\beta \rangle_s) | \Omega_\delta] \\ &= \mathbb{E}^G[\varphi(X_{t_0 - \delta, y})]_{y=X_{\delta, x_0}}. \end{aligned}$$

Therefore,

$$\begin{aligned} v(t_0, x_0) &= \mathbb{E}^G[\varphi(X_{t_0, x_0})] \\ &= \mathbb{E}^G[\mathbb{E}^G[\varphi(X_{t_0, x_0}) | \Omega_\delta]] \\ &= \mathbb{E}^G[u(t_0 - \delta, X_{\delta, x_0})] \\ &\leq \mathbb{E}^G[v(t_0 - \delta, X_{\delta, x_0})]. \end{aligned}$$

It follows that

$$\begin{aligned} 0 &\leq \mathbb{E}^G[v(t_0 - \delta, X_{\delta, x_0}) - v(t_0, x_0)] \\ &= \mathbb{E}^G[v(t_0 - \delta, X_{\delta, x_0}) - v(t_0, X_{\delta, x_0}) + v(t_0, X_{\delta, x_0}) - v(t_0, x_0)] \\ &= \mathbb{E}^G[-\delta \int_0^1 \frac{\partial v}{\partial t}(t_0 - (1 - \alpha)\delta, X_{\delta, x_0}) d\alpha + \langle \nabla v(t_0, x_0), X_{\delta, x_0} - x_0 \rangle \\ &\quad + \int_0^1 \int_0^1 \langle \nabla^2 v(t_0, x_0 + \alpha\beta(X_{\delta, x_0} - x_0))(X_{\delta, x_0} - x_0), X_{\delta, x_0} - x_0 \rangle \alpha d\alpha d\beta] \\ &\leq -\delta \frac{\partial v}{\partial t}(t_0, x_0) + \mathbb{E}^G[\langle \nabla v(t_0, x_0), X_{\delta, x_0} - x_0 \rangle \\ &\quad + \frac{1}{2} \langle \nabla^2 v(t_0, x_0)(X_{\delta, x_0} - x_0), X_{\delta, x_0} - x_0 \rangle] + \mathbb{E}^G[|J_\delta|] + \mathbb{E}^G[|J_\delta|], \end{aligned}$$

where

$$\begin{aligned} I_\delta &= -\delta \int_0^1 \left(\frac{\partial v}{\partial t}(t_0 - (1-\alpha)\delta, X_{\delta, x_0}) - \frac{\partial v}{\partial t}(t_0, x_0) \right) d\alpha, \\ J_\delta &= \int_0^1 \int_0^1 \langle (\nabla^2 v(t_0, x_0 + \alpha\beta(X_{\delta, x_0} - x_0)) \\ &\quad - \nabla^2 v(t_0, x_0))(X_{\delta, x_0} - x_0), X_{\delta, x_0} - x_0 \rangle \alpha d\alpha d\beta. \end{aligned}$$

By a standard argument one can easily show that

$$\mathbb{E}^G[|I_\delta|] + \mathbb{E}^G[|J_\delta|] \leq C\delta^{\frac{3}{2}},$$

where C is a positive constant independent of δ . On the other hand, the R.H.S. of (33) applying to

$$p = \nabla v(t_0, x_0), \quad A = \nabla^2 v(t_0, x_0),$$

is exactly the same as $G((\widehat{V_\alpha V_\beta} v(t_0, x_0))_{1 \leq \alpha, \beta \leq d})$. Therefore, by Proposition 8, we arrive at

$$\frac{\partial v}{\partial t}(t_0, x_0) - G((\widehat{V_\alpha V_\beta} v(t_0, x_0))_{1 \leq \alpha, \beta \leq d}) \leq 0.$$

Consequently, $u(t, x)$ is a viscosity subsolution of (34).

Similarly, one can show that $u(t, x)$ is a viscosity supersolution of (34). Therefore, $u(t, x)$ is a viscosity solution of (34).

The reason of uniqueness is the following. Define a function $F : \mathbb{R}^N \times \mathbb{R}^N \times S(N) \rightarrow \mathbb{R}$ by the R.H.S. of (33), that is,

$$F(x, p, A) = G\left(\frac{1}{2}\langle p, DV_\alpha(x) \cdot V_\beta(x) + DV_\beta(x) \cdot V_\alpha(x) \rangle + \langle AV_\alpha(x), V_\beta(x) \rangle\right)_{1 \leq \alpha, \beta \leq d},$$

for $(x, p, A) \in \mathbb{R}^N \times \mathbb{R}^N \times S(N)$. It is easy to prove that F is sublinear in (p, A) and monotonically increasing in $S(N)$, due to the same properties held by G . Moreover, F satisfies the continuity condition (Assumption (G) in Appendix C of [27]) for the uniqueness of the associated nonlinear PDE, due to the regularity of the given vector fields V_α . In other words, all properties of G to ensure uniqueness are preserved in F , and the space dependence of F coming out are uniformly controlled. Therefore, according to the uniqueness results (see [4], [27]), the parabolic PDE has a unique viscosity solution, which is given by $u(t, x)$.

Example 1. An example which motivates the study of G -Brownian motion on a Riemannian manifold is the following.

Let $Q \in GL(d, \mathbb{R})$, where $GL(d, \mathbb{R})$ is the group of $d \times d$ real invertible matrices. Define $B_t^Q = QB_t$, and for $\varphi \in C_b^\infty(\mathbb{R}^d)$, define

$$u(t, x) = \mathbb{E}^G[\varphi(x + B_t^Q)], \quad (t, x) \in [0, 1] \times \mathbb{R}^d.$$

Then $u(t, x)$ is the unique viscosity solution of the PDE:

$$\begin{cases} \frac{\partial u}{\partial t} - G(Q^T \cdot \nabla^2 u \cdot Q) = 0, \\ u(0, x) = \varphi(x). \end{cases}$$

In fact, it follows directly from Theorem 11 if we regard $x + B_t^Q$ as the solution of the SDE over $[0, 1]$:

$$\begin{cases} dX_{t,x} = Q_\alpha \circ dB_t^\alpha, \\ X_{0,x} = x, \end{cases} \quad (35)$$

where $Q = (Q_1, \dots, Q_d)$, and each Q_α is a constant vector field on \mathbb{R}^d (so the SDE (35) coincides exactly with the Itô type one).

The result of Theorem 34 is similar to the discussion of nonlinear Feynman-Kac formula in [27], in which the solution of a forward-backward SDE is used to represent the viscosity solution of an associated nonlinear backward parabolic PDE. In our case, the intrinsic nature of (34) is fundamental and should be emphasized below in order to develop G -Brownian motion on a Riemannian manifold.

It is not hard to see that the nonlinear second order differential operator

$$G((\widehat{V_\alpha V_\beta \cdot})_{1 \leq \alpha, \beta \leq d})$$

is intrinsically defined on \mathbb{R}^N , since V_1, \dots, V_d are vector fields independent of coordinates. Moreover, in local coordinates it preserves the same properties of the G -function which is defined under the standard coordinate system of \mathbb{R}^d . In particular, it shares the same ellipticity as G . Therefore, when the vector fields V_α are regular enough, from our results in Section 6, we are able to establish the generating PDE of a nonlinear diffusion process on a differentiable manifold. As in the last section, for technical simplicity we restrict ourselves to compact manifolds.

Assume that M is a compact manifold, and V_1, \dots, V_d are C^3 -vector fields on M . According to Section 6, the Stratonovich type SDE over $[0, 1]$

$$\begin{cases} dX_{t,x} = V_\alpha(X_{t,x}) \circ dB_t^\alpha, \\ X_{0,x} = x \in M, \end{cases} \quad (36)$$

has a unique solution. The following result is immediate from Theorem 11.

Theorem 12. *Let $\varphi \in C^\infty(M)$, and define*

$$u(t, x) = \mathbb{E}^G[\varphi(X_{t,x})], \quad (t, x) \in [0, 1] \times M,$$

then $u(t, x)$ is the unique viscosity solution of the following nonlinear parabolic PDE on M :

$$\begin{cases} \frac{\partial u}{\partial t} - G((\widehat{V_\alpha V_\beta u})_{1 \leq \alpha, \beta \leq d}) = 0, \\ u(0, x) = \varphi(x), \end{cases} \quad (37)$$

where $\widehat{V_\alpha V_\beta}$ is the symmetrization of $V_\alpha V_\beta$, defined in the same way as in Theorem 11. Here the notion of viscosity solutions for the PDE (37) can be defined in the same way as in the Euclidean case by using test functions (see [1]).

Proof. The result follows easily from an extrinsic point of view.

In fact, assume that M is embedded into an ambient Euclidean space \mathbb{R}^N as a closed submanifold, and take a C^3 -extension \widetilde{V}_α of V_α with compact support. Consider the following Stratonovich type SDE over $[0, 1]$:

$$\begin{cases} dX_{t,x} = \widetilde{V}_\alpha(X_{t,x}) \circ dB_t^\alpha, \\ X_{0,x} = x \in \mathbb{R}^N. \end{cases}$$

Let $\widetilde{\varphi}$ be a C^∞ -extension of φ with compact support, and define

$$\widetilde{u}(t,x) = \mathbb{E}^G[\widetilde{\varphi}(X_{t,x})], \quad (t,x) \in [0, 1] \times \mathbb{R}^N.$$

It follows from Theorem 11 that $\widetilde{u}(t,x)$ is the unique viscosity solution of the non-linear parabolic PDE generated by the vector fields \widetilde{V}_α .

According to Section 6, if $x \in M$, $X_{t,x}$ will never leave M quasi-surely. Therefore, when restricted on M , $\widetilde{u} = u$. In particular, we know that u is continuous. To see that u is a viscosity subsolution of (37), let $(t_0, x_0) \in (0, 1) \times M$, and $v(t,x) \in C^{2,3}([0, 1] \times M)$ be a test function such that

$$v(t_0, x_0) = u(t_0, x_0)$$

and

$$u(t,x) \leq v(t,x), \quad \forall (t,x) \in [0, 1] \times M.$$

Take an $C_b^{2,3}$ -extension \widetilde{v} of v such that

$$\widetilde{u}(t,x) \leq \widetilde{v}(t,x), \quad \forall (t,x) \in [0, 1] \times \mathbb{R}^N.$$

It follows from previous discussion that

$$\frac{\partial \widetilde{v}}{\partial t}(t_0, x_0) - G((\widehat{V_\alpha V_\beta} \widetilde{v}(t_0, x_0))_{1 \leq \alpha, \beta \leq d}) \leq 0.$$

Since

$$\widetilde{V}_\alpha|_M = V_\alpha, \quad \widetilde{v}|_M = v,$$

from the intrinsic nature of the generating PDE, we know that

$$\frac{\partial \widetilde{v}}{\partial t}(t_0, x_0) = \frac{\partial v}{\partial t}(t_0, x_0)$$

and

$$G((\widehat{V_\alpha V_\beta} \widetilde{v}(t_0, x_0))_{1 \leq \alpha, \beta \leq d}) = G((\widehat{V_\alpha V_\beta} v(t_0, x_0))_{1 \leq \alpha, \beta \leq d}).$$

It follows that

$$\frac{\partial v}{\partial t}(t_0, x_0) - G((\widehat{V_\alpha V_\beta} v)(t_0, x_0))_{1 \leq \alpha, \beta \leq d} \leq 0.$$

Therefore, $u(t, x)$ is a viscosity subsolution of (37). Similarly we can show that it is a viscosity supersolution as well, and thus a viscosity solution.

The uniqueness of (37) follows from the same reason as in the proof of Theorem 11 once we notice that the second order differential operator $G((\widehat{V_\alpha V_\beta} \cdot))_{1 \leq \alpha, \beta \leq d}$ on M shares exactly the same properties as G (in particular, the same ellipticity), which can be seen either from an extrinsic way or via local computation. Another way to see the uniqueness is to use the results in [1] as long as we assign a complete Riemannian metric on M , which is always possible according to [20]. In this case

$$G((\widehat{V_\alpha V_\beta} u))_{1 \leq \alpha, \beta \leq d} = G\left(\left(\frac{1}{2} \langle \nabla u, \nabla_{V_\alpha} V_\beta + \nabla_{V_\beta} V_\alpha \rangle + \text{Hess}u(V_\alpha, V_\beta)\right)_{1 \leq \alpha, \beta \leq d}\right),$$

where ∇ is the Levi-Civita connection corresponding to the Riemannian metric. The uniqueness of (37) follows from Theorem 5.1 in [1] directly, as the assumptions in the theorem are verified by the properties of G . Note that we don't need the Ricci curvature condition in [1] due to the compactness of M and uniform continuity of $G((\widehat{V_\alpha V_\beta} \cdot))_{1 \leq \alpha, \beta \leq d}$.

Remark 6. The study of the SDE (36) as a nonlinear diffusion process on M does not require a Riemannian metric or a connection on M . The fundamental reason is that (36) is defined in the pathwise sense as an RDE generated by the vector fields V_α on M . Such an RDE only depends on the differential structure of M . The infinitesimal diffusive nature of (36) can be studied by local computation.

Now we turn to the study of G -Brownian motion on a Riemannian manifold. The Riemannian structure (the Levi-Civita connection) is used to “roll” the Euclidean G -Brownian motion up to the manifold “without slipping” by solving an SDE generated by the fundamental horizontal vector fields on a proper frame bundle (known as horizontal lifting). This is the fundamental idea of Eells-Elworthy-Malliavin on the construction of Brownian motion on a Riemannian manifold.

As is pointed out at the beginning of this section, the essential point of such development is the invariance of the generating PDE on the frame bundle under actions by the structure group along fibers. The key of capturing such invariance is Theorem 34 and Example 1, which leads to the following important concept.

Definition 7. The invariant group $I(G)$ of G is defined by

$$I(G) = \{Q \in GL(d, \mathbb{R}) : \forall A \in S(d), G(Q^T A Q) = G(A)\}.$$

It is easy to check the $I(G)$ is a group, and hence a subgroup of $GL(d, \mathbb{R})$.

By using the representation (31) of G , we have the following equivalent characterization of the invariant group $I(G)$.

Proposition 9. Let G be represented by

$$G(A) = \frac{1}{2} \sup_{B \in \Sigma} \text{tr}(AB), \quad \forall A \in S(d),$$

where Σ is some bounded, closed and convex subset of $S_+(d)$. Then Σ is uniquely determined by G and the invariant group $I(G)$ of G is given by

$$I(G) = \{Q \in GL(d, \mathbb{R}) : Q\Sigma Q^T = \Sigma\}. \quad (38)$$

Proof. It suffices to show the uniqueness of Σ , and (38) will follow immediately from the commutativity of the trace operator and the uniqueness of Σ . Note that for any $Q \in GL(d, \mathbb{R})$, $Q\Sigma Q^T$ is also a bounded, closed and convex subset of $S_+(d)$.

Introduce a symmetric bilinear form $\langle \cdot, \cdot \rangle_{\text{tr}}$ on the finite dimensional vector space $S(d)$ by

$$\langle A_1, A_2 \rangle_{\text{tr}} = \text{tr}(A_1 A_2), \quad A_1, A_2 \in S(d).$$

It is easy to check that $\langle \cdot, \cdot \rangle_{\text{tr}}$ is indeed an inner product, thus $(S(d), \langle \cdot, \cdot \rangle_{\text{tr}})$ is a finite dimensional Hilbert space. The form $\|\cdot\|_{\text{tr}}$ induced by $\langle \cdot, \cdot \rangle_{\text{tr}}$ is equivalent to any other matrix norm on $S(d)$ since $S(d)$ is finite dimensional.

Let Σ_1, Σ_2 be two bounded, closed and convex subsets of $S_+(d)$, such that

$$\sup_{B \in \Sigma_1} \text{tr}(AB) = \sup_{B \in \Sigma_2} \text{tr}(AB), \quad \forall A \in S(d).$$

If $\Sigma_1 \neq \Sigma_2$, without loss of generality assume that $B_0 \in \Sigma_2 \setminus \Sigma_1$. According to the Mazur separation theorem in functional analysis (see [29]), there exists a bounded linear functional $f \in S(d)^*$ and some $\alpha \in \mathbb{R}$, such that

$$f(B) < \alpha < f(B_0), \quad \forall B \in \Sigma_1.$$

By the Riesz representation theorem, there exists a unique $A^* \in S(d)$, such that

$$f(B) = \langle A^*, B \rangle_{\text{tr}} = \text{tr}(A^* B), \quad \forall B \in S(d).$$

It follows that

$$\sup_{B \in \Sigma_1} \text{tr}(A^* B) \leq \alpha < \text{tr}(A^* B_0) \leq \sup_{B \in \Sigma_2} \text{tr}(A^* B),$$

which is a contradiction. Therefore, $\Sigma_1 = \Sigma_2$.

We list some examples for the invariant groups $I(G)$ of different G -functions.

Example 2. If $\Sigma = \{0\}$, then it is obvious that $I(G) = GL(d, \mathbb{R})$, which is a noncompact group.

Example 3. It is possible that $I(G)$ is a finite group.

Consider Σ is the set of diagonal matrices

$$\Lambda = \text{diag}(\lambda_1, \dots, \lambda_d)$$

such that each $\lambda_\alpha \in [0, 1]$, then Σ is a bounded, closed and convex subset of $S_+(d)$. We claim that

$$I(G) = \{(\pm e_{\sigma(1)}, \dots, \pm e_{\sigma(d)}) : \sigma \text{ is a permutation of order } d\}, \quad (39)$$

where $\{e_1, \dots, e_d\}$ is the standard orthonormal basis of \mathbb{R}^d , each e_i being regarded as a column vector.

In fact, if $Q \in GL(d, \mathbb{R})$ has the form (39), by direct computation one can show easily that

$$Q\Sigma Q^T = \Sigma. \quad (40)$$

Conversely, if Q satisfies (40), by choosing

$$\Lambda = \text{diag}(1, 0, \dots, 0),$$

we know that

$$(Q\Lambda Q^T)_{\beta}^{\alpha} = Q_1^{\alpha} Q_1^{\beta}.$$

Therefore, if $Q\Lambda Q^T \in \Sigma$, the first column of Q must contain exactly one nonzero element q_1 such that $q_1^2 \leq 1$. Similarly for other columns of Q . Moreover, the corresponding nonzero elements in any two different columns of Q must be in different rows, otherwise Q will be degenerate. Consequently, Q has the form

$$Q = (q_1 e_{\sigma(1)}, \dots, q_d e_{\sigma(d)})$$

with $q_i^2 \leq 1$ ($i = 1, 2, \dots, d$). On the other hand, for the identity matrix I_d , there exists $\Lambda \in \Sigma$, such that

$$Q\Lambda Q^T = I_d.$$

By taking determinants on both sides, we have

$$q_1^2 \cdots q_d^2 \det(\Lambda) = 1,$$

which implies that $q_{\alpha} = \pm 1$ ($\alpha = 1, 2, \dots, d$). Therefore, Q has the form of (39).

Note that in this case $I(G)$ is a finite subgroup of the orthogonal group $O(d)$ with order $2^d d!$. Moreover, G is given by

$$G(A) = \frac{1}{2} \sum_{\alpha=1}^d (A_{\alpha}^{\alpha})^+, \quad \forall A \in S(d).$$

Example 4. Now we give some examples of G such that $I(G) = O(d)$. Such case will be our main interest in this article.

(1) $\Sigma = \{I_d\}$.

Obviously (40) is equivalent to $Q \in O(d)$.

This corresponds to the case of classical Brownian motion, in which

$$G(A) = \frac{1}{2} \text{tr}(A)$$

and the generator is $\frac{1}{2}\Delta$.

(2) Σ is given by the segment joining λI_d and μI_d , where $0 \leq \lambda < \mu$.

If $Q \in GL(d, \mathbb{R})$ such that (40) holds, then

$$\mu QQ^T = tI_d,$$

for some $t \in [\lambda, \mu]$. On the other hand, there exists some $t' \in [\lambda, \mu]$ such that

$$t'QQ^T = \mu I_d.$$

The only possibility is that $QQ^T = I_d$, which means $Q \in O(d)$. The converse is trivial.

In this case, G is given by

$$G(A) = \frac{1}{2}(\mu(\text{tr}A)^+ - \lambda(\text{tr}A)^-).$$

The corresponding G -heat equation can be regarded as the generalization of the one-dimensional Barenblatt equation to higher dimensions.

(3) Σ is given by the subset of matrices $B \in S_+(d)$ such that the eigenvalues of B lie in the bounded interval $[\lambda, \mu]$, where $0 \leq \lambda < \mu$. Equivalently,

$$\Sigma = \{B \in S_+(d) : \lambda \leq x^T Bx \leq \mu, \forall x \in \mathbb{R}^d \text{ with } |x| = 1\}.$$

It follows that Σ is a bounded, closed and convex subset of $S_+(d)$.

Since Σ is characterized by eigenvalues, and the eigenvalues of a symmetric matrix is preserved under change of orthonormal basis, it follows that for any $Q \in O(d)$, (40) holds. Conversely, let $Q \in GL(d, \mathbb{R})$ with (40). Then there exists $B_1, B_2 \in \Sigma$, such that

$$\mu QQ^T = B_1, QB_2Q^T = \mu I_d.$$

It follows that all eigenvalues of QQ^T lie in $[\frac{\lambda}{\mu}, 1]$, and

$$\det(QQ^T)\det(B_2) = \mu^d.$$

Therefore, the only possibility is that all eigenvalues of QQ^T are equal to 1, which implies that Q is an orthogonal matrix.

In this case G can be expressed by

$$\begin{aligned} G(A) &= \frac{1}{2} \sup_{B \in \Sigma} \text{tr}(AB) \\ &= \frac{1}{2} \sup_{P \in O(d)} \sup_{\lambda \leq c_1, \dots, c_d \leq \mu} \text{tr}(AP^T \text{diag}(c_1, \dots, c_d)P) \\ &= \frac{1}{2} \sup_{P \in O(d)} \sup_{\lambda \leq c_1, \dots, c_d \leq \mu} \text{tr}(PAP^T \text{diag}(c_1, \dots, c_d)) \\ &= \frac{1}{2} \sup_{P \in O(d)} \sup_{\lambda \leq c_1, \dots, c_d \leq \mu} \sum_{\alpha=1}^d c_\alpha (PAP^T)_\alpha^\alpha \\ &= \frac{1}{2} \sup_{P \in O(d)} \sum_{\alpha=1}^d (\mu((PAP^T)_\alpha^\alpha)^+ - \lambda((PAP^T)_\alpha^\alpha)^-). \end{aligned}$$

Similar to Example 4, for those Σ 's characterized by eigenvalues, we can construct a large class of G such that $I(G) = O(d)$.

Remark 7. If Σ has at least one nondegenerate element, that is, there exists some positive definite matrix $B_0 \in \Sigma$, then $I(G)$ is a compact group. In fact, if we introduce a matrix norm $\|\cdot\|_{B_0}$ on the space $\text{Mat}(d, \mathbb{R})$ of real $d \times d$ matrices by

$$\|A\|_{B_0} = \sqrt{\text{tr}(AB_0A^T)}, \quad A \in \text{Mat}(d, \mathbb{R}),$$

it follows that

$$\sup_{Q \in I(G)} \|Q\|_{B_0} = \sup_{Q \in I(G)} \sqrt{\text{tr}(QB_0Q^T)} \leq \sup_{B \in \Sigma} \sqrt{\text{tr}(B)} < \infty,$$

since Σ is bounded. It is obvious that $I(G)$ is closed. Therefore, it is compact.

Now assume that (M, g) is a d -dimensional compact Riemannian manifold. If we allow explosion of a nonlinear diffusion process at some finite time, then the arguments below will carry through on a noncompact Riemannian manifold as long as the time scope is restricted from 0 up to the explosion. Here we only consider the compact case, in which explosion is not possible.

We first recall some basics about frame bundles, which is the central concept in the horizontal lifting construction. For a systematic introduction please refer to [2], [16].

Let $\mathcal{F}(M)$ be the total frame bundle over M defined by

$$\mathcal{F}(M) = \cup_{x \in M} \mathcal{F}_x(M),$$

where the fibre $\mathcal{F}_x(M)$ is the set of all frames (bases of the tangent space $T_x(M)$) at x . A frame $\xi = (\xi_1, \dots, \xi_d) \in \mathcal{F}_x(M)$ can be equivalently regarded as a linear isomorphism from \mathbb{R}^d to T_xM (also denoted by ξ) if we let

$$\xi(e_\alpha) = \xi_\alpha, \quad \alpha = 1, 2, \dots, d,$$

and extend linearly to \mathbb{R}^d , where we always fix $\{e_1, \dots, e_d\}$ to be the standard orthonormal basis of \mathbb{R}^d . $\mathcal{F}(M)$ is a principal bundle with structure group $GL(d, \mathbb{R})$ acting along fibers from the right.

Fix a frame $\xi \in \mathcal{F}_x(M)$. A vector $X \in T_\xi \mathcal{F}(M)$ is called vertical if it is tangent to the fibre $\mathcal{F}_x(M)$. The space of vertical vectors at ξ is called the vertical subspace, and it is denoted by $V_\xi \mathcal{F}(M)$. $V_\xi \mathcal{F}(M)$ is a d^2 -dimensional vector space, which is independent of the Riemannian structure.

A smooth curve $\xi_t = (\xi_{1,t}, \dots, \xi_{d,t}) \in \mathcal{F}(M)$ is called horizontal if $\xi_{\alpha,t}$ is a parallel vector field along the projection curve $x_t = \pi(\xi_t)$ for each $\alpha = 1, 2, \dots, d$. Given a smooth curve $x_t \in M$ and a frame $\xi_0 = (\xi_1, \dots, \xi_d) \in \mathcal{F}_{x_0}(M)$, by solving a first order linear ODE, we can determine a unique parallel vector field $\xi_{\alpha,t}$ along x_t with $\xi_{\alpha,0} = \xi_\alpha$ for each $\alpha = 1, 2, \dots, d$. The smooth curve

$$\xi_t = (\xi_{1,t}, \dots, \xi_{d,t}) \in \mathcal{F}(M)$$

is then the unique horizontal curve with $x_t = \pi(\xi_t)$ and initial position ξ_0 . ξ_t is called the horizontal lifting of x_t from ξ_0 . A vector $X \in T_\xi \mathcal{F}(M)$ is called horizontal if it is tangent to a horizontal curve through ξ . The space of horizontal vectors at ξ is called the horizontal subspace, and it is denoted by $H_\xi \mathcal{F}(M)$. It is a d -dimensional vector space characterized by the Levi-Civita connection ∇ .

As ξ varies, $V_\xi \mathcal{F}(M)$ (respectively, $H_\xi \mathcal{F}(M)$) determines a vertical (respectively, horizontal) subspace field on M . The following result reveals the fundamental structure of $\mathcal{F}(M)$.

Theorem 13. *The horizontal subspace field $H\mathcal{F}(M)$, which is determined by ∇ , has the following properties.*

(1) *For each $\xi \in \mathcal{F}_x(M)$, the tangent space $T_\xi \mathcal{F}(M)$ has the decomposition*

$$T_\xi \mathcal{F}(M) = H_\xi \mathcal{F}(M) \oplus V_\xi \mathcal{F}(M).$$

Moreover, $H_\xi \mathcal{F}(M)$ is isomorphic to $T_x M$ under the canonical projection $\pi : \mathcal{F}(M) \rightarrow M$.

(2) *$H\mathcal{F}(M)$ is invariant under actions by the structure group $GL(d, \mathbb{R})$. More precisely, for any $\xi \in \mathcal{F}(M)$, $Q \in GL(d, \mathbb{R})$,*

$$Q_*(H_\xi \mathcal{F}(M)) = H_{\xi Q} \mathcal{F}(M).$$

It should be pointed out that given any horizontal subspace field $H\mathcal{F}(M)$ satisfying the two properties in Theorem 13, there exists an affine connection ∇^H such that $H\mathcal{F}(M)$ is the horizontal subspace field determined by ∇^H .

On $\mathcal{F}(M)$ there is a canonical way to define a frame field globally, which is not always possible on a general Riemannian manifold. This makes $\mathcal{F}(M)$ simpler than the base space M in some sense. Fix $w \in \mathbb{R}^d$. For any $\xi \in \mathcal{F}_x(M)$ regarded as a linear isomorphism $\xi : \mathbb{R}^d \rightarrow T_x M$, $\xi(w)$ is a tangent vector in $T_x M$. By Theorem 13 (1), $\xi(w)$ corresponds to a unique vector $H_w(\xi) \in H_\xi \mathcal{F}(M)$. It follows that H_w is a globally defined horizontal vector field on $\mathcal{F}(M)$. If we take $w = e_\alpha$ ($\alpha = 1, 2, \dots, d$), then we obtain a family of horizontal vector fields $\{H_{e_1}, \dots, H_{e_d}\}$ as a basis of the horizontal subspace $H_\xi \mathcal{F}(M)$ at each frame $\xi \in \mathcal{F}(M)$. $\{H_{e_1}, \dots, H_{e_d}\}$ are called the fundamental horizontal fields of $\mathcal{F}(M)$, simply denoted by $\{H_1, \dots, H_d\}$.

Now we introduce the concept of development and anti-development (see [12]), which is crucial in the construction of G -Brownian motion on M . Assume that $x_t \in M$ is a smooth curve and ξ_t is the horizontal lifting of x_t from ξ_0 . Then we can determine a smooth curve

$$w_t = \int_0^t \xi_s^{-1} \dot{x}_s ds \in \mathbb{R}^d$$

starting from 0 (w_t is regarded as a column vector in \mathbb{R}^d). w_t is called the anti-development of x_t in \mathbb{R}^d with respect to ξ_0 . If ξ_t and η_t are two horizontal liftings of x_t with $\xi_0 = \eta_0 Q$ for some $Q \in GL(d, \mathbb{R})$, then the two corresponding anti-developments are related by

$$w_t^\eta = Q w_t^\xi.$$

The fundamental relation between the anti-development w_t of x_t and the horizontal lifting ξ_t is the following ODE on $\mathcal{F}(M)$:

$$d\xi_t = H_\alpha(\xi_t)dw_t^\alpha. \quad (41)$$

Conversely, given a smooth curve $w_t \in \mathbb{R}^d$ starting from 0, by solving the ODE (41) on $\mathcal{F}(M)$ with initial frame ξ_0 , we obtain a horizontal curve $\xi_t \in \mathcal{F}(M)$. The projection $x_t = \pi(\xi_t)$ is called the development of w_t in M with respect to ξ_0 . If we use another initial frame $\eta_0 = \xi_0 Q^{-1}$ and the driven process $v_t = Qw_t \in \mathbb{R}^d$, by solving (41) from η_0 and projection onto M we obtain the same curve x_t . In this way, we obtain a one-to-one correspondence of the Euclidean curve w_t and the manifold curve x_t via the horizontal curve ξ_t in $\mathcal{F}(M)$, which depends on the initial frame ξ_0 . The procedure of getting x_t from w_t is usually known as “rolling without slipping”.

A crucial point should be emphasized here is that such procedure is carried out by solving the ODE (41) in the pathwise sense, which fits well in the context of rough paths if the Euclidean curve w_t is interpreted as a rough path. In this case, (41) should be interpreted as an RDE. This is an important reason why we need to develop the notion of Stratonovich type SDEs on a differentiable manifold.

For a general Euclidean G -Brownian motion B_t , from Section 6 we are able to solve (41) pathwisely if the driven curve dw_t is replaced by dB_t in the Stratonovich sense (or in the RDE sense). By projecting the solution $\xi_t \in \mathcal{F}(M)$ to the manifold M , we obtain a process $X_t \in M$ pathwisely which depends on the initial position x_0 and the initial frame $\xi_0 \in \mathcal{F}_{x_0}(M)$. A disadvantage of using the total frame bundle $\mathcal{F}(M)$ is that in this way it is not possible to write down the generating PDE governing the law of X_t intrinsically on M , which does not depend on the initial frame ξ_0 . Note that the generating PDE of ξ_t is well-defined on $\mathcal{F}(M)$ according to Theorem 12, which takes the form

$$\frac{\partial u}{\partial t} - G((\widehat{H_\alpha H_\beta u})_{1 \leq \alpha, \beta \leq d}) = 0. \quad (42)$$

The main reason for such disadvantage is that the PDE (42) is not invariant under actions by $GL(d, \mathbb{R})$ along fibers, since the G -function does not have such kind of invariance.

To fix this issue, a possible way is to use the invariant group $I(G)$ of G as the structure group, so that the generating PDE will be invariant under actions by $I(G)$ along fibers due to the form (42) it takes. Therefore, we need to use a proper frame bundle (a submanifold of $\mathcal{F}(M)$ which is a principal bundle over M with structure group $I(G)$ and fibers being a suitable class of frames) instead of $\mathcal{F}(M)$. The fibers of such frame bundle should be preserved by parallel transport so the fundamental horizontal fields can be restricted on it and we are able to solve the RDE

$$d\xi_t = H_\alpha(\xi_t) \circ dB_t^\alpha$$

on the frame bundle. It will turn out that we are able to establish the generating PDE of the projection process $X_t = \pi(\xi_t)$ intrinsically on M , which does not depend on

the initial frame. Therefore, although as a process the sample paths of X_t depends on the initial frame (this is not surprising since in the Euclidean case we also don't have a canonical Brownian motion if we do not fix the frame $\{e_1, \dots, e_d\}$ in advance), the law of X_t will not. In this way we obtain a canonical PDE on M associated with the original G -function, which can be regarded as the generating PDE governing the law of X_t . The process X_t can be defined as a G -Brownian motion on M and the generating PDE will play the role of the canonical Wiener measure (the solution of the martingale problem for the operator $\frac{1}{2}\Delta_M$) on M in the nonlinear setting.

The construction of such frame bundle for a G -function with an arbitrary invariant group $I(G)$ is not clear to us at the moment. However, in the case when $I(G)$ is the orthogonal group $O(d)$, which contains a wide and interesting class of G -functions, there is a very natural frame bundle serving us well for the purpose: the orthonormal frame bundle $\mathcal{O}(M)$.

From now on, let G be given by (31) with $I(G) = O(d)$.

The orthonormal frame bundle $\mathcal{O}(M)$ over M is defined by

$$\mathcal{O}(M) = \cup_{x \in M} \mathcal{O}_x(M),$$

where the fibre $\mathcal{O}_x(M)$ is the set of orthonormal bases of $T_x M$. Since M is compact, $\mathcal{O}(M)$ is a compact submanifold of $\mathcal{F}(M)$. Moreover, since the Levi-Civita connection is compatible with the Riemannian metric g , parallel transport preserves the fibers of $\mathcal{O}(M)$. Therefore, statements about $\mathcal{F}(M)$ before on the horizontal aspect can be carried through in the case of $\mathcal{O}(M)$ directly. In particular, the fundamental horizontal fields H_α can be restricted to $\mathcal{O}(M)$. The only difference is in the vertical direction: the fibre becomes orthonormal frames, and the structure group which acts on fibers becomes the orthogonal group; the dimension in the vertical direction is reduced to $\frac{d(d-1)}{2}$.

For $\xi \in \mathcal{O}_x(M)$, according to Section 6, let $U_{t,\xi} \in \mathcal{O}(M)$ be the unique solution of the following RDE over $[0, 1]$:

$$\begin{cases} dU_{t,\xi} = H_\alpha(U_{t,\xi}) \circ dB_t^\alpha, \\ U_{0,\xi} = \xi. \end{cases} \quad (43)$$

Let $X_{t,\xi} = \pi(U_{t,\xi})$ be the projection of $U_{t,\xi}$ onto M .

Definition 8. $X_{t,\xi}$ is called a G -Brownian motion on the Riemannian manifold M with respect to the the initial orthonormal frame $\xi \in \mathcal{O}_x(M)$, and $U_{t,\xi}$ is called a horizontal G -Brownian motion in $\mathcal{O}(M)$ starting from ξ .

For any $\varphi \in C_{Lip}(M)$ (under the Riemannian distance), define

$$u(t, \xi) = \mathbb{E}^G[\varphi(X_{t,\xi})], \quad (t, \xi) \in [0, 1] \times \mathcal{O}(M).$$

Let $\hat{\varphi} = \varphi \circ \pi$ be the lifting of φ to $\mathcal{O}(M)$. It is obvious that

$$u(t, \xi) = \mathbb{E}^G[\hat{\varphi}(U_{t,\xi})].$$

By Theorem 12, we know that $u(t, \xi)$ is the unique viscosity solution of the following nonlinear parabolic PDE:

$$\begin{cases} \frac{\partial u}{\partial t} - G((\widehat{H_\alpha H_\beta} u)_{1 \leq \alpha, \beta \leq d}) = 0, \\ u(0, \xi) = \hat{\phi}(\xi), \end{cases} \quad (44)$$

on $\mathcal{O}(M)$.

The following result tells us that the law of $X_{t, \xi}$ depends only on the initial position x .

Proposition 10. *If $\xi, \eta \in \mathcal{O}_x(M)$, then*

$$u(t, \xi) = u(t, \eta).$$

Proof. For any fixed orthogonal matrix $Q \in O(d)$, let $\tilde{B}_t = QB_t$, which is an orthogonal transformation of the original G -Brownian motion B_t , and let $W_{t, \zeta}$ be the pathwise solution of the following RDE over $[0, 1]$:

$$\begin{cases} dW_{t, \zeta} = H_\alpha(W_{t, \zeta}) \circ d\tilde{B}_t^\alpha, \\ W_{0, \zeta} = \zeta \in \mathcal{O}(M), \end{cases} \quad (45)$$

on $\mathcal{O}(M)$. If we regard \tilde{B}_t as the solution of the SDE

$$d\tilde{B}_t = Q_\alpha dB_t^\alpha$$

starting from 0 with constant coefficients, then the RDE (45) is equivalent to

$$\begin{cases} dW_{t, \zeta} = H_\beta(W_{t, \zeta}) Q_\alpha^\beta \circ dB_t^\alpha, \\ W_{0, \zeta} = \zeta, \end{cases}$$

in which the generating vector fields are $H_\beta Q_\alpha^\beta$. Since the invariant group $I(G)$ of G is the orthogonal group, by Theorem 12 we know that the function

$$v(t, \zeta) = \mathbb{E}^G[\hat{\phi}(W_{t, \zeta})], \quad (t, \zeta) \in [0, 1] \times \mathcal{O}(M)$$

is the unique viscosity solution of the same PDE (44) on $\mathcal{O}(M)$. Therefore,

$$u(t, \xi) = v(t, \zeta), \quad \forall (t, \zeta) \in [0, 1] \times \mathcal{O}(M).$$

Now since $\xi, \eta \in \mathcal{O}_x(M)$, there exists some $Q \in O(d)$ such that $\xi = \eta Q$. Define $W_{t, \zeta}$ as before. By the previous discussion on the relation between different anti-developments, we know that

$$X_{t, \xi} = \pi(U_{t, \xi}) = \pi(W_{t, \eta}), \quad \forall t \in [0, 1].$$

Therefore,

$$\begin{aligned}
u(t, \xi) &= \mathbb{E}^G[\varphi \circ \pi(U_{t, \xi})] \\
&= \mathbb{E}^G[\varphi \circ \pi(W_{t, \eta})] \\
&= v(t, \eta) \\
&= u(t, \eta).
\end{aligned}$$

From Proposition 10, we know that $u(t, \xi)$ is invariant along each fibre. Therefore, the law of $X_{t, \xi}$ depends only on the initial position $x \in M$ but not on the initial frame ξ . We use $u(t, x)$ to denote $u(t, \xi)$, where x is the base point of ξ . In this situation it is possible to establish the PDE for $u(t, x)$ intrinsically on M by “projecting down” (44), which should become the generating PDE governing the law of $X_{t, \xi}$.

For any $u \in C^\infty(M)$, take an orthonormal frame $\xi = (\xi_1, \dots, \xi_d) \in \mathcal{O}_x(M)$, and consider the quantity

$$G((\text{Hess}u(\xi_\alpha, \xi_\beta))_{1 \leq \alpha, \beta \leq d}).$$

Since $I(G) = O(d)$, it is easy to see that the above quantity is independent of the orthonormal frame $\xi \in \mathcal{O}_x(M)$. In other words, G can be regarded as a functional of the Hessian, and the nonlinear second order differential operator $G(\text{Hess}(\cdot))$ is globally well-defined on M .

Now we have the following result.

Theorem 14. $u(t, x)$ is the unique viscosity solution of the following nonlinear heat equation on M :

$$\begin{cases} \frac{\partial u}{\partial t} - G(\text{Hess}u) = 0, \\ u(0, x) = \varphi(x). \end{cases} \quad (46)$$

Proof. It suffices to show that: if $f \in C^\infty(M)$, and $\hat{f} = f \circ \pi$ is the lifting of f to $\mathcal{O}(M)$, then for any $\xi = (\xi_1, \dots, \xi_d) \in \mathcal{O}_x(M)$,

$$\text{Hess}f(\xi_\alpha, \xi_\beta)(x) = H_\alpha H_\beta \hat{f}(\xi).$$

Note that uniqueness follows from the same reason as pointed out in the proof of Theorem 12 by using results in [1].

In fact, for any $\xi = (\xi_1, \dots, \xi_d) \in \mathcal{O}_x(M)$, let ξ_t be a horizontal curve through ξ such that $H_\beta(\xi)$ is tangent to ξ_t at $t = 0$, and let x_t be its projection onto M . It follows that the tangent vector of x_t at $t = 0$ is ξ_β , and

$$\begin{aligned}
H_\beta \hat{f}(\xi) &= \frac{d\hat{f}(\xi_t)}{dt} \Big|_{t=0} \\
&= \frac{df(x_t)}{dt} \Big|_{t=0} \\
&= \langle \xi_\beta, \nabla f(x) \rangle_g.
\end{aligned}$$

Therefore, if now assume that ξ_t is a horizontal curve through ξ with tangent vector $H_\alpha(\xi)$ at ξ and still $x_t = \pi(\xi_t)$, then

$$\begin{aligned}
H_\alpha H_\beta \hat{f}(\xi) &= H_\alpha \langle \xi_\beta, \nabla f(\pi(\xi)) \rangle_g \\
&= \frac{d}{dt} \Big|_{t=0} \langle \xi_{\beta,t}, \nabla f(x_t) \rangle_g \\
&= \left\langle \frac{D\xi_{\beta,t}}{dt} \Big|_{t=0}, \nabla f(x) \right\rangle_g + \langle \xi_\beta, \nabla_{\xi_\alpha} \nabla f(x) \rangle_g \\
&= \text{Hess}f(\xi_\alpha, \xi_\beta)(x),
\end{aligned}$$

where we've used the fact that $\xi_{\beta,t}$ is parallel along x_t .

Since $X_{t,\xi}$ is the projection of $U_{t,\xi}$ and $U_{t,\xi}$ is the solution of the RDE (43) which is equivalent to an Itô type SDE from an extrinsic point of view, by Theorem 14 we can see that as a process on M the law of the G -Brownian motion $X_{t,\xi}$ is characterized by the nonlinear parabolic PDE (46).

Example 5. When G is given by a functional of trace, as in Example 4 (1), (2), the generating PDE (46) takes a more explicit form in terms of the Laplace-Beltrami operator Δ_M on M . This is due to the fact that

$$\Delta_M = \text{tr}(\text{Hess}).$$

For instance, if $G(A) = \frac{1}{2} \text{tr}(A)$, then (46) becomes the classical heat equation on M :

$$\frac{\partial u}{\partial t} - \frac{1}{2} \Delta_M u = 0,$$

which governs the law of classical Brownian motion on M (see [12], [13]). If G is given by

$$G(A) = \frac{1}{2} (\mu(\text{tr}A)^+ - \lambda(\text{tr}A)^-),$$

where $0 \leq \lambda < \mu$, then (46) becomes

$$\frac{\partial u}{\partial t} - \frac{1}{2} (\mu(\Delta_M u)^+ - \lambda(\Delta_M u)^-) = 0.$$

It is a generalization of the one-dimensional Barenblatt equation to higher dimensions in a Riemannian geometric setting.

As pointed out before, as a process the G -Brownian motion $X_{t,\xi}$ on M depends on the initial orthonormal frame ξ and hence there is not a canonical choice of a particular one. However, if we consider the path space $W(M) = C([0, 1]; M)$, then for each $x \in M$, it is possible to define a canonical sublinear expectation \mathbb{E}_x on the space $\mathcal{H}(M)$ of functionals on $W(M)$ of the form

$$f(x_{t_1}, \dots, x_{t_n}),$$

where $0 \leq t_1 < \dots < t_n \leq 1$ and $f \in C_{Lip}(M)$, such that under \mathbb{E}_x the law of the coordinate process is characterized by the PDE (46) with $\mathbb{E}_x[\varphi(x_0)] = \varphi(x)$ for any $\varphi \in C_{Lip}(M)$.

To see this, we will define \mathbb{E}_x explicitly. We use $u_\varphi(t, x)$ to denote the solution of (46), emphasizing the dependence on φ . For a functional of the form $f(x_t)$, we simply define

$$\mathbb{E}_x[f(x_t)] := u_f(t, x).$$

For a functional of the form $f(x_s, x_t)$, $\mathbb{E}_x f(x_s, x_t)$ should be defined by $\mathbb{E}^G[f(X_{s,\xi}, X_{t,\xi})]$, where $X_{t,\xi}$ is a G -Brownian motion on M with respect to an initial orthonormal frame $\xi \in \mathcal{O}_x(M)$. Similar to the proof of Theorem 11 we know that

$$\begin{aligned} \mathbb{E}^G[f(X_{s,\xi}, X_{t,\xi})] &= \mathbb{E}^G[\mathbb{E}^G[f(X_{s,\xi}, X_{t,\xi}) | \Omega_s]] \\ &= \mathbb{E}^G[\mathbb{E}^G[f(\pi(U_{s,\xi}), \pi(U_{t,\xi})) | \Omega_s]] \\ &= \mathbb{E}^G[\mathbb{E}^G[f(\pi(\eta), X_{t-s,\eta}) | \eta = U_{s,\xi}]]. \end{aligned}$$

But since the law of $X_{t-s,\eta}$ does not depend on the initial orthonormal frame η , we obtain that

$$\mathbb{E}^G[f(\pi(\eta), X_{t-s,\eta}) | \eta = U_{s,\xi}] = u_{f(X_{s,\xi}, \cdot)}(t-s, X_{s,\xi}).$$

Therefore, we define

$$\mathbb{E}_x[f(x_s, x_t)] := \mathbb{E}^G[f(X_{s,\xi}, X_{t,\xi})] = u_g(s, x),$$

where

$$g(y) := u_{f(y, \cdot)}(t-s, y), \quad y \in M.$$

Inductively, assume that

$$u_f^{(n)}(t_1, \dots, t_n, x) = \mathbb{E}_x[f(x_{t_1}, \dots, x_{t_n})]$$

is already defined. For a functional of the form $f(x_{t_1}, \dots, x_{t_{n+1}})$, define

$$\mathbb{E}_x[f(x_{t_1}, \dots, x_{t_{n+1}})] := u_g(t_1, x),$$

where

$$g(y) := u_{f(y, \dots, \cdot)}^{(n)}(t_2 - t_1, \dots, t_{n+1} - t_1, y), \quad y \in M.$$

Then \mathbb{E}_x is the desired sublinear expectation on $\mathcal{H}(M)$.

Remark 8. As we've pointed out before, for noncompact Riemannian manifolds, the RDE (43) may possibly explode at some finite time and so may the corresponding G -Brownian motion as well. An interesting question is the study of explosion criterion. It might depend on the curvature and topology of the Riemannian manifold.

On the other hand, for those G -functions with the same invariant group, they may have some special features in common; while for those with different invariant groups, their structure should be very different. The study of classification of G -functions in terms of the invariant group is interesting, and it might give us some hints on generalizing our results to the case when $I(G) \neq O(d)$. We believe that in some cases it is still possible to construct a proper frame bundle with structure group $I(G)$ on which we can apply similar techniques in this section. But in some extreme

cases, for instance when $I(G)$ is a finite group as in Example 3, it seems difficult to proceed along this direction unless we have a globally defined frame field over the Riemannian manifold M , which is usually not true. We might need some very different methods for those extreme cases.

8 Conclusion

To summarize, the motivation of the present article is to understand nonlinear diffusion processes from the view of rough path theory, and in particular, to understand the intrinsic and geometric nature of the nonlinear heat flow. Along this direction, the pathwise approach under the framework of rough paths seems to be a right tool, since it is not natural to use Itô's formulation from the geometric point of view, and localization technique from Itô's perspective is not well understood in the nonlinear situation.

In this article, we study the geometric rough path nature of sample paths of G -Brownian motion, which enables us to establish differential equations driven by G -Brownian motion in a pathwise sense. Furthermore, we establish the fundamental relation between the two types of differential equations driven by G -Brownian motion. Such relation enables us to develop nonlinear diffusion processes on a differentiable manifold and G -Brownian motion on a Riemannian manifold from the pathwise point of view. The pathwise approach seems to be quite natural if we aim at understanding the intrinsic and geometric nature of nonlinear diffusion processes. Finally, we establish the generating nonlinear heat equation of G -Brownian motion on a Riemannian manifold and construct the associated canonical sublinear expectation on the path space for a class of nonlinear G -functions whose invariant group is the orthogonal group. Although such class of G -functions contains the most important and interesting examples so far, it remains an open problem for the general case (in particular, the singular case) where the use of orthogonal frame bundle is no longer applicable.

Acknowledgements The authors wish to thank Professor Shige Peng for so many valuable suggestions on the present article. The authors are supported by the Oxford-Man Institute at the University of Oxford.

References

1. D. Azagra, J. Ferrera, and B. Sanz, Viscosity solutions to second order partial differential equations on riemannian manifolds. *J. Differential Equations*. **245**(2), 307–336 (2009)
2. S. Chern, W. Chen, and K. Lam, *Lectures on Differential Geometry* (World Scientific, 1999)
3. G. Choquet, Theory of capacities. *Ann. Inst. Fourier*. **5**, 87 (1953)
4. M.G. Crandall, H. Ishii, and P.L. Lions, User's guide to viscosity solutions of second order partial differential equations. *Bull. Amer. Math. Soc.* **27**(1), 1–67 (1992)

5. G. de Rham, *Riemannian Manifolds* (Springer-Verlag, 1984)
6. C. Dellacherie, *Capacités et Processus Stochastiques* (Springer-Verlag, 1972)
7. L. Denis, M. Hu, and S. Peng, Function spaces and capacity related to a sublinear expectation: application to G-Brownian motion paths. *Potential Anal.* **34**(2), 139–161 (2011)
8. K.D. Elworthy, *Stochastic Differential Equations on Manifolds* (Springer, 1998)
9. P.K. Friz and N.B. Victoir, *Multidimensional Stochastic Processes as Rough Paths : Theory and Applications* (Cambridge University Press, 2010)
10. F. Gao, Pathwise properties and homeomorphic flows for stochastic differential equations driven by G-Brownian motion. *Stochastic Process. Appl.* **119**(10), 3356–3382 (2009)
11. B.M. Hambly and T.J. Lyons, Stochastic area for Brownian motion on the sierpinski gasket. *Ann. Probab.* **26**(1), 132–148 (1998)
12. E.P. Hsu, *Stochastic Analysis on Manifolds* (Amer Mathematical Society, 2002)
13. N. Ikeda and S.Watanabe, *Stochastic Differential Equations and Diffusion Processes* (North- Holland, 1989)
14. M. Kac, On distributions of certain wiener functionals. *Trans. Amer. Math. Soc.* **65**(1): 1–13, 1949.
15. I.A. Karatzas and S.E. Shreve, *Brownian Motion and Stochastic Calculus* (Springer-Verlag, 1991)
16. S. Kobayashi and K. Nomizu, *Foundations of Differential Geometry, I and II* (Interscience, New York, 1963 and 1969)
17. T.J. Lyons, Differential equations driven by rough signals. *Rev. Mat. Iberoam.* **14**(2), 215–310 (1998)
18. T.J. Lyons, C. Michael, and T. Lévy, *Differential Equations Driven by Rough Paths* (Springer, Berlin, 2007)
19. T.J. Lyons and Z. Qian, *System Control and Rough Paths* (Oxford University Press, 2002)
20. K. Nomizu and H. Ozeki, The existence of complete riemannian metrics. *Proc. Amer. Math. Soc.* **12**(6), 889–891 (1961)
21. É. Pardoux and S. Peng, Adapted solution of a backward stochastic differential equation. *Syst. Control Lett.* **14**(1), 55–61 (1990)
22. É. Pardoux and S. Peng, Backward stochastic differential equations and quasilinear parabolic partial differential equations. *Stochastic Partial Differential Equations and Their Applications*. pp. 200–217 (Springer, 1992)
23. É. Pardoux and S. Peng, Backward doubly stochastic differential equations and systems of quasilinear SPDEs. *Probab. Theory Related Fields.* **98**(2), 209–227 (1994)
24. S. Peng, Backward SDE and related g-expectation. *Pitman Research Notes in Mathematics series*. pp. 141–160 (1997)
25. S. Peng, G-expectation, G-Brownian motion and related stochastic calculus of Itô type. *Stoch. Anal. Appl.* **2**, 541–567 (2007)
26. S. Peng, Multi-dimensional G-Brownian motion and related stochastic calculus under G-expectation. *Stochastic Process. Appl.* **118**(12), 2223–2253 (2008)
27. S. Peng, Nonlinear expectations and stochastic calculus under uncertainty. *arXiv preprint, arXiv : 1002.4546* (2010)
28. E. Wong and M. Zakai, On the relation between ordinary and stochastic differential equations. *Internat. J. Engrg. Sci.* **3**(2), 213–229 (1965)
29. K. Yosida, *Functional Analysis* (Springer-Verlag, 1980)