

Backward problems for stochastic differential equations on the Sierpinski gasket

Xuan Liu^{*} and Zhongmin Qian[†]

Abstract

In this paper, we study the non-linear backward problems (with deterministic or stochastic durations) of stochastic differential equations on the Sierpinski gasket. We prove the existence and uniqueness of solutions of backward stochastic differential equations driven by Brownian martingale (defined in Section 2) on the Sierpinski gasket constructed by S. Goldstein and S. Kusuoka. The exponential integrability of quadratic processes for martingale additive functionals is obtained, and as an application, a Feynman-Kac representation formula for weak solutions of semi-linear parabolic PDEs on the gasket is also established.

Keywords Sierpinski gasket, backward stochastic differential equations, semi-linear parabolic equations

Mathematics Subject Classification (2000) 28A80, 60H10, 60H30

1 Introduction

Interests on diffusions on fractals initially came from mathematical physics such as percolation clusters near the percolation thresholds. (See [32] and references therein.) Diffusions on fractals are also of significance from mathematical point of views:

(i) Calculus can be established on a manifold with a differential structure, while there are interesting spaces appearing in applications which possess no suitable smooth structures. It is important and tempting to consider calculus on fractals as archetypical examples of singular spaces;

(ii) By utilizing the general theory of Dirichlet forms, linear analysis may be well studied via the corresponding Markov semigroups in a quite general setting.

Dynamic systems in physics (e.g. fluid dynamics) and mathematical finance (e.g. optimal stochastic controls) are usually described by non-linear partial differential equations. As suggested by the situation on Euclidean spaces, knowledge on non-linear PDEs on the Sierpinski gasket can be obtained by studying the corresponding stochastic dynamic systems, which will be the subject of this paper as a part of an investigation of non-linear analysis on fractals.

Brownian motion on the Sierpinski gasket was first constructed by S. Goldstein [12] and S. Kusuoka [26] as the limit of sequences of random walks. A similar construction was used by M. Barlow and E. Perkins in [3], where heat kernel estimates were obtained. Brownian motion on the gasket is a diffusion process symmetric with respect to the Hausdorff measure, and therefore, gives rise to a Dirichlet form. The standard Laplacian on the Sierpinski gasket is defined to be the associated self-adjoint operator. J. Kigami [19] gave an analytic construction of the Dirichlet form using products of stochastic matrices. There have been many works on the study of Brownian motion and the Laplacian on the Sierpinski gasket (see, for example, [3, 11, 23, 2, 21], and etc.). Several definitions of gradients of functions with finite energies have been introduced and studied with applications to the study of non-linear PDEs and differential forms on fractals (cf. [20, 35, 37, 6, 14, 15, 17, 4] and references

^{*}Mathematical Institute, University of Oxford, Oxford, OX2 6GG, United Kingdom. Email: xuan.liu@maths.ox.ac.uk

[†]Research supported partly by an ERC grant(Grant Agreement No. 291244 ESig). Mathematical Institute, University of Oxford, Oxford, OX2 6GG, United Kingdom. Email: zhongmin.qian@maths.ox.ac.uk

in these works). See Remark 3.12-(ii) for more details and comments on the connection between the definition of gradients introduced in Definition 3.11 below and those in the aforementioned literature.

In this paper, we develop a theory of backward stochastic differential equations (BSDEs hereafter for short) on the Sierpinski gasket. As an application, we derive a representation formula for solutions of semi-linear parabolic PDEs on the gasket, which will be formulated later in Section 3. In contrast to the BSDE theory on Euclidean spaces, BSDEs on the gasket have two mutually singular drift terms, which is due to the singularity between the Hausdorff measure as the volume measure and the Kusuoka measure as the energy measure (see Section 2 for more details). Singularity of different drift terms introduces significant difficulties in the study of BSDEs on the gasket, one of which is the exponential integrability of a quadratic process (cf. Theorem 3.5 below). Our approach to the exponential integrability is based on moment estimates for the quadratic processes by expressing the moments as iterated integrals and by using the heat kernel estimate. The study of BSDEs on the gasket here is a specific case and a continuation of the papers [24, 25] investigating BSDEs associated with Dirichlet forms and their application to partial differential equations. (See Section 3 for more related works and connections between BSDEs studied in the current paper and those in literature.) Though results in this paper are stated and proved specifically for 2-dimensional Sierpinski gasket, we however believe that our results also hold for higher-dimensional Sierpinski gaskets, and the proofs given in this paper should be easily adapted to higher-dimensional cases.

The paper is organized as follows. In Section 2, we introduce notations used in the paper and recall several results that will be needed in following sections. The main results are formulated in Section 3. Section 4 is devoted to several results on Brownian motion which will be needed in later sections. In particular, a representation theorem for square-integrable martingales is given as an immediate consequence of the results in [27], [13] and [31]. A result on the exponential integrability for quadratic processes is also given in this section, which is of interests in itself. In Section 5, we prove the existence and uniqueness of solutions to BSDEs with deterministic or stochastic durations. In the last section, Section 6, we give the proof of a Feynman-Kac representation of solutions of semi-linear parabolic PDEs, and as a result, derive the uniqueness of weak solutions.

2 Notations and several related results

In this section, we introduce several notations and notions which will be in force throughout this paper.

2.1 Diffusions and Dirichlet forms on the gasket

Let $V_0 = \{p_1, p_2, p_3\} \subseteq \mathbb{R}^2$ with $p_1 = (0, 0)$, $p_2 = (1, 0)$, $p_3 = (\frac{1}{2}, \frac{\sqrt{3}}{2})$, and $F_i : \mathbb{R}^2 \rightarrow \mathbb{R}^2$, $i = 1, 2, 3$ be the contraction mappings given by $F_i(x) = \frac{1}{2}(x + p_i)$, $x \in \mathbb{R}^2$, $i = 1, 2, 3$. Let $W_* = \{\omega = \omega_1\omega_2\omega_3 \dots : \omega_i \in \{1, 2, 3\}, i \in \mathbb{N}_+\}$. For $\omega = \omega_1\omega_2\omega_3 \dots \in W_*$ and $m \in \mathbb{N}_+$, we denote $[\omega]_m = \omega_1 \dots \omega_m$ and $F_{[\omega]_m} = F_{\omega_1} \circ F_{\omega_2} \circ \dots \circ F_{\omega_m}$. Let $V_m = \bigcup_{\omega \in W_*} F_{[\omega]_m}(V_0)$, $m \in \mathbb{N}_+$, and $V_* = \bigcup_{m=0}^{\infty} V_m$. The (2-dimensional) *Sierpinski gasket* is defined to be the closure $\mathbb{S} = \text{cl}(V_*)$.

The Dirichlet form on \mathbb{S} can be introduced via a finite difference scheme described below. For any functions u, v on V_* , define

$$\mathcal{E}^{(m)}(u, v) = \sum_{x, y \in V_m, |x-y|=2^{-m}} \frac{1}{2} \left(\frac{5}{3}\right)^m [u(x) - u(y)][v(x) - v(y)], \quad m \in \mathbb{N}.$$

For each function u on V_* , $\mathcal{E}^{(m)}(u, u)$ is non-decreasing in m , and $\mathcal{E}(u, u) = \lim_{m \rightarrow \infty} \mathcal{E}^{(m)}(u, u)$ exists (possibly infinite). Let $\mathcal{F}(\mathbb{S}) = \{u : \mathcal{E}(u, u) < \infty\}$. Every $u \in \mathcal{F}(\mathbb{S})$ is continuous on V_* and therefore can be extended to a continuous function on \mathbb{S} ([22, Section 2.2 and Section 3.1]). In other words, $\mathcal{F}(\mathbb{S}) \subseteq C(\mathbb{S})$, which in fact follows easily from the following

$$\text{osc}_{\mathbb{S}}(u) \leq C_* \sqrt{\mathcal{E}(u, u)}, \quad u \in \mathcal{F}(\mathbb{S}), \quad (2.1)$$

where $C_* > 0$ is a universal constant. (See [22, Lemma 2.3.9 and Theorem 3.3.4].) Since $\mathcal{F}(\mathbb{S}) \subseteq C(\mathbb{S})$, the empty set \emptyset is the only subset of \mathbb{S} with zero capacity. As a consequence of the definition of $\mathcal{E}^{(m)}$, it is seen that $\mathcal{E}^{(m+1)}(u, v) = \sum_{i=1,2,3} \frac{5}{3} \mathcal{E}^{(m)}(u \circ F_i, v \circ F_i)$, $m \in \mathbb{N}$, which implies the following self-similar property of \mathcal{E} :

$$\mathcal{E}(u, v) = \sum_{i=1,2,3} \frac{5}{3} \mathcal{E}(u \circ F_i, v \circ F_i), \quad u, v \in \mathcal{F}(\mathbb{S}). \quad (2.2)$$

For any function u on V_0 , there exists a unique $h \in \mathcal{F}(\mathbb{S})$ such that $h|_{V_0} = u$ and

$$\mathcal{E}(h, h) = \min\{\mathcal{E}(v, v) : v \in \mathcal{F}(\mathbb{S}) \text{ and } v|_{V_0} = u\}.$$

(See [22, Corollary 3.2.15].) The above function h is called the *harmonic function* in \mathbb{S} with boundary value u , and denoted by $h = Hu$. For the harmonic function Hu with boundary value u , we have

$$\mathcal{E}(Hu, Hu) = \mathcal{E}^{(0)}(u, u) = \frac{3}{2} u^t \mathbf{P} u, \quad (2.3)$$

where

$$\mathbf{P} = \frac{1}{3} \begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix}. \quad (2.4)$$

The values of Hu on V_* is given by

$$(Hu) \circ F_{[\omega]_m} = \mathbf{A}_{[\omega]_m} u \quad \text{for all } \omega \in W_*, m \in \mathbb{N},$$

where

$$\mathbf{A}_{[\omega]_m} = \mathbf{A}_{\omega_m} \mathbf{A}_{\omega_{m-1}} \cdots \mathbf{A}_{\omega_1} \quad (2.5)$$

with \mathbf{A}_i , $i = 1, 2, 3$ being the linear operators having matrix representations

$$\mathbf{A}_1 = \frac{1}{5} \begin{bmatrix} 5 & 0 & 0 \\ 2 & 2 & 1 \\ 2 & 1 & 2 \end{bmatrix}, \quad \mathbf{A}_2 = \frac{1}{5} \begin{bmatrix} 2 & 2 & 1 \\ 0 & 5 & 0 \\ 1 & 2 & 2 \end{bmatrix}, \quad \mathbf{A}_3 = \frac{1}{5} \begin{bmatrix} 2 & 1 & 2 \\ 1 & 2 & 2 \\ 0 & 0 & 5 \end{bmatrix}. \quad (2.6)$$

Notice that, in contrast to the definition of $F_{[\omega]}$, the order of matrices \mathbf{A}_{ω_i} on the right hand side of (2.5) is reversed.

Let μ be the (normalized) *Hausdorff measure* on \mathbb{S} ; that is, μ is the unique Borel probability measure on \mathbb{S} such that $\mu(F_{[\omega]_m}(\mathbb{S})) = 3^{-m}$ for all $\omega \in W_*$, $m \in \mathbb{N}$. Then $(\mathcal{E}, \mathcal{F}(\mathbb{S}))$ is a regular Dirichlet form on $L^2(\mathbb{S}; \mu)$, called the *standard Dirichlet form* on \mathbb{S} , of which the associated non-positive self-adjoint operator will be denoted by \mathcal{L} . According to the theory of Dirichlet forms and Markov processes (see [10, Chapter 7]), there exists a standard Hunt process $(\Omega, \mathcal{F}, \{X_t\}, \{\theta_t\}, \{\mathbb{P}_x\}_{x \in \mathbb{S} \cup \{\Delta\}})$ with state space \mathbb{S} , where Δ is an isolated point adjoined to \mathbb{S} , and $\theta_t : \Omega \rightarrow \Omega$, $t \geq 0$ are the shift operators. The Hunt process $\{X_t\}$ is a diffusion on \mathbb{S} , called (the reflected) *Brownian motion* on \mathbb{S} , of which the associated Markov semigroup will be denoted by $\{P_t\}$. Let $\mathcal{P}(\mathbb{S})$ be the family of all Borel probability measures on \mathbb{S} . For each $\lambda \in \mathcal{P}(\mathbb{S})$, the probability measure \mathbb{P}_λ on Ω is defined by

$$\mathbb{P}_\lambda(E) = \int_{\mathbb{S}} \mathbb{P}_x(E) \lambda(dx), \quad E \in \mathcal{F}.$$

The expectation with respect to \mathbb{P}_λ will be denoted by \mathbb{E}_λ , and we denote by $\{\mathcal{F}_t\}$ the *minimal admissible filtration* determined by $\{X_t\}$; that is, $\mathcal{F}_t = \bigcap_{\lambda \in \mathcal{P}(\mathbb{S})} \mathcal{F}_t^\lambda$, $t \geq 0$, where \mathcal{F}_t^λ is the \mathbb{P}_λ -completion of $\sigma(X_r : r \leq t)$ in \mathcal{F} . An $\{\mathcal{F}_t\}$ -adapted process A_t is called an *additive functional* if $A_{t+s} = A_s + A_t \circ \theta_s$ for all $t, s \geq 0$.

Let $\mathcal{F}(\mathbb{S} \setminus V_0) = \{u \in \mathcal{F}(\mathbb{S}) : u|_{V_0} = 0\}$. The restriction of \mathcal{E} on $\mathcal{F}(\mathbb{S} \setminus V_0)$ is also a Dirichlet form with $\mathcal{F}(\mathbb{S} \setminus V_0)$ as its Dirichlet space. Dirichlet spaces $\mathcal{F}(\mathbb{S})$ and $\mathcal{F}(\mathbb{S} \setminus V_0)$ correspond to the Neumann boundary conditions and the Dirichlet boundary conditions respectively. (See [22, Theorem 3.7.9].) Let σ_{V_0} be the hitting time $\sigma_{V_0} = \inf \{t > 0 : X_t \in V_0\}$. We define the *killed Brownian motion* $\{X_t^0\}$ by killing $\{X_t\}$ on hitting V_0 ; that is, $X_t^0 = X_t$ if $t < \sigma_{V_0}$, and $X_t^0 = \Delta$ if $t \geq \sigma_{V_0}$. Then $\{X_t^0\}$ is a μ -symmetric Hunt process on $\mathbb{S} \setminus V_0$ with $(\mathcal{E}, \mathcal{F}(\mathbb{S} \setminus V_0))$ as its associated Dirichlet form, and the associated semigroup, denoted by $\{P_t^0\}$, is given by $P_t^0(x, E) = \mathbb{P}_x(X_t \in E, t < \sigma_{V_0})$ for all $x \in \mathbb{S} \setminus V_0$, $E \in \mathcal{B}(\mathbb{S} \setminus V_0)$. It can be easily shown that if A_t is an additive functional, then $A_t^{\sigma_{V_0}} = A_{t \wedge \sigma_{V_0}}$ is an additive functional with the shift operators θ_t replaced by $\theta_{t \wedge \sigma_{V_0}}$.

2.2 Kusuoka measure and Brownian martingale

Let \mathbf{P} and \mathbf{A}_i , $i = 1, 2, 3$ be the matrices in (2.4) and (2.6) respectively, and let $\mathbf{Y}_i = \mathbf{P}\mathbf{A}_i\mathbf{P}$, $i = 1, 2, 3$.

Definition 2.1. The *Kusuoka measure* ν is defined to be the unique probability measure ν on \mathbb{S} such that

$$\nu(F_{[\omega]_m}(\mathbb{S})) = \frac{1}{2} \left(\frac{5}{3}\right)^m \text{tr}(\mathbf{Y}_{[\omega]_m}^t \mathbf{Y}_{[\omega]_m}) \quad \text{for all } \omega \in W_*, m \in \mathbb{N},$$

where, similar to (2.5), $\mathbf{Y}_{[\omega]_m} = \mathbf{Y}_{\omega_m} \mathbf{Y}_{\omega_{m-1}} \cdots \mathbf{Y}_{\omega_1}$.

Remark 2.2. The Kusuoka measure ν is singular to the Hausdorff measure μ . (See [27, Corollary (2.15) and Example 1, p. 678].)

Definition 2.3. For each $u \in \mathcal{F}(\mathbb{S})$, the *energy measure* $\nu_{\langle u \rangle}$ of u is defined to be the unique Borel measure on \mathbb{S} such that

$$\int_{\mathbb{S}} \phi d\nu_{\langle u \rangle} = \mathcal{E}(\phi u, u) - 2\mathcal{E}(\phi, u^2) \quad \text{for all } \phi \in \mathcal{F}(\mathbb{S}).$$

For any $u, v \in \mathcal{F}(\mathbb{S})$, the *mutual energy measure* $\nu_{\langle u, v \rangle}$ of u and v is defined by polarization $\nu_{\langle u, v \rangle} = \frac{1}{4}(\nu_{\langle u+v \rangle} - \nu_{\langle u-v \rangle})$.

Remark 2.4. The measure ν is an energy dominant; that is, $\nu_{\langle u \rangle} \ll \nu$ for all $u \in \mathcal{F}(\mathbb{S})$. (See [27, Lemma (5.1)].)

Definition 2.5. For any positive additive functional A_t , the *Revuz measure* ν_A of A is defined to be the unique Radon measure on \mathbb{S} such that

$$\int_{\mathbb{S}} \phi d\nu_A = \lim_{t \rightarrow 0} \frac{1}{t} \mathbb{E}_\mu \left(\int_0^t \phi(X_r) dA_r \right) \quad \text{for all } \phi \in \mathcal{B}_+(\mathbb{S}).$$

According to stochastic analysis for Brownian motions on \mathbb{R}^d , any square-integrable martingale adapted to the Brownian filtration can be written as an Itô integral against Brownian motion, which is the crucial ingredient in solving SDEs backward. On fractal, representations for martingale additive functionals were first proved by S. Kusuoka [27], and generalized to a large class of self-similar sets including nested fractals by M. Hino [13]. The following representation theorem is the specific case of [27, Theorem (5.4)] applied to the Sierpinski gasket.

Theorem 2.6. *There exists a martingale additive functional W_t satisfying the following:*

- (i) W_t has ν as its energy measure; that is, $\nu_{\langle W \rangle} = \nu$;
- (ii) for any $u \in \mathcal{F}(\mathbb{S})$, there exists a unique $\zeta \in L^2(\mathbb{S}; \nu)$ such that

$$M_t^{[u]} = \int_0^t \zeta(X_r) dW_r \quad \text{for all } t \geq 0, \quad (2.7)$$

where $M^{[u]}$ is the martingale part of $u(X_t) - u(X_0)$.

Remark 2.7. Notice that the martingale additive functional W_t in Theorem 2.6 is not unique. In fact, let g be any Borel measurable function on \mathbb{S} such that $|g| = 1$. Then $\int_0^t g(X_r) dW_r$ is also a martingale additive functional satisfying the properties in Theorem 2.6. However, we may have a canonical choice if the sign of the integrand $\zeta \in L^2(\mathbb{S}; \nu)$ in (2.7) is specified. More precisely, we make the following definition.

Definition 2.8. The *Brownian martingale* is defined to be the unique martingale additive functional W_t such that W_t satisfies the properties in Theorem 2.6, and that, if h is the harmonic function with boundary value $1_{\{p_1\}}$ and $M_t^{[h]} = \int_0^t \zeta(X_r) dW_r$, then $\zeta < 0$ ν -a.e.

Remark 2.9. The sign of ζ in Definition 2.8 is chosen as a convention.

We shall always denote by W_t the Brownian martingale in the rest of this paper.

3 Formulation of main results

Linear BSDEs were first introduced by J. Bismut to establish the Pontryagin maximum principle in stochastic control theory (see [5, 39] and etc.), while the theory of non-linear BSDEs was developed in E. Pardoux and S. Peng [29]. The celebrated Feynman-Kac formula was also generalized to non-linear cases in [30] using BSDEs. There have been a large amount of works on the theory of BSDEs. For example, BSDEs associated with (non-symmetric) second-order elliptic operators of divergence forms on Euclidean spaces were studied in [34, 1] and applied to study semi-linear parabolic PDEs involving divergences of measurable vector fields, where an Itô-Fukushima decomposition for the diffusion process associated to the elliptic operator was derived in terms of forward-backward martingales, and a representation formula for solutions of parabolic PDEs was obtained. BSDEs and semi-linear parabolic equations on Hilbert spaces were investigated in [40] using methods from functional analysis and generalized Dirichlet forms. A martingale representation with countably many representing martingales for the infinite dimensional case was also proved in [40] in order to solve BSDEs on Hilbert spaces. In [24, 25], existence and uniqueness of solutions to a class of BSDEs associated with (not necessarily local) regular (or quasi-regular) Dirichlet forms were established, together with a probabilistic representation of solutions to semi-linear elliptic equations perturbed by smooth measures. It was also shown in [24] that the probabilistic solutions yielded by BSDEs coincide with the notion of weak solutions (called *solutions in the sense of duality* in [24]) under a transience assumption on the Dirichlet form.

In the current paper, we consider BSDEs, with deterministic or stochastic durations, driven by Brownian martingale W_t . We shall give a Feynman-Kac representation for solutions of semi-linear parabolic equations on \mathbb{S} , of which the meaning will be formulated later. On the one hand, the BSDEs considered here can be regarded as a specific case of those studied in [24, 25] for quite general (quasi-regular) Dirichlet forms. On the other hand, solutions to BSDEs in our case can be formulated in a more specific way (e.g. the martingale parts are given as stochastic integrals, and the exponential integrability assumption on quadratic processes as drifts can be verified), which is due to the specific setting of the gasket. Regarding probabilistic interpretations of parabolic equations in terms of BSDEs, the representation in Theorem 3.19 is an analogue of the representations established in [30, 34, 1].

To simplify notations, we shall adopt the convention

$$u(t) = u(t, \cdot) \text{ and } g(t, u(t)) = g(t, u(t, \cdot), \cdot)$$

for any function $u : [0, \infty) \times \mathbb{S} \rightarrow \mathbb{R}$ and $g : [0, \infty) \times \mathbb{S} \times \mathbb{R} \rightarrow \mathbb{R}$. Let $T \in (0, \infty)$ and $\lambda \in \mathcal{P}(\mathbb{S})$. Consider the BSDE on $(\Omega, \{\mathcal{F}_t^\lambda\}, \mathbb{P}_\lambda)$

$$\begin{cases} dY_t = -g(t, Y_t)dt - f(t, Y_t, Z_t)d\langle W \rangle_t + Z_t dW_t, & t \in [0, T), \\ Y_T = \xi, \end{cases} \quad (3.1)$$

where ξ is an \mathcal{F}_T^λ -measurable random variable, and the coefficients g, f satisfy the following measurability condition:

(M) $g : [0, \infty) \times \mathbb{R} \times \Omega \rightarrow \mathbb{R}$ and $f : [0, \infty) \times \mathbb{R} \times \mathbb{R} \times \Omega \rightarrow \mathbb{R}$ are measurable, and $t \mapsto g(t, y)$ and $t \mapsto f(t, y, z)$ are $\{\mathcal{F}_t^\lambda\}$ -adapted processes for all $y, z \in \mathbb{R}$.

As in the classical case, the BSDE (3.1) is interpreted as the corresponding (backward) integral equation

$$Y_t = \xi + \int_t^T g(r, Y_r) dr + \int_t^T f(r, Y_r, Z_r) d\langle W \rangle_r - \int_t^T Z_r dW_r, \quad t \in [0, T] \quad \mathbb{P}_\lambda\text{-a.s.}, \quad (3.2)$$

provided that all integrals can be well-defined.

Remark 3.1. Let us make some remarks on the different drift terms in the BSDE (3.1). The drift $g(t, Y_t)dt$ corresponds to Brownian motion X_t , while the drift $f(t, Y_t, Z_t)d\langle W \rangle_t$ corresponds to Brownian martingale W_t . As we shall see later (Lemma 4.3), the processes X_t and W_t have singular “speeds”, which is different from the situations on Euclidean spaces and reveals the fractal nature of \mathbb{S} . We also note that the process Z_t does not appear in the drift $g(t, Y_t)dt$. This is because, intuitively, Z_t represents the projection of Y_t on the “ W_t -direction”, which is orthogonal to the projection of Y_t on the “ X_t -direction”.

We first introduce the Banach spaces for solutions of BSDEs. As BSDEs with stochastic durations will also be considered, it is convenient to define directly the spaces of processes with stochastic running times.

Definition 3.2. Let τ be an $\{\mathcal{F}_t^\lambda\}$ -stopping time, and $\beta = (\beta_0, \beta_1) \in \mathbb{R}_+^2$. We define $\mathcal{V}_\lambda^\beta[0, \tau]$ to be the Banach space of all pairs (y, z) of $\{\mathcal{F}_t^\lambda\}$ -adapted processes such that

$$\begin{aligned} \|(y, z)\|_{\mathcal{V}_\lambda^\beta[0, \tau]}^2 &\triangleq \mathbb{E}_\lambda \left[\sup_{0 \leq t \leq \tau} \left(y_t^2 e^{2\beta_0 t + 2\beta_1 \langle W \rangle_t} + \int_t^\tau y_r^2 e^{2\beta_0 r + 2\beta_1 \langle W \rangle_r} dr \right. \right. \\ &\quad \left. \left. + \int_t^\tau (y_r^2 + z_r^2) e^{2\beta_0 r + 2\beta_1 \langle W \rangle_r} d\langle W \rangle_r \right) \right] < \infty. \end{aligned}$$

We shall simply write $\mathcal{V}_x^\beta[0, \tau]$ when $\lambda = \delta_x$ is the Dirac measure concentrated at $x \in \mathbb{S}$.

Definition 3.3. Let $\beta \in \mathbb{R}_+^2$. We say that the BSDE (3.1) admits a *solution* (Y, Z) in $\mathcal{V}_\lambda^\beta[0, T]$ if $(Y, Z) \in \mathcal{V}_\lambda^\beta[0, T]$ and satisfies (3.2). The solution (Y, Z) is said to be *unique* in $\mathcal{V}_\lambda^\beta[0, T]$ if $\|(Y - \bar{Y}, Z - \bar{Z})\|_{\mathcal{V}_\lambda^\beta[0, T]} = 0$ for any solution (\bar{Y}, \bar{Z}) of (3.1) in $\mathcal{V}_\lambda^\beta[0, T]$.

We should point out that the uniqueness of solutions defined in Definition 3.3 is the uniqueness of solutions in the space $\mathcal{V}_\lambda^\beta[0, T]$. In fact, uniqueness can also be defined for solutions not necessarily in $\mathcal{V}_\lambda^\beta[0, T]$. More precisely, we have the following definition for uniqueness of solutions.

Definition 3.4. The BSDE (3.1) is said to *admit at most one solution*, if

$$Y_t = \bar{Y}_t, \quad t \in [0, T], \quad \text{and} \quad \int_0^T (Z_r - \bar{Z}_r)^2 d\langle W \rangle_r = 0, \quad \mathbb{P}_\lambda\text{-a.s.} \quad (3.3)$$

for any two pairs (Y, Z) and (\bar{Y}, \bar{Z}) of $\{\mathcal{F}_t^\lambda\}$ -adapted processes satisfying (3.2).

Theorem 3.5. Let $\beta = (\beta_0, \beta_1) \in [1, \infty)^2$. Suppose that

$$\mathbb{E}_\lambda(\xi^2 e^{2\beta_1 \langle W \rangle_T}) < \infty, \quad (A.1)$$

and that, for all $t \in [0, T)$ and all $y, \bar{y}, z, \bar{z} \in \mathbb{R}$,

$$|g(t, y, \omega) - g(t, \bar{y}, \omega)| \leq \frac{K_0}{2} |y - \bar{y}| \quad \mathbb{P}_\lambda\text{-a.s.}, \quad (A.2)$$

$$|f(t, y, z, \omega) - f(t, \bar{y}, \bar{z}, \omega)| \leq \frac{K_0}{2} |y - \bar{y}| + K_1 |z - \bar{z}| \quad \mathbb{P}_\lambda\text{-a.s.}, \quad (\text{A.3})$$

where $K_0, K_1 > 0$ are some constants.

(a) The BSDE (3.1) admits at most one solution.

(b) If, in addition,

$$\mathbb{E}_\lambda \left(\int_0^T g(r, 0)^2 e^{2\beta_1 \langle W \rangle_r} dr \right) < \infty, \quad \mathbb{E}_\lambda \left(\int_0^T f(r, 0, 0)^2 e^{2\beta_1 \langle W \rangle_r} d\langle W \rangle_r \right) < \infty, \quad (\text{A.4})$$

for sufficiently large β_0, β_1 ($\beta_i > 36K_i^2$, $i = 0, 1$ will suffice), then (3.1) admits a unique solution $(Y, Z) \in \mathcal{V}_\lambda^\beta[0, T]$. Moreover,

$$\begin{aligned} \|(Y, Z)\|_{\mathcal{V}_\lambda^\beta[0, T]}^2 &\leq C \mathbb{E}_\lambda \left(\xi^2 e^{2\beta_0 T + 2\beta_1 \langle W \rangle_T} + \int_0^T g(r, 0)^2 e^{2\beta_0 r + 2\beta_1 \langle W \rangle_r} dr \right. \\ &\quad \left. + \int_0^T f(r)^2 e^{2\beta_0 r + 2\beta_1 \langle W \rangle_r} d\langle W \rangle_r \right), \end{aligned} \quad (\text{3.4})$$

where $C > 0$ is a constant depending only on K_0, K_1, β .

Remark 3.6. As we shall see in Theorem 3.19 below, (Dirichlet) terminal-boundary value problems for semi-linear parabolic PDEs correspond to BSDEs with bounded stochastic durations. Results of Theorem 3.5 can be extended to such case with no essential difficulties. More precisely, let τ be a bounded $\{\mathcal{F}_t^\lambda\}$ -stopping time and ξ be an \mathcal{F}_τ^λ -adaptive random variable. The definition of solutions of the BSDE on $(\Omega, \{\mathcal{F}_t^\lambda\}, \mathbb{P}_\lambda)$

$$\begin{cases} dY_t = -g(t, Y_t)dt - f(t, Y_t, Z_t)d\langle W \rangle_t + Z_t dW_t, & t \in [0, \tau), \\ Y_\tau = \xi, \end{cases} \quad (\text{3.5})$$

can be given by replacing T by τ in Definition 3.3, and the conclusions of Theorem 3.5 also hold with T replaced by τ .

We may also consider the BSDE (3.5) with stochastic duration τ which is not necessarily bounded, where g, f satisfy the measurability condition (M).

Definition 3.7. Let $\beta = (\beta_0, \beta_1) \in \mathbb{R}_+^2$. We say that (Y, Z) is a *solution of (3.5) in $\mathcal{V}_\lambda^\beta[0, \tau]$* if $(Y, Z) \in \mathcal{V}_\lambda^\beta[0, \tau]$ and satisfies the following:

$$\begin{aligned} Y_{t \wedge \tau} &= Y_{T \wedge \tau} + \int_{t \wedge \tau}^{T \wedge \tau} g(r, Y_r) dr + \int_{t \wedge \tau}^{T \wedge \tau} f(r, Y_r, Z_r) d\langle W \rangle_r \\ &\quad - \int_{t \wedge \tau}^{T \wedge \tau} Z_r dW_r \quad \text{for all } 0 \leq t \leq T < \infty \quad \mathbb{P}_\lambda\text{-a.s.}, \end{aligned} \quad (\text{3.6})$$

and

$$\lim_{T \rightarrow \infty} \mathbb{E}_\lambda \left(|Y_{T \wedge \tau} - \xi|^2 e^{2\beta_0(T \wedge \tau) + 2\beta_1 \langle W \rangle_{T \wedge \tau}} \right) = 0. \quad (\text{3.7})$$

The solution (Y, Z) is said to be *unique in $\mathcal{V}_\lambda^\beta[0, \tau]$* if $\|(Y - \bar{Y}, Z - \bar{Z})\|_{\mathcal{V}_\lambda^\beta[0, \tau]} = 0$ for any solution (\bar{Y}, \bar{Z}) of (3.5) in $\mathcal{V}_\lambda^\beta[0, \tau]$.

Remark 3.8. Clearly, if τ is bounded, then Definition 3.7 coincides with Definition 3.3 with T replaced by τ . (See Remark 3.6.)

Similar to Definition 3.3, we may discuss uniqueness of solutions not necessarily in the space $\mathcal{V}_\lambda^\beta[0, \tau]$.

Definition 3.9. The BSDE (3.5) is said to *admit at most one solution*, if (3.3) holds with T replaced by τ for any two pairs (Y, Z) and (\bar{Y}, \bar{Z}) of $\{\mathcal{F}_t^\lambda\}$ -adapted processes satisfying (3.6) and (3.7).

Theorem 3.10. Let $\beta = (\beta_0, \beta_1) \in [1, \infty)^2$. Suppose that (A.1)–(A.3) hold with T replaced by τ :

$$\mathbb{E}_\lambda(\xi^2 e^{2\beta_0\tau+2\beta_1\langle W \rangle_\tau}) < \infty, \quad (\text{A'.1})$$

$$|g(t, y, \omega) - g(t, \bar{y}, \omega)| \leq \frac{K_0}{2}|y - \bar{y}| \quad \mathbb{P}_\lambda\text{-a.s.}, \quad (\text{A'.2})$$

$$|f(t, y, z, \omega) - f(t, \bar{y}, \bar{z}, \omega)| \leq \frac{K_0}{2}|y - \bar{y}| + K_1|z - \bar{z}| \quad \mathbb{P}_\lambda\text{-a.s.}, \quad (\text{A'.3})$$

(a) The BSDE (3.5) admits at most one solution.

(b) If, in addition to (A'.1)–(A'.3),

$$(y - \bar{y})(g(t, y, \omega) - g(t, \bar{y}, \omega)) \leq -\kappa_0|y - \bar{y}|^2, \quad (\text{A'.4})$$

$$(y - \bar{y})(f(t, y, z, \omega) - f(t, \bar{y}, \bar{z}, \omega)) \leq -\kappa_1|y - \bar{y}|^2, \quad (\text{A'.5})$$

for all $t \in [0, \tau(\omega))$, $y, \bar{y}, z, \bar{z} \in \mathbb{R}$, and \mathbb{P}_λ -a.e. $\omega \in \Omega$, and

$$\mathbb{E}_\lambda\left(\int_0^\tau g(t, 0)^2 e^{2\beta_0 t+2\beta_1\langle W \rangle_t} dt\right) < \infty, \quad \mathbb{E}_\lambda\left(\int_0^\tau f(t, 0, 0)^2 e^{2\beta_0 t+2\beta_1\langle W \rangle_t} d\langle W \rangle_t\right) < \infty, \quad (\text{A'.6})$$

where $\kappa_0, \kappa_1 \in \mathbb{R}$ are some constants satisfying

$$\beta_0 - \kappa_0 > 0, \quad \beta_1 - \kappa_1 + \frac{K_1^2}{2} > 0. \quad (\text{A'.7})$$

Then BSDE (3.5) admits a unique solution $(Y, Z) \in \mathcal{V}_\lambda^\beta[0, \tau]$, and

$$\begin{aligned} \|(Y, Z)\|_{\mathcal{V}_\lambda^\beta[0, \tau]}^2 &\leq C \mathbb{E}_\lambda\left(\xi^2 e^{2\beta_0\tau+2\beta_1\langle W \rangle_\tau} + \int_0^\tau g(t, 0)^2 e^{2\beta_0 t+2\beta_1\langle W \rangle_t} dt \right. \\ &\quad \left. + \int_0^\tau f(t, 0, 0)^2 e^{2\beta_0 t+2\beta_1\langle W \rangle_t} d\langle W \rangle_t\right), \end{aligned} \quad (\text{3.8})$$

where $C > 0$ is a constant depending only on $\beta_0, \beta_1, \kappa_0, \kappa_1, K_0, K_1$.

The last result of this paper addresses the representation of weak solutions of semi-linear parabolic PDEs by those of corresponding BSDEs. We first formulate the meaning of semi-linear parabolic PDEs on \mathbb{S} . To this end, we introduce the following definition of (weak) gradients of functions in $\mathcal{F}(\mathbb{S})$.

Definition 3.11. For any $u \in \mathcal{F}(\mathbb{S})$, in view of Theorem 2.6, we define the gradient ∇u of u as the unique element in $L^2(\mathbb{S}; \nu)$ such that $M_t^{[u]} = \int_0^t \nabla u(X_r) dW_r$, $t \geq 0$.

Remark 3.12. (i) Clearly, for the (reflected) Brownian motion on $Q = [0, 1]^d$, the definition of gradients given above coincides with that of weak derivatives for functions in the Sobolev space $W^{1,2}(Q)$ of $L^2(Q)$ functions with distributional derivatives in $L^2(Q)$.

(ii) As pointed out in Section 1, there exist several definitions of gradients on fractals in literature (for example, [20, 35, 37, 6, 14, 15, 17, 4]) specifically introduced to address different questions. Our definition of gradients of functions is a slight variant of those in [20, 35, 37], and equivalent to those in [14, 15, 4]. In [15] and [4], gradient operators were introduced for Dirichlet forms admitting (measure-valued) carré du champ operators $\Gamma(\cdot, \cdot)$, which send functions $u, v \in \mathcal{F}(\mathbb{S})$ to $\Gamma(u, v) = \nu_{\langle u, v \rangle}$ (cf. Definition 2.3). For the standard Dirichlet form on the gasket, the isometry $\int_{\mathbb{S}} |\nabla u|^2 d\nu = \int_{\mathbb{S}} d\Gamma(u, u)$ is valid as an immediate consequence of definitions, and the pointwise equality $\nabla u = d\Gamma(u, \psi)/d\nu$ holds ν -a.e. with a suitable choice of $\psi \in \mathcal{F}(\mathbb{S})$,¹ which is possible due to the results of [?, Lemma 3.2, Theorem 4.1, and Theorem 4.2] (see also [4, Proposition 4.2]).

¹The Radon-Nikodym derivative $(d\Gamma(u, \psi)/d\nu)(x)$ is denoted by $\langle u, \psi \rangle_{\mathcal{H}_x}$ in [15, 4].

As an immediate consequence of Definition 3.11, we have, for any $u, v \in \mathcal{F}(\mathbb{S})$, that $\mathcal{E}(u, v) = \langle \nabla u, \nabla v \rangle_\nu$ and $\nu_{\langle u, v \rangle} = \nabla u \cdot \nabla v \cdot \nu$. Let $\Phi \in C^1(\mathbb{R}^m)$ and $u_1, \dots, u_m \in \mathcal{F}(\mathbb{S})$. Then $\Phi(\mathbf{u}) \in \mathcal{F}(\mathbb{S})$ and $\nabla \Phi(\mathbf{u}) = \sum_{i=1}^m \partial_i \Phi(\mathbf{u}) \cdot \nabla u_i$ ν -a.e., where $\mathbf{u} = (u_1, \dots, u_m)$. In particular, for any $u, v \in \mathcal{F}(\mathbb{S})$, $\nabla(uv) = \nabla u \cdot v + u \cdot \nabla v$ ν -a.e.

As an application of the theory of BSDEs, we consider semi-linear parabolic equations on \mathbb{S} . On Euclidean spaces, the non-linear parabolic PDE $\partial_t u + \mathcal{L}u = f(t, x, u, \nabla u)$ is interpreted as $(\partial_t u + \mathcal{L}u)dx = f(t, x, u, \nabla u)dx$ when one considers its weak solutions. The natural analogue of these PDEs on \mathbb{S} would be $(\partial_t u + \mathcal{L}u) \cdot \mu = f(t, x, u, \nabla u) \cdot \nu$. However, for any $h \in \mathcal{F}(\mathbb{S})$, the gradient ∇h is defined only as a function in $L^2(\mathbb{S}; \nu)$ and ν is singular to μ , therefore, the above formulation does not have a proper meaning for functions u such that $u(t) \in \mathcal{F}(\mathbb{S})$ for $t \in [0, T]$. In view of this, it is more appropriate to formulate (Dirichlet) terminal-boundary value problems for semi-linear parabolic PDEs as those for measure equations.

Definition 3.13. For any $v \in L^2(\mu)$, let $\|v\|_{\mathcal{F}^{-1}} = \sup \{ \langle u, v \rangle_\mu : u \in \mathcal{F}(\mathbb{S}), \mathcal{E}_1(u) \leq 1 \}$. The space $\mathcal{F}^{-1}(\mathbb{S})$ is defined to be the $\|\cdot\|_{\mathcal{F}^{-1}}$ -completion of $L^2(\mu)$.

Definition 3.14. Let $u \in L^2(0, T; \mathcal{F}(\mathbb{S}))$; that is, $\int_0^T \mathcal{E}_1(u(t)) dt < \infty$. The $\mathcal{F}(\mathbb{S})$ -valued function u is said to have *weak derivative in $L^2(0, T; \mathcal{F}^{-1}(\mathbb{S}))$* , if there exists an $\mathcal{F}^{-1}(\mathbb{S})$ -valued function $\partial_t u$ on $[0, T]$ such that

$$\left(\int_0^T \|\partial_t u(t)\|_{\mathcal{F}^{-1}}^2 dt \right)^{1/2} < \infty \quad \text{and} \quad \int_0^T \langle u(t), \partial_t v(t) \rangle_\mu dt = - \int_0^T \langle \partial_t u(t), v(t) \rangle_\mu dt$$

for all $v \in C^1(0, T; \mathcal{F}(\mathbb{S}))$ with $v(0) = v(T) = 0$.

Remark 3.15. Clearly, if $u \in L^2(0, T; \mathcal{F}(\mathbb{S})) \cap C^1(0, T; L^2(\mu))$, then u has a weak derivative in $L^2(0, T; \mathcal{F}^{-1}(\mathbb{S}))$.

The following can be easily shown by a standard mollifier argument. (See, for example, [8, Theorem 3, Section 5.9].)

Lemma 3.16. Suppose $u \in L^2(0, T; \mathcal{F}(\mathbb{S}))$ has weak derivative $\partial_t u$ in $L^2(0, T; \mathcal{F}^{-1}(\mathbb{S}))$. Then $t \mapsto \|u(t)\|_{L^2(\mu)}^2$, $t \in [0, T]$ is absolutely continuous and $\frac{d}{dt} \|u(t)\|_{L^2(\mu)}^2 = 2 \langle \partial_t u(t), u(t) \rangle_\mu$ a.e. $t \in [0, T]$.

As an application, we consider the following semi-linear parabolic PDEs on \mathbb{S} . We should point out that there exist several formulations of PDEs on fractals, and our formulation should be regarded as an extension in this direction. (See, for example, [36, 15, 16] and references therein.)

Definition 3.17. Let $\varphi \in C^{1,0}([0, T] \times V_0)$ and $\psi \in L^2(\mu)$. A function u on $[0, T] \times \mathbb{S}$ is said to be a *weak solution* of the (Dirichlet) terminal-boundary value problem

$$\begin{cases} (\partial_t u + \mathcal{L}u) \cdot \mu = -g(t, x, u) \cdot \mu - f(t, x, u, \nabla u) \cdot \nu, & (t, x) \in [0, T] \times \mathbb{S} \setminus V_0, \\ u(t, x) = \varphi(t, x) \text{ on } [0, T] \times V_0, & u(T) = \psi, \end{cases} \quad (3.9)$$

if the following are satisfied:

- (WS.1) $u \in C([0, T] \times \mathbb{S}) \cap L^2(0, T; \mathcal{F}(\mathbb{S}))$ and u has weak derivative $\partial_t u$ in $L^2(0, T; \mathcal{F}^{-1}(\mathbb{S}))$;
- (WS.2) for any $v \in \mathcal{F}(\mathbb{S} \setminus V_0)$,

$$\frac{d}{dt} \langle u(t), v \rangle_\mu - \mathcal{E}(u(t), v) = -\langle g(t, u(t)), v \rangle_\mu - \langle f(t, u(t), \nabla u(t)), v \rangle_\nu \quad \text{a.e. } t \in [0, T]; \quad (3.10)$$

- (WS.3) $u(t, x) = \varphi(t, x)$ for each $x \in V_0$ and a.e. $t \in [0, T]$, and $\lim_{t \rightarrow T} u(t) = \psi$ in $L^2(\mu)$.

Remark 3.18. (i) We point out that the term $\langle f(t, u(t), \nabla u(t), v) \rangle_\nu$ in (3.10) is well-defined. In fact, ∇u is ν -a.e. defined and u is pointwise defined by virtue of the fact $\mathcal{F}(\mathbb{S}) \subseteq C(\mathbb{S})$, which is, as pointed out in Section 2.1, a corollary of (2.1).

(ii) The equation (3.10) is well-posed by virtue of Lemma 3.16.

(iii) In view of (WS.2) and the singularity of ν and μ , we see that if $f \neq 0$ then the PDE (3.9) does not admit a solution u such that $u \in C^{1,0}([0, T] \times \mathbb{S})$ and $u(t) \in \text{Dom}(\mathcal{L})$, $t \in [0, T]$. This suggests that the theory of PDEs on \mathbb{S} is quite different from that on \mathbb{R}^d .

To construct weak solutions of the PDE (3.9), a natural idea is to show that the solution mapping of a related linear equation is a contraction in some suitable Banach space, then iterate solutions of this linear equation. However, difficulties arise immediately due to the singularity of μ and ν . To address this difficulty, our idea is that, though calculus on fractals might be considerably different from those on \mathbb{R}^d , stochastic calculus however remains similar to its classical counterpart. Specifically, we have following Feynman-Kac representation, which gives a BSDE approach for semi-linear parabolic PDEs on \mathbb{S} .

Theorem 3.19. *Let $\varphi \in C^{1,0}([0, T] \times V_0)$, $\psi \in L^2(\mu)$. Let $g \in C([0, T] \times \mathbb{S} \times \mathbb{R})$ and $f \in C([0, T] \times \mathbb{S} \times \mathbb{R}^2)$. If the PDE (3.9) admits a solution u , then, for each $s \in [0, T]$ and each $x \in \mathbb{S}$,*

$$(Y_t^{(s)}, Z_t^{(s)}) = (u(t + s, X_t), \nabla u(t + s, X_t))$$

is the unique solution (in the sense of Theorem 3.5.(a)) of the BSDE

$$\begin{cases} dY_t^{(s)} = -g(t + s, X_t, Y_t^{(s)})dt \\ \quad - f(t + s, X_t, Y_t^{(s)}, Z_t^{(s)})d\langle W \rangle_t + Z_t^{(s)}dW_t, \quad t \in [0, \sigma^{(s)}), \\ Y_{\sigma^{(s)}}^{(s)} = \Psi(\sigma^{(s)}, X_{\sigma^{(s)}}), \end{cases} \quad (3.11)$$

on $(\Omega, \{\mathcal{F}_t\}, \mathbb{P}_x)$ for each $x \in \mathbb{S}$, where $\sigma^{(s)} = (T - s) \wedge \sigma_{V_0}$, $s \in [0, T]$, and

$$\Psi(t, x) = \begin{cases} \varphi(t, x), & \text{if } (t, x) \in [0, T] \times V_0, \\ \psi(x), & \text{if } (t, x) \in \{T\} \times \mathbb{S} \setminus V_0. \end{cases}$$

Moreover, the solution of (3.9) is unique, and has the following representation

$$u(t, x) = \mathbb{E}_x(Y_0^{(t)}) \quad \text{for all } (t, x) \in [0, T] \times \mathbb{S}. \quad (3.12)$$

4 Several results on Brownian martingale

In this section, we collect several results, which will be needed in later sections, on Brownian martingale. In particular, a representation theorem for square-integrable martingale and a time-dependent Itô-Fukushima decomposition are given. Moreover, we prove the exponential integrability of the quadratic process of Brownian martingale, which is the main technical result in this section.

4.1 Martingale representations

In this subsection, we consider representations for square-integrable martingales adapted to filtrations induced by Brownian motion. The following was proved in [31, Theorem 3] for general continuous Hunt processes.

Theorem 4.1. *Let $(\Omega, \{X_t\}, \{\mathbb{P}_x\})$ be a continuous Hunt process with state space S , and $\lambda \in \mathcal{P}(S)$. Let \mathcal{F}_t^λ be the \mathbb{P}_λ -completion of $\sigma(X_r : r \leq t)$. Then any local martingales on $(\Omega, \{\mathcal{F}_t^\lambda\}, \mathbb{P}_\lambda)$ are continuous. Furthermore, suppose that there exist a sub-algebra \mathcal{A} of $L^2(S; \lambda) \cap \mathcal{B}_b(S)$ and finitely many continuous martingales W^1, \dots, W^d such that the following are satisfied:*

(i) $\sigma(\mathcal{A}) = \mathcal{B}(S)$ and $R_\alpha(\mathcal{A}) \subseteq \mathcal{A}$ for each $\alpha > 0$, where R_α denotes the α -resolvent of $\{X_t\}$;
(ii) for any $f \in \mathcal{A}$ and any $\alpha > 0$, there exist $\{\mathcal{F}_t^\lambda\}$ -predictable processes f^1, \dots, f^d such that $M_t^{\alpha, f} = \sum_{j=1}^d \int_0^t f^j dW_r^j$, $t \geq 0$ \mathbb{P}_λ -a.s., where $M^{\alpha, f}$ is the martingale part of $R_\alpha u(X_t) - R_\alpha u(X_0)$;

(iii) the matrix $(\langle W^i, W^j \rangle_t)_{i,j}$ is strictly positive definite for all $t \geq 0$ \mathbb{P}_λ -a.s.

Then, for any square-integrable martingale M on $(\Omega, \{\mathcal{F}_t^\lambda\}, \mathbb{P}_\lambda)$, there uniquely exist $\{\mathcal{F}_t^\lambda\}$ -predictable processes f^1, \dots, f^d such that $M_t = M_0 + \sum_{j=1}^d \int_0^t f^j(r) dW_r^j$, $t \geq 0$ \mathbb{P}_λ -a.s.

We now return to the setting of Brownian motion on the Sierpinski gasket. Since $\mathcal{F}(\mathbb{S}) \subseteq C(\mathbb{S})$, we that $\mathcal{F}(\mathbb{S})$ is an algebra. Let $\lambda \in \mathcal{P}(\mathbb{S})$. The assumptions (i) and (iii) in Theorem 4.1 are clearly satisfied with $\mathcal{A} = \mathcal{F}(\mathbb{S})$. The assumption (ii) is also satisfied in view of Theorem 2.6. Therefore, we have the following representation theorem for square-integrable martingales adapted to filtrations induced by the Brownian motion.

Theorem 4.2. *Let $\lambda \in \mathcal{P}(\mathbb{S})$. For any square-integrable martingale M on $(\Omega, \{\mathcal{F}_t^\lambda\}, \mathbb{P}_\lambda)$, there exists an $\{\mathcal{F}_t^\lambda\}$ -predictable process f such that*

$$M_t = M_0 + \int_0^t f(r) dW_r, \quad t \geq 0 \quad \mathbb{P}_\lambda\text{-a.s.}$$

The process f is unique; that is, if \bar{f} is another $\{\mathcal{F}_t^\lambda\}$ -predictable process satisfying the above, then $\mathbb{E}_\lambda[\int_0^\infty (f(r) - \bar{f}(r))^2 d\langle W \rangle_r] = 0$.

We end this subsection by showing the uniqueness of decompositions of semi-martingales Y_t of the form

$$Y_t = Y_0 + \int_0^t g(r) dr + \int_0^t f(r) d\langle W \rangle_r + M_t, \quad t \geq 0, \quad (4.1)$$

where M is a martingale on $(\Omega, \{\mathcal{F}_t^\mu\}, \mathbb{P}_\mu)$.

Lemma 4.3. *The Lebesgue-Stieltjes measure $d\langle W \rangle_t$ is singular to dt \mathbb{P}_μ -a.s.*

Proof. Let \mathcal{P} be the σ -field of predictable sets in $[0, \infty) \times \Omega$. Let Q be the unique measure on $([0, \infty) \times \Omega, \mathcal{P})$ satisfying $Q(\llbracket \sigma, \tau \rrbracket) = \mathbb{E}_\mu(\langle W \rangle_\tau - \langle W \rangle_\sigma)$, $\sigma, \tau \in \mathcal{T}_p$, where \mathcal{T}_p is the family of all $\{\mathcal{F}_t^\mu\}$ -predictable times, and $\llbracket \sigma, \tau \rrbracket = \{(t, \omega) \in [0, \infty) \times \Omega : \sigma(\omega) \leq t < \tau(\omega)\}$. By the Lebesgue decomposition, $Q = f \cdot (dt \times \mathbb{P}_\mu) + Q_s$, where $f \geq 0$ is a predictable process and is σ -integrable with respect to $dt \times \mathbb{P}_\mu$, and Q_s is a σ -finite positive measure singular to $dt \times \mathbb{P}_\mu$.

Note that, for any non-negative predictable process g ,

$$\int_{[0, T) \times \Omega} g(r) Q(dr, d\omega) = \mathbb{E}_\mu\left(\int_0^T g(r) d\langle W \rangle_r\right), \quad (4.2)$$

which can be easily shown by the definition of Q and a standard monotone class argument. Now, let $B \in \mathcal{B}(\mathbb{S})$ such that $\mu(B) = 1 = \nu(\mathbb{S} \setminus B) = 1$. By $\mathbb{E}_\mu(\int_0^T 1_{\mathbb{S} \setminus B}(X_r) dr) = T\mu(\mathbb{S} \setminus B) = 0$, we see that $1_B(X_t) = 1$, a.e. $t \geq 0$, \mathbb{P}_μ -a.s. Therefore, by (4.2),

$$\begin{aligned} \int_0^T \mathbb{E}_\mu(f(t)) dt &= \int_0^T \mathbb{E}_\mu(f(t) 1_B(X_r)) dt \leq \int_{[0, T) \times \Omega} 1_B(X_t(\omega)) Q(dt, d\omega) \\ &= \mathbb{E}_\mu\left(\int_0^T 1_B(X_t) d\langle W \rangle_t\right) = T\nu(B) = 0, \end{aligned}$$

which implies the conclusion of the lemma. \square

Corollary 4.4. *Let Y be a semi-martingale on $(\Omega, \{\mathcal{F}_t^\mu\}, \mathbb{P}_\mu)$ of the form (4.1). Then the decomposition (4.1) is unique; that is, if (4.1) also holds with g, f, M replaced by $\bar{g}, \bar{f}, \bar{M}$, then $\mathbb{E}_\mu(\int_0^\infty |g(r) - \bar{g}(r)| dr) = \mathbb{E}_\mu(\int_0^\infty |f(r) - \bar{f}(r)| d\langle W \rangle_r) = 0$.*

We shall need the following time-dependent Itô-Fukushima decomposition (see also [28, 9] for similar results in different settings), which follows from Theorem 4.2 and the decomposition theorem in [38, Theorem 4.5] applied to the (non-symmetric) *generalized Dirichlet form* $(u, v) \mapsto \Lambda(u, v) + \int_0^T \mathcal{E}(u(t), v(t)) dt$, where

$$\Lambda(u, v) = \begin{cases} \int_0^T \langle u(t), \partial_t v(t) \rangle_\mu dt, & \text{if } u \in L^2(0, T; \mathcal{F}(\mathbb{S})), v \in L^2(0, T; \mathcal{F}(\mathbb{S})) \cap C^1(0, T; L^2(\mu)), \\ - \int_0^T \langle \partial_t u(t), v(t) \rangle_\mu dt, & \text{if } u \in L^2(0, T; \mathcal{F}(\mathbb{S})) \cap C^1(0, T; L^2(\mu)), v \in L^2(0, T; \mathcal{F}(\mathbb{S})). \end{cases}$$

(See, for example, [33, 38] and etc., for the theory of generalized Dirichlet forms.)

Lemma 4.5. *Suppose $u \in C([0, T] \times \mathbb{S}) \cap L^2(0, T; \mathcal{F}(\mathbb{S}))$ and that u has weak derivative $\partial_t u$ in $L^2(0, T; \mathcal{F}^{-1}(\mathbb{S}))$. Then*

$$u(t, X_t) = u(0, X_0) + \int_0^t \nabla u(r, X_r) dW_r + N_t, \quad t \in [0, T],$$

where N_t is a continuous processes with zero quadratic variation; that is, for each $t \geq 0$,

$$\lim_{n \rightarrow \infty} \mathbb{E}_\mu \left[\sum_{i=1}^n (N_{t_i} - N_{t_{i-1}})^2 \right] = 0,$$

where $t_i = \frac{i}{n}t$, $i = 0, 1, \dots, n$.

4.2 Exponential integrability of quadratic processes

We now turn to the main technical result of this section, the exponential integrability of the quadratic process $\langle W \rangle$, which is a sufficient condition for the Girsanov theorem to hold and is crucial for the existence of solutions of BSDEs. We shall need the following heat kernel estimates. (See [22, Theorem 5.3.1].)

Lemma 4.6. (a) $\{X_t\}$ admits (jointly) continuous transition kernels $p_t(x, y)$, $t > 0$, and there exists universal constants $C_{*,1}, C_{*,2} > 0$ such that

$$C_{*,1} t^{-d_s/2} \leq p_t(x, y) \leq C_{*,2} t^{-d_s/2}, \quad t \in (0, 1], \quad x, y \in \mathbb{S},$$

where $d_s = 2 \log 3 / \log 5$ is the spectral dimension of $\{X_t\}$.

(b) $\{X_t^0\}$ admits (jointly) continuous transition kernels $p_t^0(x, y)$, $t > 0$, and there are universal constants $C_{*,3}, C_{*,4} > 0$ such that

$$C_{*,3} t^{-d_s/2} \leq p_t^0(x, y) \leq C_{*,4} t^{-d_s/2}, \quad t \in (0, 1], \quad x, y \in \mathbb{S} \setminus V_0.$$

In general, $P_t f$ and $P_t^0 f$ are well-defined only for $f \in L^2(\mu)$. However, using the corresponding transition densities and their continuities, these definitions can be extended to measures (even possibly singular to μ) as follows.

Definition 4.7. For each Radon measure λ on \mathbb{S} and each $t > 0$, define

$$P_t \lambda(x) \triangleq \int_{\mathbb{S}} p_t(x, y) \lambda(dy), \quad x \in \mathbb{S}. \quad (4.3)$$

Remark 4.8. Clearly, $P_t \lambda \in C(\mathbb{S})$ and $|P_t \lambda| \leq |\lambda|(\mathbb{S}) < \infty$ for any Radon measure λ on \mathbb{S} .

Lemma 4.9. Let A be a positive continuous additive functional such that $\nu_A(\mathbb{S}) < \infty$, where ν_A is the Revuz measure of A . Then, for each $t > 0$ and each $f \in \mathcal{B}_b([0, \infty) \times \mathbb{S})$,

$$\mathbb{E}_x \left(\int_0^t f(r, X_r) dA_r \right) = \int_0^t P_r(f(r) \nu_A)(x) dr, \quad x \in \mathbb{S}. \quad (4.4)$$

Similarly, for positive continuous additive functional A with respect to $\{X_t^0\}$,

$$\mathbb{E}_x \left(\int_0^t f(r, X_r^0) dA_r \right) = \int_0^t P_r^0(f(r) \nu_A)(x) dr, \quad x \in \mathbb{S} \setminus V_0. \quad (4.5)$$

Proof. By and the definition of $P_r(g\nu_A)$ and a monotone class argument, $\mathbb{E}_\mu((\int_0^t f(r, X_r) dA_r)h(X_0)) = \int_0^t \langle P_r(f\nu_A), h \rangle_\mu dr$ for all $h \in \mathcal{B}_b(\mathbb{S})$, which implies (4.4) in view of the continuity of both sides of (4.4). The other equality can be proved similarly. \square

Lemma 4.10. Let $A^{(i)}$, $i = 1, \dots, n$ be positive continuous additive functionals such that $\nu_i(\mathbb{S}) < \infty$, where ν_i is the Revuz measure of $A^{(i)}$, $i = 1, \dots, n$. Then, for each $t > 0$ and each $f_i \in \mathcal{B}_b([0, \infty) \times \mathbb{S})$, $i = 1, \dots, n$,

$$\begin{aligned} & \mathbb{E}_x \left(f(X_t) \int_{0 < t_1 < \dots < t_n < t} f_1(t_1, X_{t_1}) \cdots f_n(t_n, X_{t_n}) dA_{t_1}^{(1)} \cdots dA_{t_n}^{(n)} \right) \\ &= \int_{0 < t_1 < \dots < t_n < t} P_{t_1}(\nu_1 f_1(t_1) P_{t_2-t_1}(\cdots \nu_n f_n(t_n) P_{t-t_n} f) \cdots)(x) dt_1 \cdots dt_{n-1} dt_n, \quad x \in \mathbb{S}. \end{aligned} \quad (4.6)$$

Proof. By (4.4),

$$\mathbb{E}_x \left(f(X_t) \int_0^t f_1(t_1, X_{t_1}) dA_{t_1}^{(1)} \right) = \mathbb{E}_x \left(\int_0^t f_1(t_1, X_{t_1}) \mathbb{E}_x(f(X_t) | \mathcal{F}_{t_1}) dA_{t_1}^{(1)} \right) = \int_0^t P_{t_1}(\nu_1 f_1(t_1) P_{t-t_1} f)(x) dt_1.$$

This proves (4.6) for $n = 1$. The conclusion for a general n follows readily from an induction argument. \square

Lemma 4.11. Let A be a positive continuous additive functional such that $\nu_A(\mathbb{S}) < \infty$, where ν_A is the Revuz measure of A .

(a) For each $\beta > 0$,

$$\sup_{x \in \mathbb{S}} \mathbb{E}_x(e^{\beta A_t}) \leq E_{\gamma_s, 1}[C_* \nu_A(\mathbb{S}) \beta \max\{t, t^{\gamma_s}\}], \quad t \geq 0. \quad (4.7)$$

where $C_* > 0$ denotes any (possibly different) universal constant, $\gamma_s = 1 - d_s/2 \in (0, 1/2)$, $E_{a,b}(z) = \sum_{p=0}^{\infty} \frac{z^p}{\Gamma(ap+b)}$, $z \in \mathbb{C}$, $a, b > 0$ is the Mittag-Leffler function, and Γ is the Gamma function.

(b) For each $f \in L_+^1(\mu)$ and $\beta > 0$,

$$\sup_{x \in \mathbb{S}} \mathbb{E}_x(f(X_t) e^{\beta A_t}) \leq \max\{1, t^{-d_s/2}\} \|f\|_{L^1(\mu)} E_{\gamma_s, \gamma_s}[C_* \nu_A(\mathbb{S}) \beta \max\{t, t^{\gamma_s}\}], \quad t > 0.$$

Remark 4.12. To see the asymptotics of $\mathbb{E}_x(e^{\beta A_t})$ as $t \rightarrow 0$, we note that the Mittag-Leffler function $E_{a,b}$ is an entire function of order $1/a$, that is,

$$\frac{1}{a} = \limsup_{r \rightarrow \infty} \frac{\log \log(\sup_{|z| \leq r} |E_{a,b}(z)|)}{\log r}.$$

Proof. (a) Suppose first that $t \in (0, 1]$. For each $p \in \mathbb{N}_+$, by Lemma 4.10,

$$\begin{aligned}\mathbb{E}_x(A_t^p) &= p! \mathbb{E}_x \left(\int_{0 < t_1 < \dots < t_p < t} dA_{t_1} \cdots dA_{t_p} \right) \\ &= p! \int_{0 < t_1 < \dots < t_p < t} P_{t_1}(\nu_A P_{t_2-t_1}(\cdots \nu_A P_{t_p-t_{p-1}}(\mu_A) \cdots))(x) dt_1 \cdots dt_p.\end{aligned}$$

For each non-negative $f \in C(\mathbb{S})$, by Lemma 4.6,

$$P_r(\nu_A f)(x) \leq \|f\|_{L^\infty} P_r \nu_A(x) \leq C_* \nu_A(\mathbb{S}) r^{-d_s/2}, \quad r > 0, x \in \mathbb{S} \quad (4.8)$$

so that

$$\|P_{t_1}(\nu_A P_{t_2-t_1}(\cdots \nu_A P_{t_p-t_{p-1}}(\nu_A) \cdots))\|_{L^\infty} \leq (C_* \nu_A(\mathbb{S}))^p [t_1(t_2-t_1) \cdots (t_p-t_{p-1})]^{-d_s/2},$$

which implies that

$$\begin{aligned}\frac{1}{p!} \mathbb{E}_x(\langle W \rangle_t^p) &\leq (C_* \nu_A(\mathbb{S}))^p \int_{0 < t_1 < \dots < t_p < t} [t_1(t_2-t_1) \cdots (t_p-t_{p-1})]^{-d_s/2} dt_1 \cdots dt_p \\ &= (C_* \nu_A(\mathbb{S}) t^{\gamma_s})^p \int_{0 < \theta_1 < \dots < \theta_p < 1} [\theta_1(\theta_2-\theta_1) \cdots (\theta_p-\theta_{p-1})]^{-d_s/2} d\theta_1 \cdots d\theta_p.\end{aligned} \quad (4.9)$$

Let

$$\beta_0(\gamma_1, \dots, \gamma_p) = \int_{0 < \theta_1 < \dots < \theta_p < 1} \theta_1^{\gamma_1-1} (\theta_2-\theta_1)^{\gamma_2-1} \cdots (\theta_p-\theta_{p-1})^{\gamma_p-1} d\theta_1 \cdots d\theta_p.$$

Then

$$\begin{aligned}&\beta_0(\gamma_1, \dots, \gamma_p) \\ &= \int_{0 < \theta_2 < \dots < \theta_p < 1} \left(\int_0^{\theta_2} \theta_1^{\gamma_1-1} (\theta_2-\theta_1)^{\gamma_2-1} d\theta_1 \right) (\theta_3-\theta_2)^{\gamma_3-1} \cdots (\theta_p-\theta_{p-1})^{\gamma_p-1} d\theta_2 \cdots d\theta_p \\ &= \int_{0 < \theta_2 < \dots < \theta_p < 1} \left(\int_0^1 \theta^{\gamma_1-1} (1-\theta)^{\gamma_2-1} d\theta \right) \theta_2^{\gamma_1+\gamma_2-1} (\theta_3-\theta_2)^{\gamma_3-1} \cdots (\theta_p-\theta_{p-1})^{\gamma_p-1} d\theta_2 \cdots d\theta_p \\ &= B(\gamma_1, \gamma_2) \beta_0(\gamma_1 + \gamma_2, \gamma_3, \dots, \gamma_p),\end{aligned}$$

where $B(\cdot, \cdot)$ is the Beta function. By induction, we see that

$$\begin{aligned}\beta_0(\gamma_1, \dots, \gamma_p) &= B(\gamma_1, \gamma_2) B(\gamma_1 + \gamma_2, \gamma_3) \cdots B(\gamma_1 + \cdots + \gamma_{p-1}, \gamma_p) \int_0^1 \theta^{\gamma_1 + \cdots + \gamma_{p-1}} d\theta \\ &= \frac{B(\gamma_1, \gamma_2) B(\gamma_1 + \gamma_2, \gamma_3) \cdots B(\gamma_1 + \cdots + \gamma_{p-1}, \gamma_p)}{\gamma_1 + \cdots + \gamma_p}.\end{aligned}$$

Therefore, by (4.9),

$$\begin{aligned}\frac{1}{p!} \mathbb{E}_x(A_t^p) &\leq (C_* \nu_A(\mathbb{S}) t^{\gamma_s})^p \beta(\gamma_s, \dots, \gamma_s) \\ &= \frac{(C_* \nu_A(\mathbb{S}) t^{\gamma_s})^p}{p \gamma_s} \prod_{i=1}^{p-1} B(i \gamma_s, \gamma_s) = \frac{(C_* \nu_A(\mathbb{S}) t^{\gamma_s})^p}{\Gamma(p \gamma_s + 1)},\end{aligned}$$

for each $p \in \mathbb{N}$, $t \in (0, 1]$, which implies (4.7) for $t \in [0, 1]$.

For $t \in (1, \infty)$, we claim that

$$\sup_{x \in \mathbb{S}} \mathbb{E}_x(e^{\beta(A_{k+1}-A_k)}) \leq \mathbb{E}_{\gamma_s,1}(C_* \nu_A(\mathbb{S})\beta), \quad \beta > 0, k \in \mathbb{N}. \quad (4.10)$$

In fact, for any $f \in L^1(\mu)$,

$$\begin{aligned} \left| \int_{\mathbb{S}} f(x) \mathbb{E}_x(e^{\beta(A_{k+1}-A_k)}) \mu(dx) \right| &= \left| \mathbb{E}_\mu(f(X_0) e^{\beta(A_{k+1}-A_k)}) \right| \\ &= \left| \mathbb{E}_\mu[f(X_0) \mathbb{E}_{X_k}(e^{\beta A_1})] \right| \leq \mathbb{E}_{\gamma_s,1}[C_* \nu_A(\mathbb{S})\beta] \|f\|_{L^1(\mu)}, \end{aligned}$$

which implies that $\mathbb{E}_x(e^{\beta(A_{k+1}-A_k)}) \leq \mathbb{E}_{\gamma_s,1}(C_* \Gamma(\gamma_s) \nu_A(\mathbb{S})\beta)$, μ -a.e. $x \in \mathbb{S}$. In particular,

$$\sum_{p=0}^N \frac{1}{p!} \mathbb{E}_x((A_{k+1} - A_k)^p) \leq \mathbb{E}_{\gamma_s,1}[C_* \nu_A(\mathbb{S})\beta], \quad N \in \mathbb{N}, \mu\text{-a.e. } x \in \mathbb{S}.$$

By Lemma 4.10, for each $N \in \mathbb{N}$, the function $x \mapsto \sum_{p=0}^N \frac{1}{p!} \mathbb{E}_x((A_{k+1} - A_k)^p)$ is continuous. Therefore, the above holds for each $N \in \mathbb{N}$ and each $x \in \mathbb{S}$. Letting $N \rightarrow \infty$ proves (4.10).

For $n \in \mathbb{N}_+$ such that $n < t \leq n+1 \leq 2t$, by Hölder's inequality and (4.10), we have

$$\mathbb{E}_x(e^{\beta A_t}) \leq \prod_{i=0}^n [\mathbb{E}_x(e^{(n+1)\beta(A_{i+1}-A_i)})]^{1/(n+1)} \leq \mathbb{E}_{\gamma_s,1}[C_* \nu_A(\mathbb{S})\beta t].$$

which proves (a).

(b) Suppose $t \in (0, 1]$. By Lemma 4.10,

$$\mathbb{E}_x(f(X_t) A_t^p) = p! \int_{0 < t_1 < \dots < t_p < t} P_{t_1}(\nu_A P_{t_2-t_1}(\dots \nu_A P_{t_p-t_{p-1}}(\nu_A P_{t-t_p} f) \dots))(x) dt_1 \dots dt_p.$$

According to (4.8),

$$\|P_{t_1}(\nu_A P_{t_2-t_1}(\dots \nu_A P_{t_p-t_{p-1}}(\nu_A P_{t-t_p} f) \dots))\|_{L^\infty} \leq \|f\|_{L^1(\mu)} (C_* \nu_A(\mathbb{S}))^p [t_1(t_2-t_1) \dots (t-t_p)]^{-d_s/2}$$

so that

$$\begin{aligned} \frac{1}{p!} \mathbb{E}_x(f(X_t) \langle W \rangle_t^p) &\leq (C_* \nu_A(\mathbb{S}))^p \int_{0 < t_1 < \dots < t_p < t} [t_1(t_2-t_1) \dots (t-t_p)]^{-d_s/2} dt_1 \dots dt_p \\ &= t^{-d_s/2} \|f\|_{L^1(\mu)} (C_* \nu_A(\mathbb{S}) t^{\gamma_s})^p \int_{0 < \theta_1 < \dots < \theta_p < 1} [\theta_1(\theta_2-\theta_1) \dots (1-\theta_p)]^{-d_s/2} d\theta_1 \dots d\theta_p. \end{aligned}$$

The proof now proceeds similarly to that of (a). □

Corollary 4.13. For each $f \in L^1_+(\mu)$ and $\beta, t > 0$,

$$\sup_{x \in \mathbb{S}} \mathbb{E}_x(f(X_t) e^{\beta \langle W \rangle_t}) \leq \max\{1, t^{-d_s/2}\} \|f\|_{L^1(\mu)} \mathbb{E}_{\gamma_s, \gamma_s}[C_* \beta \max\{t, t^{\gamma_s}\}],$$

where $C_* > 0$ is a universal constant.

Corollary 4.14. (a) For any $\beta \in \mathbb{R}$, $Z_t = e^{\beta W_t - \frac{1}{2} \beta^2 \langle W \rangle_t}$, $t \geq 0$ is a martingale on $(\Omega, \{\mathcal{F}_t\}, \mathbb{P}_x)$ for each $x \in \mathbb{S}$.

(b) Let $x \in \mathbb{S}$, $T > 0$. Let $\{M_t\}_{t \geq 0}$ be a continuous local \mathbb{P}_x -martingale. Then $\tilde{M}_t = M_t - \langle M, W \rangle_t$, $0 \leq t \leq T$ is a continuous local \mathbb{Q}_x -martingale, where $\mathbb{Q}_x = Z_T \mathbb{P}_x$.

We now prove the exponential integrability of $\langle W \rangle$ up to the hitting time σ_{V_0} .

Proposition 4.15. *There exists a $\beta_0 > 0$ such that $\sup_{x \in \mathbb{S}} \mathbb{E}_x(e^{\beta_0 \langle W \rangle_{\sigma_{V_0}}}) < \infty$.*

To prove Proposition 4.15, we need the following two lemmas.

Lemma 4.16. *Let h_i be the harmonic functions with boundary values $h_i|_{V_0} = 1_{\{p_i\}}$, $i = 1, 2, 3$ respectively. Then*

$$\langle W \rangle = \frac{1}{3}(\langle M^{[h_1]} \rangle + \langle M^{[h_2]} \rangle + \langle M^{[h_3]} \rangle), \quad (4.11)$$

where $M^{[h_i]}$ are the martingale parts of $h_i(X_t) - h_i(X_0)$, $i = 1, 2, 3$ respectively.

Proof. Let $\mathbf{e}_i = 1_{\{p_i\}}$, $i = 1, 2, 3$. For any $\omega \in W_*$, since $h_i \circ F_{[\omega]_m}$ is the harmonic function with boundary value $\mathbf{A}_{[\omega]_m} \mathbf{e}_i$, it follows from (2.3) that

$$\mathcal{E}(h_i \circ F_{[\omega]_m}) = \frac{3}{2} \mathbf{e}_i^t \mathbf{A}_{[\omega]_m}^t \mathbf{P} \mathbf{A}_{[\omega]_m} \mathbf{e}_i = \frac{3}{2} \mathbf{e}_i^t \mathbf{Y}_{[\omega]_m}^t \mathbf{Y}_{[\omega]_m} \mathbf{e}_i, \quad i = 1, 2, 3.$$

Therefore, by the self-similar property (2.2) of \mathcal{E} , we deduce that

$$\nu_{\langle h_i \rangle}(F_{[\omega]_m}(\mathbb{S})) = \left(\frac{5}{3}\right)^m \mathcal{E}(h_i \circ F_{[\omega]_m}) = \frac{3}{2} \left(\frac{5}{3}\right)^m \mathbf{e}_i^t \mathbf{Y}_{[\omega]_m}^t \mathbf{Y}_{[\omega]_m} \mathbf{e}_i, \quad i = 1, 2, 3.$$

In particular,

$$\sum_{i=1,2,3} \nu_{\langle h_i \rangle}(F_{[\omega]_m}(\mathbb{S})) = \frac{3}{2} \left(\frac{5}{3}\right)^m \text{tr}(\mathbf{Y}_{[\omega]_m}^t \mathbf{Y}_{[\omega]_m}) = 3\nu(F_{[\omega]_m}(\mathbb{S})).$$

Therefore, $\nu = \frac{1}{3} \sum_{i=1,2,3} \nu_{\langle h_i \rangle}$, from which (4.11) follows readily in view of the one-to-one correspondence between Revuz measures and positive additive functionals. \square

Lemma 4.17. *Let $h \in \mathcal{F}(\mathbb{S})$ be a harmonic function. Then, for each $x \in \mathbb{S}$, $M_t = h(X_{t \wedge \sigma_{V_0}})$, $t \geq 0$ is a BMO martingale on $(\Omega, \{\mathcal{F}_t\}, \mathbb{P}_x)$ and $\|M\|_{\text{BMO}} \leq \max_{V_0} |h|$.*

Proof. We first show that $\{M_t\}$ is a martingale. Since h is a harmonic function, we have $h(x) = \mathbb{E}_x(h(X_{\sigma_{V_0}}))$, $x \in \mathbb{S}$. Therefore

$$\begin{aligned} M_t &= \mathbb{E}_{X_t}(h(X_{\sigma_{V_0}}))1_{\{t < \sigma_{V_0}\}} + h(X_{\sigma_{V_0}})1_{\{t \geq \sigma_{V_0}\}} \\ &= \mathbb{E}_x(h(X_{t+\sigma_{V_0} \circ \theta_t})1_{\{t < \sigma_{V_0}\}} | \mathcal{F}_t) + h(X_{\sigma_{V_0}})1_{\{t \geq \sigma_{V_0}\}} \\ &= \mathbb{E}_x(h(X_{\sigma_{V_0}})1_{\{t < \sigma_{V_0}\}} | \mathcal{F}_t) + h(X_{\sigma_{V_0}})1_{\{t \geq \sigma_{V_0}\}}. \end{aligned}$$

Note that $h(X_{\sigma_{V_0}})1_{\{t \geq \sigma_{V_0}\}} \in \mathcal{F}_t$. We see that $M_t = \mathbb{E}_x(h(X_{\sigma_{V_0}}) | \mathcal{F}_t)$, $t \geq 0$. This implies that $\{M_t\}$ is a martingale on $(\Omega, \{\mathcal{F}_t\}, \mathbb{P}_x)$.

By the maximum principle for harmonic functions (see [22, Theorem 3.2.5]), $|M_t| \leq \max_{V_0} |h|$. Note that $M_t = M_{t \wedge \sigma_{V_0}}$, $t \geq 0$. We see that

$$\mathbb{E}_x(\langle M \rangle_\infty - \langle M \rangle_\tau) = \mathbb{E}_x(M_{\sigma_{V_0}}^2 - M_{\tau \wedge \sigma_{V_0}}^2) \leq \max_{V_0} |h|^2,$$

which implies $\|M\|_{\text{BMO}} \leq \max_{V_0} |h|$. \square

Proof of Proposition 4.15. The conclusion follows immediately from Lemma 4.17, Lemma 4.16 and the John-Nirenberg inequality for BMO martingales. \square

5 BSDEs driven by Brownian martingale

5.1 BSDEs with deterministic durations

In this subsection, we prove the existence and uniqueness of solutions of the BSDE (3.1). Results and proofs in this subsection are also valid if T is replaced by a bounded $\{\mathcal{F}_t^\lambda\}$ -stopping time τ .

We start with the simple case where g, f do not depend on y or z ; that is,

$$\begin{cases} dY_t = -g(t)dt - f(t)d\langle W \rangle_t + Z_t dW_t, & t \in [0, T], \\ Y_T = \xi, \end{cases} \quad (5.1)$$

where $g(t) = g(t, \omega)$ and $f(t) = f(t, \omega)$ are $\{\mathcal{F}_t^\lambda\}$ -adapted processes.

Lemma 5.1. *Let $\beta = (\beta_0, \beta_1) \in [1, \infty)^2$, and let $\xi \in \mathcal{F}_T^\lambda$ satisfy (A.1). Suppose that $g(t), f(t)$ are $\{\mathcal{F}_t^\lambda\}$ -adapted processes such that*

$$\mathbb{E}_\lambda \left(\int_0^T g(r)^2 e^{2\beta_1 \langle W \rangle_r} dr \right) < \infty, \quad \mathbb{E}_\lambda \left(\int_0^T f(r)^2 e^{2\beta_1 \langle W \rangle_r} d\langle M^1 \rangle_r \right) < \infty.$$

Then the BSDE (5.1) admits a unique solution (Y, Z) in $\mathcal{V}_\lambda^\beta[0, T]$. Moreover,

$$\begin{aligned} \|(Y, Z)\|_{\mathcal{V}_\lambda^\beta[0, T]}^2 &\leq 10 \left[\mathbb{E}_\lambda \left(\xi^2 e^{2\beta_0 T + 2\beta_1 \langle W \rangle_T} \right) + \mathbb{E}_\lambda \left(\int_0^T g(r)^2 e^{2\beta_0 r + 2\beta_1 \langle W \rangle_r} dr \right) \right. \\ &\quad \left. + \mathbb{E}_\lambda \left(\int_0^T f(r)^2 e^{2\beta_0 r + 2\beta_1 \langle W \rangle_r} d\langle W \rangle_r \right) \right]. \end{aligned} \quad (5.2)$$

Proof. Let

$$Y_t = \mathbb{E}_\lambda \left(\xi + \int_t^T g(r) dr + \int_t^T f(r) d\langle W \rangle_r \mid \mathcal{F}_t \right), \quad t \in [0, T].$$

Then $Y_T = \xi$, and

$$Y_t + \int_0^t g(r) dr + \int_0^t f(r) d\langle W \rangle_r = \mathbb{E}_\lambda \left(\xi + \int_0^T g(r) dr + \int_0^T f(r) d\langle W \rangle_r \mid \mathcal{F}_t \right)$$

is a martingale on $(\Omega, \{\mathcal{F}_t^\lambda\}, \mathbb{P}_\lambda)$. By Theorem 4.2, there exists a unique $\{\mathcal{F}_t^\lambda\}$ -predictable process Z such that

$$Y_t - Y_0 + \int_0^t g(r) dr + \int_0^t f(r) d\langle W \rangle_r = \int_0^t Z_r dW_r, \quad t \in [0, T] \quad \mathbb{P}_\lambda\text{-a.s.}$$

which, together with $Y_T = \xi$, implies that

$$Y_t = \xi + \int_t^T g(r) dr + \int_t^T f(r) d\langle W \rangle_r - \int_t^T Z_r dW_r, \quad t \in [0, T] \quad \mathbb{P}_\lambda\text{-a.s.}$$

Therefore, (Y, Z) is a solution of the BSDE (5.1).

We now turn to the proof of (5.2), from which the uniqueness of the solution in $\mathcal{V}_\lambda^\beta[0, T]$ follows immediately. Denote $e_t = \exp(2\beta_0 t + 2\beta_1 \langle W \rangle_t)$. By Itô's formula,

$$\begin{aligned} Y_t^2 e_t &= \xi^2 e_T + \int_t^T (-2\beta_0 Y_r^2 + 2Y_r g(r)) e_r dr \\ &\quad + \int_t^T (-2\beta_1 Y_r^2 + 2Y_r f(r) - Z_r^2) e_r d\langle W \rangle_r - 2 \int_t^T Y_r Z_r e_r dW_r, \end{aligned}$$

which implies that

$$\begin{aligned} & Y_t^2 e_t + 2\beta_0 \int_t^T Y_r^2 e_r dr + \int_t^T (2\beta_1 Y_r^2 + Z_r^2) e_r d\langle W \rangle_r \\ & \leq \xi^2 e_T + \int_t^T \left(\beta_0 Y_r^2 + \frac{1}{\beta_0} g(r)^2 \right) e_r dr + \int_t^T \left(\beta_1 Y_r^2 + \frac{1}{\beta_1} f(r)^2 \right) e_r d\langle W \rangle_r - 2 \int_t^T Y_r Z_r e_r dW_r. \end{aligned} \quad (5.3)$$

Taking expectations on both sides of the above and using a localization argument give

$$\mathbb{E}_\lambda \left(\int_0^T Z_r^2 e_r d\langle W \rangle_r \right) \leq \mathbb{E}_\lambda \left(\xi^2 e_T + \frac{1}{\beta_0} \int_0^T g(r)^2 e_r dr + \frac{1}{\beta_1} \int_0^T f(r)^2 e_r d\langle W \rangle_r \right). \quad (5.4)$$

By (5.3) again,

$$\begin{aligned} & Y_t^2 e_t + \beta_0 \int_t^T Y_r^2 e_r dr + \int_t^T (\beta_1 Y_r^2 + Z_r^2) e_r d\langle W \rangle_r \\ & \leq \xi^2 e_T + \frac{1}{\beta_0} \int_t^T g(r)^2 e_r dr + \frac{1}{\beta_1} \int_t^T f(r)^2 e_r d\langle W \rangle_r + 2 \left| \int_t^T Y_r Z_r e_r dW_r \right|. \end{aligned} \quad (5.5)$$

By Doob's maximal inequality,

$$\begin{aligned} \mathbb{E}_\lambda \left(\sup_{0 \leq t \leq T} \left| \int_t^T Y_r Z_r e_r dW_r \right| \right) & \leq 2 \mathbb{E}_\lambda \left(\sup_{0 \leq t \leq T} \left| \int_0^t Y_r Z_r e_r dW_r \right| \right) \\ & \leq 2 \mathbb{E}_\lambda \left[\left(\int_0^T Y_r^2 Z_r^2 e_r^2 d\langle W \rangle_r \right)^{1/2} \right] \leq \frac{1}{4} \|(Y, Z)\|_{\mathcal{V}_\lambda^\beta[0, T]}^2 + 4 \mathbb{E}_\lambda \left(\int_0^T Z_r^2 e_r d\langle W \rangle_r \right), \end{aligned} \quad (5.6)$$

By the above, (5.5) and (5.4), we obtain that

$$\|(Y, Z)\|_{\mathcal{V}_\lambda^\beta[0, T]}^2 \leq \frac{1}{2} \|(Y, Z)\|_{\mathcal{V}_\lambda^\beta[0, T]}^2 + 5 \mathbb{E}_\lambda \left(\xi^2 e_T + \frac{1}{\beta_0} \int_0^T g(r)^2 e_r dr + \frac{1}{\beta_1} \int_0^T f(r)^2 e_r d\langle W \rangle_r \right).$$

which, together with a localization argument if necessary, completes the proof. \square

The following a priori estimate is crucial to the proof of Theorem 3.5.

Lemma 5.2. *Let $\beta = (\beta_0, \beta_2) \in [1, \infty)^2$. Suppose that ξ, g, f satisfy (A.1)–(A.4). In view of Lemma 5.1, the BSDE*

$$\begin{cases} dY_t = -g(t, y_t)dt - f(t, y_t, z_t)d\langle W \rangle_t + Z_t dW_t, & t \in [0, T], \\ Y_T = \xi, \end{cases}$$

admits a unique solution (Y, Z) in $\mathcal{V}_\lambda^\beta[0, T]$ for any $(y, z) \in \mathcal{V}_\lambda^\beta[0, T]$. Let $F : \mathcal{V}_\lambda^\beta[0, T] \rightarrow \mathcal{V}_\lambda^\beta[0, T]$ be the solution map $(y, z) \mapsto (Y, Z)$. If $(\bar{y}, \bar{z}) \in \mathcal{V}_\lambda^\beta[0, T]$ and $(\bar{Y}, \bar{Z}) = F(\bar{y}, \bar{z})$, then

$$\|(\hat{Y}, \hat{Z})\|_{\mathcal{V}_\lambda^\beta[0, T]} \leq 3\sqrt{2}K_\beta \|(\hat{y}, \hat{z})\|_{\mathcal{V}_\lambda^\beta[0, T]},$$

where $\hat{\eta} = \eta - \bar{\eta}$ for $\eta = y, z, Y, Z$, and

$$K_\beta^2 = \frac{K_0^2}{\beta_0} + \frac{K_1^2}{\beta_1}. \quad (5.7)$$

Moreover, F is a $\|\cdot\|_{\mathcal{V}_\lambda^\beta[0, T]}$ -contraction when β_0, β_1 are sufficiently large ($\beta_i \geq 36K_i^2$, $i = 0, 1$ will suffice).

Proof. Let $\hat{g}_t = g(t, y_t) - g(t, \bar{y}_t)$, $\hat{f}_t = f(t, y_t, z_t) - f(t, \bar{y}_t, \bar{z}_t)$. Then $|\hat{g}_t| \leq \frac{K_0}{2} |\hat{y}_t|$, $|\hat{f}_t| \leq \frac{K_0}{2} |\hat{y}_t| + K_1 |\hat{z}_t|$, and

$$d\hat{Y}_t = -\hat{g}_t dt - \hat{f}_t d\langle W \rangle_t + \hat{Z}_t dW_t, \quad \hat{Y}_T = 0.$$

Let $e_t = \exp(2\beta_0 t + 2\beta_1 \langle W \rangle_t)$. Similar to the derivation of (5.5), we have

$$\begin{aligned} & \hat{Y}_t^2 e_t + \beta_0 \int_t^T \hat{Y}_r^2 e_r dr + \int_t^T (\beta_1 \hat{Y}_r^2 + \hat{Z}_r^2) e_r d\langle W \rangle_r \\ & \leq K_\beta^2 \int_t^T \hat{y}_r^2 e_r dr + K_\beta^2 \int_t^T \hat{z}_r^2 e_r d\langle W \rangle_r - 2 \int_t^T \hat{Y}_r \hat{Z}_r e_r dW_r, \end{aligned} \quad (5.8)$$

where $K_\beta > 0$ is given by (5.7). Proceeding as in the proof of Lemma 5.1, we obtain

$$\|(\hat{Y}, \hat{Z})\|_{\mathcal{V}_\lambda^\beta[0, T]}^2 \leq \frac{1}{2} \|(\hat{Y}, \hat{Z})\|_{\mathcal{V}_\lambda^\beta[0, T]}^2 + 9K_\beta^2 \|(\hat{y}, \hat{z})\|_{\mathcal{V}_\lambda^\beta[0, T]}^2,$$

which completes the proof. \square

We are now in a position to prove Theorem 3.5.

Proof of Theorem 3.5. (a) Suppose that (Y, Z) and (\bar{Y}, \bar{Z}) are two pairs of $\{\mathcal{F}_t^\lambda\}$ -adapted processes satisfying (3.2). Denote $\hat{\eta} = \eta - \bar{\eta}$ for $\eta = y, z, Y, Z$, and let $\hat{g}_t = g(t, Y_t) - g(t, \bar{Y}_t)$, $\hat{f}_t = f(t, Y_t, Z_t) - f(t, \bar{Y}_t, \bar{Z}_t)$. Similar to the derivation of (5.8), we have

$$\begin{aligned} & \hat{Y}_t^2 e_t + \beta_0 \int_t^T \hat{Y}_r^2 e_r dr + \int_t^T (\beta_1 \hat{Y}_r^2 + \hat{Z}_r^2) e_r d\langle W \rangle_r \\ & \leq K_\beta^2 \int_t^T \hat{Y}_r^2 e_r dr + K_\beta^2 \int_t^T \hat{Z}_r^2 e_r d\langle W \rangle_r - 2 \int_t^T \hat{Y}_r \hat{Z}_r e_r dW_r, \end{aligned}$$

where $K_\beta > 0$ is given by (5.7). Setting $\beta_i = 4K_i^2$, $i = 0, 1$ in the above gives

$$\mathbb{E}_\lambda \left(\hat{Y}_t^2 e_t + \frac{1}{2} \int_t^T \hat{Y}_r^2 e_r dr + \int_t^T \left(\frac{1}{2} \hat{Y}_r^2 + \hat{Z}_r^2 \right) e_r d\langle W \rangle_r \right) \leq 0, \quad t \in [0, T],$$

which completes the proof of (a).

(b) Suppose, in addition, that (A.4) is satisfied. Let $(Y^{(0)}, Z^{(0)}) = (0, 0)$. By virtue of Lemma 5.2, define inductively $(Y^{(n)}, Z^{(n)}) \in \mathcal{V}_\lambda^\beta[0, T]$, $n \in \mathbb{N}_+$ to be the unique solution in $\mathcal{V}_\lambda^\beta[0, T]$ of the BSDE

$$\begin{cases} dY_t^{(n)} = -g(t, Y_t^{(n-1)})dt - f(t, Y_t^{(n-1)}, Z_t^{(n-1)})d\langle W \rangle_t + Z_t^{(n)}dW_t, & t \in [0, T], \\ Y_T^{(n)} = \xi. \end{cases}$$

By Lemma 5.2,

$$\begin{aligned} & \|(Y^{(n+1)} - Y^{(n)}, Z^{(n+1)} - Z^{(n)})\|_{\mathcal{V}_\lambda^\beta[0, T]} \\ & \leq K_\beta \|(Y^{(n)} - Y^{(n-1)}, Z^{(n)} - Z^{(n-1)})\|_{\mathcal{V}_\lambda^\beta[0, T]}, \quad n \in \mathbb{N}_+, \end{aligned} \quad (5.9)$$

where $K_\beta > 0$ is given by (5.7). By Lemma 5.1,

$$\begin{aligned} \|(Y^{(1)}, Z^{(1)})\|_{\mathcal{V}_\lambda^\beta[0, T]}^2 & \leq 10 \left[\mathbb{E}_\lambda(\xi^2 e^{2\beta_0 T + 2\beta_1 \langle W \rangle_T}) + \mathbb{E}_\lambda \left(\int_0^T g(r, 0)^2 e^{2\beta_0 r + 2\beta_1 \langle W \rangle_r} dr \right) \right. \\ & \quad \left. + \mathbb{E}_\lambda \left(\int_0^T f(r)^2 e^{2\beta_0 r + 2\beta_1 \langle W \rangle_r} d\langle W \rangle_r \right) \right]. \end{aligned} \quad (5.10)$$

Choose $\beta_0, \beta_1 > 0$ sufficiently large so that $K_\beta < 1$ ($\beta_i > 36K_i^2$, $i = 0, 1$ will suffice). By (5.9) and (5.10), we conclude that $(Y^{(n)}, Z^{(n)})$, $n \in \mathbb{N}_+$ is a Cauchy sequence in $\mathcal{V}_\lambda^\beta[0, T]$. Moreover, $\lim_{n \rightarrow \infty} \|(Y^{(n)} - Y, Z^{(n)} - Z)\|_{\mathcal{V}_\lambda^\beta[0, T]} = 0$ for some $(Y, Z) \in \mathcal{V}_\lambda^\beta[0, T]$ satisfying (3.4).

Clearly, (Y, Z) is a solution of (3.1), and the proof is completed. \square

5.2 BSDEs with stochastic durations

In this subsection, we prove the existence and uniqueness of solutions of the BSDE (3.5). As in [30] (see also [39, Section 7.3.2]), we shall use the method of continuity. As before, we start with the simple case where g, f do not depend on y or z , that is,

$$\begin{cases} dY_t = -g(t)dt - f(t)d\langle W \rangle_t + Z_t dW_t, & t \in [0, \tau), \\ Y_\tau = \xi, \end{cases} \quad (5.11)$$

where $g(t) = g(t, \omega)$ and $f(t) = f(t, \omega)$ are $\{\mathcal{F}_t^\lambda\}$ -adapted processes.

Lemma 5.3. *Let $\beta = (\beta_0, \beta_1) \in [1, \infty)^2$, and $\xi \in \mathcal{F}_\tau^\lambda$ which satisfy (A'.1). Suppose that $g(t), f(t)$ are $\{\mathcal{F}_t^\lambda\}$ -adapted processes such that*

$$\mathbb{E}_\lambda \left(\int_0^\tau g(t)^2 e^{2\beta_0 t + 2\beta_1 \langle W \rangle_t} dt \right) < \infty, \quad \mathbb{E}_\lambda \left(\int_0^\tau f(t)^2 e^{2\beta_0 t + 2\beta_1 \langle W \rangle_t} d\langle W \rangle_t \right) < \infty. \quad (5.12)$$

Then BSDE (5.11) admits a unique solution (Y, Z) in $\mathcal{V}_\lambda^\beta[0, \tau]$. Moreover,

$$\begin{aligned} \|(Y, Z)\|_{\mathcal{V}_\lambda^\beta[0, \tau]}^2 &\leq C \left[\mathbb{E}_\lambda \left(\xi^2 e^{2\beta_0 \tau + 2\beta_1 \langle W \rangle_\tau} \right) + \mathbb{E}_\lambda \left(\int_0^\tau g(t)^2 e^{2\beta_0 t + 2\beta_1 \langle W \rangle_t} dt \right) \right. \\ &\quad \left. + \mathbb{E}_\lambda \left(\int_0^\tau f(t)^2 e^{2\beta_0 t + 2\beta_1 \langle W \rangle_t} d\langle W \rangle_t \right) \right], \end{aligned} \quad (5.13)$$

where $C > 0$ is a constant depending only on β .

Proof. Let $M_t = \mathbb{E}_\lambda \left(\xi + \int_0^\tau g(r)dr + \int_0^\tau f(r)d\langle W \rangle_r \mid \mathcal{F}_t^\lambda \right)$, $t \geq 0$. Then, by (A'.1) and (5.12), M_t is an $\{\mathcal{F}_t^\lambda\}$ -adapted square-integrable martingale and $\mathbb{E}_\lambda(M_\tau^2) < \infty$. By Theorem 4.2, there exists a unique $\{\mathcal{F}_t^\lambda\}$ -predictable process Z such that $M_t - M_0 = \int_0^t Z_r dW_r$, $t \geq 0$.

Let $Y_t = M_{t \wedge \tau} - \int_0^{t \wedge \tau} g(r)dr - \int_0^{t \wedge \tau} f(r)dW_r$, $t \geq 0$. Then (3.6) is satisfied. Let $e_t = \exp(2\beta_0 t + 2\beta_1 \langle W \rangle_t)$, $t \geq 0$. By the definitions of Y_t and M_t , we have

$$|Y_{T \wedge \tau} - \xi|^2 e_{T \wedge \tau} = \left| \mathbb{E}_\lambda \left(\xi + \int_{T \wedge \tau}^\tau g(r)dr + \int_{T \wedge \tau}^\tau f(r)d\langle W \rangle_r \mid \tilde{\mathcal{F}}_{T \wedge \tau}^\lambda \right) - \xi \right|^2 e_{T \wedge \tau},$$

which implies that

$$\begin{aligned} \mathbb{E}_\lambda(|Y_{T \wedge \tau} - \xi|^2 e_{T \wedge \tau}) &\leq 3\mathbb{E}_\lambda[|\mathbb{E}_\lambda(\xi | \tilde{\mathcal{F}}_{T \wedge \tau}^\lambda) - \xi|^2 e_{T \wedge \tau}] \\ &\quad + 3\mathbb{E}_\lambda \left(\int_{T \wedge \tau}^\tau g(r)^2 e_r dr \right) + 3\mathbb{E}_\lambda \left(\int_{T \wedge \tau}^\tau f(r)^2 e_r d\langle W \rangle_r \right). \end{aligned}$$

By the Lebesgue dominated convergence theorem, the last two expectations on the right hand side of the above converge to zero as $T \rightarrow \infty$. For the first expectation on the right hand side of the above, note that $\lim_{T \rightarrow \infty} \mathbb{E}_\lambda(\xi | \mathcal{F}_{T \wedge \tau}^\lambda) = \xi$ \mathbb{P}_λ -a.s. by the martingale convergence theorem, which, together with (A'.1) and the dominated convergence theorem, implies that

$$\lim_{T \rightarrow \infty} \mathbb{E}_\lambda[|\mathbb{E}_\lambda(\xi | \mathcal{F}_{T \wedge \tau}^\lambda) - \xi|^2 e_{T \wedge \tau}] = 0.$$

This completes the proof of (3.7).

The proof of (5.13) is similar to that of (5.2). As a corollary of (5.13), we see that (Y, Z) is the unique solution of (5.11) in $\mathcal{V}_\lambda^\beta[0, \tau]$. \square

Next, we consider the following BSDE parametrized by $\alpha \in [0, 1]$:

$$\begin{cases} dY_t = -(g_0(t) + \alpha g(t, Y_t))dt \\ \quad - (f_0(t) + \alpha f(t, Y_t, Z_t))d\langle W \rangle_t + Z_t dW_t, & t \in [0, \tau), \\ Y_\tau = \xi, \end{cases} \quad (5.14)$$

where $g_0(t)$ and $f_0(t)$ are $\{\mathcal{F}_t^\lambda\}$ -adapted processes, and f, g are functions satisfying the measurability condition (M) in Section 3. The following a priori estimate is crucial to the proof of existence and uniqueness of solutions of BSDEs with random durations.

Lemma 5.4. *Let $\xi, \bar{\xi} \in \mathcal{F}_\tau^\lambda$ satisfy (A'.1). Let g_0, f_0 and \bar{g}_0, \bar{f}_0 be $\{\mathcal{F}_t^\lambda\}$ -adapted processes satisfying (5.12). Suppose that g, f satisfy (A'.2)–(A'.7) with $\kappa_0 = 0, \kappa_1 = \frac{K_1^2}{2}$. Let (Y, Z) be a solution in $\mathcal{V}_\lambda^\beta[0, \tau]$ of (5.14), and (\bar{Y}, \bar{Z}) be a solution in $\mathcal{V}_\lambda^\beta[0, \tau]$ of the BSDE obtained by replacing (ξ, g_0, f_0) by $(\bar{\xi}, \bar{g}_0, \bar{f}_0)$ in (5.14). Then*

$$\|(\hat{Y}, \hat{Z})\|_{\mathcal{V}_\lambda^\beta[0, \tau]}^2 \leq C\mathbb{E}_\lambda \left(\int_0^\tau \hat{g}_0(r)^2 e^{2\beta_0 r + 2\beta_1 \langle W \rangle_r} dr + \int_0^\tau \hat{f}_0(r)^2 e^{2\beta_0 r + 2\beta_1 \langle W \rangle_r} d\langle W \rangle_r \right), \quad (5.15)$$

where $\hat{\eta} = \eta - \bar{\eta}$ for $\eta = g_0, f_0, Y, Z$, and $C > 0$ is a constant depending only on β .

Proof. Let $e_t = \exp(2\beta_0 t + 2\beta_1 \langle W \rangle_t)$, and $\hat{g}(t) = g(t, Y_t) - g(t, \bar{Y}_t)$, $\hat{f}(t) = f(t, Y_t, Z_t) - f(t, \bar{Y}_t, \bar{Z}_t)$. By Itô's formula and by using (A'.2)–(A'.5), we have

$$\begin{aligned} \hat{Y}_{t \wedge \tau}^2 e_{t \wedge \tau} &\leq \hat{Y}_{T \wedge \tau}^2 e_{T \wedge \tau} + \int_{t \wedge \tau}^{T \wedge \tau} (2\hat{Y}_r \hat{g}_0(r) - 2\beta_0 \hat{Y}_r^2) e_r dr \\ &\quad + \int_{t \wedge \tau}^{T \wedge \tau} [2\hat{Y}_r \hat{f}_0(r) + 2\alpha K_1 |\hat{Y}_r| |\hat{Z}_r| - (2\alpha \kappa_1 + 2\beta_1) \hat{Y}_r^2 - \hat{Z}_r^2] e_r d\langle W \rangle_r \\ &\quad - 2 \int_{t \wedge \tau}^{T \wedge \tau} \hat{Y}_r \hat{Z}_r e_r dW_r \\ &\leq \hat{Y}_{T \wedge \tau}^2 e_{T \wedge \tau} - \beta_0 \int_{t \wedge \tau}^{T \wedge \tau} \hat{Y}_r^2 e_r dr + \frac{1}{\beta_0} \int_{t \wedge \tau}^{T \wedge \tau} \hat{g}_0(r)^2 e_r dr \\ &\quad + \int_{t \wedge \tau}^{T \wedge \tau} \left[(a - 2\alpha \kappa_1 + b\alpha^2 K_1^2 - 2\beta_1) \hat{Y}_r^2 + \left(\frac{1}{b} - 1 \right) \hat{Z}_r^2 + \frac{1}{a} \hat{f}_0(r)^2 \right] e_r d\langle W \rangle_r \\ &\quad - 2 \int_{t \wedge \tau}^{T \wedge \tau} \hat{Y}_r \hat{Z}_r e_r dW_r, \end{aligned}$$

where a and b are positive constants to be determined.

Since $\kappa_1 = \frac{K_1^2}{2}$, we may choose $b > 1$ sufficiently close to 1, and choose accordingly $a > 0$ sufficiently small such that $a - 2\kappa_1 + bK_1^2 - 2\beta_1 < 0$. Since $\alpha \mapsto a - 2\alpha \kappa_1 + b\alpha^2 K_1^2 - 2\beta_1$ is convex and is negative at $\alpha = 0$ and $\alpha = 1$, we see that $a - 2\alpha \kappa_1 + b\alpha^2 K_1^2 - 2\beta_1 < 0$ for each $\alpha \in [0, 1]$. With such a and b , (5.15) follows easily from an argument similar to the proof of (5.2). \square

Corollary 5.5. *Let g, f satisfy (A'.2)–(A'.7). Then there exists an $\epsilon_0 > 0$, depending only on K_0, K_1 and β , such that the following holds: If, for some $\alpha \in [0, 1]$, (5.14) admits a unique solution (Y, Z) in $\mathcal{V}_\lambda^\beta[0, \tau]$ such that*

$$\begin{aligned} \|(Y, Z)\|_{\mathcal{V}_\lambda^\beta[0, \tau]}^2 &\leq C\mathbb{E}_\lambda \left(\xi^2 e^{2\beta_0 \tau + 2\beta_1 \langle W \rangle_\tau} + \int_0^\tau (g_0(t)^2 + g(t, 0)^2) e^{2\beta_0 t + 2\beta_1 \langle W \rangle_t} dt \right. \\ &\quad \left. + \int_0^\tau (f_0(t)^2 + f(t, 0, 0)^2) e^{2\beta_0 t + 2\beta_1 \langle W \rangle_t} d\langle W \rangle_t \right), \end{aligned} \quad (5.16)$$

for any ξ satisfying (A'.1) and any g_0, f_0 satisfying (5.12), where $C > 0$ is a constant depending only on K_0, K_1 and β , then the same is valid when replacing α by $\alpha + \epsilon$ with $\epsilon \in [0, \epsilon_0]$ and $\alpha + \epsilon \leq 1$, and (5.16) holds for a possibly different constant $C > 0$ depending only on K_0, K_1 and β .

Proof. Suppose that (5.14) admits a unique solution in $\mathcal{V}_\lambda^\beta[0, \tau]$ satisfying (5.16) for some $\alpha \in [0, 1]$. Let $\epsilon > 0$ and $(Y_0, Z_0) = (0, 0)$. By (A'.6) and Lemma 5.3, define inductively $(Y^{(n)}, Z^{(n)})$, $n \in \mathbb{N}_+$

as the unique solution in $\mathcal{V}_\lambda^\beta[0, \tau]$ of the following BSDE

$$\begin{cases} dY_t^{(n)} = -[g_0(t) + \epsilon g(t, Y_t^{(n-1)}) + \alpha g(t, Y_t^{(n)})]dt \\ \quad - [f_0(t) + \epsilon f(t, Y_t^{(n-1)}, Z_t^{(n-1)}) + \alpha f(t, Y_t^{(n)}, Z_t^{(n)})]d\langle W \rangle_t \\ \quad + Z_t^{(n)}dW_t, & t \in [0, \tau), \\ Y_\tau^{(n)} = \xi. \end{cases}$$

According to (5.16) and Lemma 5.4, we have

$$\begin{aligned} \|(Y^{(1)}, Z^{(1)})\|_{\mathcal{V}_\lambda^\beta[0, \tau]}^2 &\leq C \mathbb{E}_\lambda \left(\xi^2 e^{2\beta_0\tau + 2\beta_1\langle W \rangle_\tau} + \int_0^\tau (g_0(t)^2 + g(t, 0)^2) e^{2\beta_0t + 2\beta_1\langle W \rangle_t} dt \right. \\ &\quad \left. + \int_0^\tau (f_0(t)^2 + f(t, 0, 0)^2) e^{2\beta_0t + 2\beta_1\langle W \rangle_t} d\langle W \rangle_t \right), \end{aligned} \quad (5.17)$$

and

$$\|(Y^{(n+1)} - Y^{(n)}, Z^{(n+1)} - Z^{(n)})\|_{\mathcal{V}_\lambda^\beta[0, \tau]} \leq \epsilon C \|(Y^{(n)} - Y^{(n-1)}, Z^{(n)} - Z^{(n-1)})\|_{\mathcal{V}_\lambda^\beta[0, \tau]}, \quad n \in \mathbb{N}_+,$$

where $C > 0$ is a constant depending only on K_0, K_1 and β . In particular, C is independent of α or ϵ . Let $\epsilon_0 = (4C)^{-1/2}$. Then for each $\epsilon \in [0, \epsilon_0]$ with $\alpha + \epsilon \leq 1$,

$$\|(Y^{(n+1)} - Y^{(n)}, Z^{(n+1)} - Z^{(n)})\|_{\mathcal{V}_\lambda^\beta[0, \tau]} \leq 2^{-n} \|(Y^{(1)}, Z^{(1)})\|_{\mathcal{V}_\lambda^\beta[0, \tau]}, \quad n \in \mathbb{N}_+.$$

This implies that $\lim_{n \rightarrow \infty} \|(Y^{(n)} - Y, Z^{(n)} - Z)\|_{\mathcal{V}_\lambda^\beta[0, \tau]} = 0$ for some $(Y, Z) \in \mathcal{V}_\lambda^\beta[0, \tau]$.

Clearly, (Y, Z) is the unique solution in $\mathcal{V}_\lambda^\beta[0, \tau]$ for the BSDE obtained by replacing α by $\alpha + \epsilon$ in (5.14). Moreover, $\|(Y, Z)\|_{\mathcal{V}_\lambda^\beta[0, \tau]} \leq 2\|(Y^{(1)}, Z^{(1)})\|_{\mathcal{V}_\lambda^\beta[0, \tau]}$. This, together with (5.17), completes the proof. \square

We now give the proof of Theorem 3.10.

Proof of Theorem 3.10. (a) This can be proved similarly to Theorem 3.5-(a).

(b) Suppose first that $\kappa_0 = 0, \kappa_1 = K_1^2/2$. By Lemma 5.3, when $\alpha = 0$,

$$\begin{cases} dY_t = -\alpha g(t, X_t, Y_t)dt - \alpha f(t, X_t, Y_t, Z_t)d\langle W \rangle_t + Z_t dW_t, & t \in [0, \tau), \\ Y_\tau = \xi, \end{cases}$$

admits a unique solution $(Y, Z) \in \mathcal{V}_\lambda^\beta[0, \tau]$ satisfying $\|(Y, Z)\|_{\mathcal{V}_\lambda^\beta[0, \tau]}^2 \leq C \mathbb{E}_\lambda (\xi^2 e^{2\beta_0\tau + 2\beta_1\langle W \rangle_\tau})$,

where and thereafter, $C > 0$ denotes any instance of a generic constant depending only on $\kappa_0, \kappa_1, K_0, K_1, \beta$.

By Corollary 5.5, there exists an $\epsilon_0 > 0$ depending only on K_0, K_1, β and satisfying the property stated in Corollary 5.5. Applying Corollary 5.5 successively shows that 3.5 admits a unique solution $(Y, Z) \in \mathcal{V}_\lambda^\beta[0, \tau]$ satisfying (3.8).

For the general case, we use the exponential martingale $\tilde{e}_t = \exp \left[-\kappa_0 t + \left(\frac{K_1^2}{2} - \kappa_1 \right) \langle W \rangle_t \right]$, $t \geq 0$, and let

$$\tilde{\xi} = \xi \tilde{e}_\tau, \quad \tilde{g}(t, y) = g(t, y \tilde{e}_t) \tilde{e}_t^{-1}, \quad \tilde{f}(t, y, z) = f(t, y \tilde{e}_t, z \tilde{e}_t) \tilde{e}_t^{-1},$$

for each $t \geq 0, y, z \in \mathbb{R}$. Then \tilde{g}, \tilde{f} satisfy the assumptions of the above case. Let

$$\tilde{\beta}_0 = \beta_0 - \kappa_0 > 0, \quad \tilde{\beta}_1 = \beta_1 - \kappa_1 + \frac{K_1^2}{2} > 0.$$

Then

$$\begin{cases} d\tilde{Y}_t = -\tilde{g}(t, \tilde{Y}_t)dt - \tilde{f}(t, \tilde{Y}_t, \tilde{Z}_t)d\langle W \rangle_t + \tilde{Z}_t dW_t, & t \in [0, \tau), \\ \tilde{Y}_\tau = \tilde{\xi}, \end{cases}$$

admits a unique solution $(\tilde{Y}, \tilde{Z}) \in \mathcal{V}_\lambda^{\tilde{\beta}}[0, \tau]$ such that

$$\begin{aligned} \|(\tilde{Y}, \tilde{Z})\|_{\mathcal{V}_\lambda^{\tilde{\beta}}[0, \tau]}^2 &\leq C \mathbb{E}_\lambda \left(\tilde{\xi}^2 e^{2\tilde{\beta}_0\tau + 2\tilde{\beta}_1\langle W \rangle_\tau} + \int_0^\tau \tilde{g}(t, 0)^2 e^{2\tilde{\beta}_0t + 2\tilde{\beta}_1\langle W \rangle_t} dt \right. \\ &\quad \left. + \int_0^\tau \tilde{f}(t, 0, 0)^2 e^{2\tilde{\beta}_0t + 2\tilde{\beta}_1\langle W \rangle_t} d\langle W \rangle_t \right) \\ &= C \mathbb{E}_\lambda \left(\xi^2 e^{2\beta_0\tau + 2\beta_1\langle W \rangle_\tau} + \int_0^\tau g(t, 0)^2 e^{2\beta_0t + 2\beta_1\langle W \rangle_t} dt \right. \\ &\quad \left. + \int_0^\tau f(t, 0, 0)^2 e^{2\beta_0t + 2\beta_1\langle W \rangle_t} d\langle W \rangle_t \right). \end{aligned}$$

Let $Y_t = \tilde{Y}_t \tilde{e}_t^{-1}$, $Z_t = \tilde{Z}_t \tilde{e}_t^{-1}$, $t \geq 0$. It is easily seen that $(Y, Z) \in \mathcal{V}_\lambda^\beta[0, \tau]$ is a solution of (3.5), and (Y, Z) satisfies (3.8). Thus we have completed the proof. \square

5.3 Example

Let $\lambda \in \mathcal{P}(\mathbb{S})$, and let τ be an $\{\mathcal{F}_t^\lambda\}$ -stopping time such that $\tau \leq T$ \mathbb{P}_λ -a.s. for some $T > 0$. We present a worked-out solution of linear BSDEs on $(\Omega, \{\mathcal{F}_t^\lambda\}, \mathbb{P}_\lambda)$. In contrast with BSDEs on Euclidean spaces, linear BSDEs on \mathbb{S} can take the following two forms:

$$\begin{cases} dY_t = -aY_t dt - cZ_t d\langle W \rangle_t + Z_t dW_t, & t \in [0, \tau), \\ Y_\tau = \xi, \end{cases}$$

and

$$\begin{cases} dY_t = -(aY_t + cZ_t) d\langle W \rangle_t + Z_t dW_t, & t \in [0, \tau), \\ Y_\tau = \xi, \end{cases}$$

where $a, c \in \mathbb{R}$ are constants, and $\xi \in L^p(\mathcal{F}_\tau^\lambda, \mathbb{P}_\lambda)$ for some $p > 2$. Clearly, these cases can be unified as

$$\begin{cases} dY_t = -aY_t dt - (bY_t + cZ_t) d\langle W \rangle_t + Z_t dW_t, & t \in [0, \tau), \\ Y_\tau = \xi, \end{cases} \quad (5.18)$$

where $a, b, c \in \mathbb{R}$ are constants.

To solve (5.18), let

$$\Phi_t = \exp \left[at + \left(b - \frac{c^2}{2} \right) \langle W \rangle_t + cW_t \right], \quad t \geq 0.$$

Then, by Lemma 4.11 and Corollary 4.14.(a), $\mathbb{E}_\lambda(\Phi_t^q) < \infty$ for any $t \geq 0$ and any $q > 0$.

By Itô's formula,

$$d\Phi_t = a\Phi_t dt + b\Phi_t d\langle W \rangle_t + c\Phi_t dW_t, \quad t \geq 0. \quad (5.19)$$

Furthermore, if (Y, Z) is the solution of (5.18) then, by (5.18) and (5.19),

$$d(\Phi_t Y_t) = Y_t d\Phi_t + \Phi_t dY_t + d\langle \Phi, Y \rangle_t = \Phi_t (cY_t + Z_t) dW_t.$$

Therefore,

$$\Phi_{t \wedge \tau} Y_{t \wedge \tau} = \Phi_\tau \xi - \int_{t \wedge \tau}^\tau \Phi_r (cY_r + Z_r) dW_r, \quad t \geq 0.$$

Taking conditional expectations on both sides of the above gives that

$$\Phi_t Y_t = \Phi_{t \wedge \tau} Y_{t \wedge \tau} = \mathbb{E}_\lambda(\Phi_\tau \xi | \mathcal{F}_{t \wedge \tau}^\lambda) = \mathbb{E}_\lambda(\Phi_\tau \xi | \mathcal{F}_t^\lambda), \quad t \geq 0.$$

Equivalently,

$$Y_t = \Phi_t^{-1} \mathbb{E}_\lambda(\Phi_\tau \xi | \mathcal{F}_t^\lambda), \quad t \geq 0. \quad (5.20)$$

Since $\xi \in L^p(\mathcal{F}_\tau^\lambda, \mathbb{P}_\lambda)$ for some $p > 2$, $\Phi_\tau \xi \in L^2(\mathcal{F}_\tau^\lambda, \mathbb{P}_\lambda)$. Therefore, by Theorem 4.2, there exists a unique $\{\mathcal{F}_t^\lambda\}$ -predictable process $\zeta(t)$ such that

$$\mathbb{E}_\lambda(\Phi_\tau \xi | \mathcal{F}_t^\lambda) = \mathbb{E}_\lambda(\Phi_\tau \xi) + \int_0^t \zeta(r) dW_r, \quad t \geq 0. \quad (5.21)$$

By a similar argument for (5.19), we have

$$d\Phi_t^{-1} = -a\Phi_t^{-1}dt - (b - c^2)\Phi_t^{-1}d\langle W \rangle_t - c\Phi_t^{-1}dW_t, \quad t \geq 0.$$

By (5.20), (5.21) and the above equation,

$$\begin{aligned} dY_t &= \Phi_t^{-1}\zeta(t)dW_t + \mathbb{E}_\lambda(\Phi_\tau \xi | \mathcal{F}_t^\lambda)d\Phi_t^{-1} - c\Phi_t^{-1}\zeta(t)d\langle W \rangle_t \\ &= -aY_tdt - [bY_t + c(\Phi_t^{-1}\zeta(t) - cY_t)]d\langle W \rangle_t + (\Phi_t^{-1}\zeta(t) - cY_t)dW_t. \end{aligned} \quad (5.22)$$

Let

$$Z_t = \Phi_t^{-1}\zeta(t) - cY_t = \Phi_t^{-1}[\zeta(t) - c\mathbb{E}_\lambda(\Phi_\tau \xi | \mathcal{F}_t^\lambda)], \quad t \geq 0. \quad (5.23)$$

Then, by (5.22), (Y, Z) given by (5.20) and (5.23) is the solution of (5.18).

6 Representations for solutions of semi-linear parabolic PDEs

In this section, we give the proof of Theorem 3.19.

Proof of Theorem 3.19. We prove the theorem by several steps.

Step 1. *Let*

$$g^{(s)}(r, x) = g(r + s, x, u(r + s, x)), \quad f^{(s)}(r, x) = f(r + s, x, u(r + s, x), \nabla u(t + s, x)).$$

Then, for any $\eta \in \mathcal{F}(\mathbb{S} \setminus V_0)$,

$$\begin{aligned} \frac{d}{dt}\mathbb{E}_\mu[u(t \wedge \sigma^{(s)} + s, X_{t \wedge \sigma^{(s)}})\eta(X_0)] &= \langle \partial_t u(t + s), P_t^0 \eta \rangle_\mu \\ &\quad - \mathcal{E}(u(t + s), P_t^0 \eta) \quad \text{a.e. } t \in [0, T], \end{aligned} \quad (6.1)$$

$$\frac{d}{dt}\mathbb{E}_\mu\left[\left(\int_0^{t \wedge \sigma^{(s)}} g^{(s)}(r, X_r) dr\right)\eta(X_0)\right] = \langle g^{(s)}(t), P_t^0 \eta \rangle_\mu \quad \text{a.e. } t \in [0, T], \quad (6.2)$$

and

$$\frac{d}{dt}\mathbb{E}_\mu\left[\left(\int_0^{t \wedge \sigma^{(s)}} f^{(s)}(r, X_r) d\langle W \rangle_r\right)\eta(X_0)\right] = \langle f^{(s)}(t), P_t^0 \eta \rangle_\nu \quad \text{a.e. } t \in [0, T]. \quad (6.3)$$

Proof of Step 1. Without loss of generality, we may assume $s = 0$. Let $H(\varphi(t))$ be the harmonic function with boundary value $\varphi(t)$ and $u^0(t, x) = u(t, x) - H(\varphi(t))(x)$. Let $0 < \delta < T - t$. Since

$$u^0(t \wedge \sigma_{V_0}, X_{t \wedge \sigma_{V_0}}) = u^0(t, X_t)1_{\{t < \sigma_{V_0}\}},$$

and $u^0(t) = 0$ on V_0 , using the μ -symmetry of $\{X_t^0\}$, we obtain

$$\begin{aligned} \mathbb{E}_\mu[u^0((t + \delta) \wedge \sigma_{V_0}, X_{(t + \delta) \wedge \sigma_{V_0}})\eta(X_0)] &= \mathbb{E}_\mu[u^0(t + \delta, X_0^0)\eta(X_{t + \delta}^0)] \\ &= \langle u^0(t + \delta), P_{t + \delta}^0 \eta \rangle_\mu. \end{aligned} \quad (6.4)$$

Similarly, $\mathbb{E}_\mu[(u^0(t \wedge \sigma_{V_0}, X_{t \wedge \sigma_{V_0}})\eta(X_0)] = \langle u^0(t), P_t^0 \eta \rangle_\mu$. Therefore,

$$\begin{aligned} \mathbb{E}_\mu[(u((t + \delta) \wedge \sigma_{V_0}, X_{(t+\delta) \wedge \sigma_{V_0}}) - u(t \wedge \sigma_{V_0}, X_{t \wedge \sigma_{V_0}}))\eta(X_0)] &= \langle u^0(t + \delta) - u^0(t), P_{t+\delta}^0 \eta \rangle_\mu \\ &\quad + \langle u^0(t), P_{t+\delta}^0 \eta - P_t^0 \eta \rangle_\mu + \langle H[\varphi(t + \delta) - \varphi(t)], P_t^0 \eta \rangle_\mu. \end{aligned} \quad (6.5)$$

Note that $u^0(t) \in \mathcal{F}(\mathbb{S} \setminus V_0)$. We have

$$\lim_{\delta \rightarrow 0} \frac{1}{\delta} \langle u^0(t), P_{t+\delta}^0 \eta - P_t^0 \eta \rangle_\mu = -\mathcal{E}(u^0(t), P_t^0 \eta) = -\mathcal{E}(u(t), P_t^0 \eta), \quad (6.6)$$

where we have used in the second equality the fact that $\mathcal{E}(H(\varphi(t)), v) = 0$ for any $v \in \mathcal{F}(\mathbb{S} \setminus V_0)$.

By Lemma 3.16 and $\lim_{\delta \rightarrow 0} \|P_{t+\delta}^0 \eta - P_t^0 \eta\|_{L^2(\mu)} = 0$, we deduce that

$$\begin{aligned} \lim_{\delta \rightarrow 0} \frac{1}{\delta} \langle u^0(t + \delta) - u^0(t), P_{t+\delta}^0 \eta \rangle_\mu &= \lim_{\delta \rightarrow 0} \frac{1}{\delta} \langle u(t + \delta) - u(t), P_{t+\delta}^0 \eta \rangle_\mu - \langle H(\partial_t \varphi(t)), P_t^0 \eta \rangle_\mu \\ &= \langle \partial_t u(t), P_t^0 \eta \rangle_\mu - \langle H(\partial_t \varphi(t)), P_t^0 \eta \rangle_\mu. \end{aligned} \quad (6.7)$$

The equality (6.1) now follows readily from (6.5), (6.6) and (6.7).

Similar to (6.4), we have

$$\mathbb{E}_\mu \left[\left(\int_{t \wedge \sigma_{V_0}}^{(t+\delta) \wedge \sigma_{V_0}} g^{(0)}(r, X_r) dr \right) \eta(X_0) \right] = \int_0^\delta \langle g^{(0)}(t+r), P_{t+r}^0 \eta \rangle_\mu dr.$$

Using the Lebesgue dominated convergence theorem and fact that $\lim_{r \rightarrow 0} P_{t+r}^0 \eta(x) = P_t^0 \eta(x)$, $x \in \mathbb{S} \setminus V_0$, we obtain (6.2).

We now prove (6.3). Similar to the above, by the μ -symmetry of $\{X_t^0\}$ again, we have

$$\mathbb{E}_\mu \left[\left(\int_{t \wedge \sigma_{V_0}}^{(t+\delta) \wedge \sigma_{V_0}} f^{(0)}(r, X_r) d\langle W \rangle_r \right) \eta(X_0) \right] = \mathbb{E}_\mu \left[\left(\int_0^\delta f^{(0)}(t_n + r, X_r^0) d\langle W \rangle_r \right) P_t^0 \eta(X_0^0) \right].$$

Now we apply Lemma 4.10 and conclude that

$$\frac{d}{dt} \mathbb{E}_\mu \left[\left(\int_0^t f^{(0)}(r, X_r^0) d\langle W \rangle_r \right) \eta(X_0) \right] = \lim_{\delta \rightarrow 0} \frac{1}{\delta} \int_0^\delta \langle f^{(0)}(r+t), P_t^0 \eta \rangle_\nu dr = \langle f^{(0)}(t), P_t^0 \eta \rangle_\nu.$$

This completes the proof of Step 1.

Step 2. *Let*

$$\begin{aligned} M_t^{(s)} &= u(t \wedge \sigma^{(s)} + s, X_{t \wedge \sigma^{(s)}}) - u(s, X_0) + \int_0^{t \wedge \sigma^{(s)}} g^{(s)}(r, X_r) dr \\ &\quad + \int_0^{t \wedge \sigma^{(s)}} f^{(s)}(r, X_r) d\langle W \rangle_r, \quad t \geq 0. \end{aligned}$$

Then $\{M_t^{(s)}\}$ is a \mathbb{P}_x -martingale for each $x \in \mathbb{S} \setminus V_0$.

Proof of Step 2. By Step 1,

$$\frac{d}{dt} \mathbb{E}_\mu [M_t^{(s)} \eta(X_{t_0}^0)] = 0 \quad \text{for all } \eta \in \mathcal{F}(\mathbb{S} \setminus V_0) \text{ a.e. } t \geq t_0,$$

which, together with the continuity of $t \mapsto \mathbb{E}_x(M_t^{(s)})$, implies that

$$\mathbb{E}_x(M_t^{(s)}) = 0 \quad \text{for all } t \geq 0 \quad \mu\text{-a.e. } x \in \mathbb{S} \setminus V_0. \quad (6.8)$$

We claim that $x \mapsto \mathbb{E}_x(M_t^{(s)})$, $x \in \mathbb{S} \setminus V_0$ is continuous, and therefore, (6.8) holds for all $t \geq 0$ and all $x \in \mathbb{S} \setminus V_0$. Note that

$$M_t^{(s)} = u(t+s, X_t^0) - u(s, X_0^0) + \int_0^t g^{(s)}(r, X_r^0) dr + \int_0^t f^{(s)}(r, X_r^0) d\langle W \rangle_r, \quad 0 \leq t \leq T-s.$$

Recall that, in the above, we have used the convention that $\eta(\Delta) = 0$ for any function η on \mathbb{S} . Hence

$$\mathbb{E}_x(M_t^{(s)}) = P_t^0(u^0(t+s))(x) - u^0(s, x) + \int_0^t [P_r^0(g^{(s)}(r))(x) + P_r^0(f^{(s)}(r)\nu)(x)] dr.$$

Therefore, it suffices to prove the continuity of $x \mapsto \int_0^t P_r^0(f^{(s)}(r)\nu)(x) dr$, $x \in \mathbb{S} \setminus V_0$.

By Lemma 4.6-(b) and (A.3), we have

$$\begin{aligned} \|P_r^0(f^{(s)}(r)\nu)\|_{L^\infty} &\leq C \min\{1, r^{-d_s/2}\} \|f^{(s)}(r)\|_{L^1(\nu)} \\ &\leq C \min\{1, r^{-d_s/2}\} [\|f(r+s, 0, 0)\|_{L^1(\nu)} + \|u(r+s)\|_{L^1(\nu)} \\ &\quad + \|\nabla u(r+s)\|_{L^1(\nu)}] \\ &\leq C \min\{1, r^{-d_s/2}\} [1 + \max_{[0, T] \times \mathbb{S}} |u| + \mathcal{E}(u(r+s))^{1/2}] \end{aligned} \quad (6.9)$$

for all $r > 0$, where $C > 0$ is a constant depending only on K_0, K_1 and $\max_{[0, T] \times \mathbb{S}} |f(t, x, 0, 0)|$. Note that $\int_0^T \mathcal{E}(u(t)) dt < \infty$, and that, for each $r > 0$,

$$P_r^0(f^{(s)}(r)\nu)(x) = \int_{\mathbb{S}} f^{(s)}(r, y) p_r^0(x, y) \nu(dy)$$

is continuous in $x \in \mathbb{S} \setminus V_0$. Now the continuity of $x \mapsto \int_0^t P_r^0(f^{(s)}(r)\nu)(x) dr$ follows readily from (6.9) and the Lebesgue dominated convergence theorem.

Therefore, $\mathbb{E}_x(M_t^{(s)}) = 0$ for all $t \geq 0$, $x \in \mathbb{S} \setminus V_0$, which, together with the Markov property of $\{X_t^0\}$, completes the proof of Step 2.

Step 3. For each $s \in [0, T)$ and each $x \in \mathbb{S} \setminus V_0$,

$$(Y_t^{(s)}, Z_t^{(s)}) = (u(t \wedge \sigma^{(s)} + s, X_{t \wedge \sigma^{(s)}}), \nabla u(t \wedge \sigma^{(s)} + s, X_{t \wedge \sigma^{(s)}}))$$

is the solution of the BSDE (3.11). Moreover, the representation (3.12) holds, and the solution of (3.9) is unique.

Proof of Step 3. By Lemma 4.5,

$$u(t+s, X_t) = u(s, X_0) + \int_0^t \nabla u(r+s, X_r) dW_r + N_t^{(s)}, \quad t \geq 0,$$

where $N^{(s)}$ is a continuous process with zero quadratic variation. Let

$$Q_t^{(s)} = N_{t \wedge \sigma^{(s)}}^{(s)} + \int_0^{t \wedge \sigma^{(s)}} g^{(s)}(r, X_r) dr + \int_0^{t \wedge \sigma^{(s)}} f^{(s)}(r, X_r) d\langle W \rangle_r, \quad t \geq 0.$$

Then

$$Q_t^{(s)} = M_t^{(s)} - \int_0^{t \wedge \sigma^{(s)}} \nabla u(r+s, X_r) dW_r, \quad t \geq 0, \quad (6.10)$$

and therefore $\{Q_t^{(s)}\}$ is a \mathbb{P}_x -martingale for all $x \in \mathbb{S} \setminus V_0$.

Let $t_i = \frac{i}{n}t$, $0 \leq i \leq n$. Then

$$\lim_{n \rightarrow \infty} \sum_{i=1}^n (Q_{t_i}^{(s)} - Q_{t_{i-1}}^{(s)})^2 = \langle Q^{(s)} \rangle_t \text{ in } L^1(\mathbb{P}_x) \text{ for each } x \in \mathbb{S} \setminus V_0.$$

Since $\lim_{n \rightarrow \infty} \mathbb{E}_\mu \left[\sum_{i=1}^n (N_{t_i}^{(s)} - N_{t_{i-1}}^{(s)})^2 \right] = 0$, there exists a subsequence $\{n_k\}$ such that

$$\lim_{k \rightarrow \infty} \sum_{i=1}^{n_k} (N_{t_i}^{(s)} - N_{t_{i-1}}^{(s)})^2 = 0 \quad \mathbb{P}_\mu\text{-a.s.},$$

which implies that $\langle Q^{(s)} \rangle_t = 0$ \mathbb{P}_μ -a.s., as $t \mapsto \int_0^{t \wedge \sigma^{(s)}} g^{(s)}(r, X_r) dr$ and $t \mapsto \int_0^{t \wedge \sigma^{(s)}} f^{(s)}(r, X_r) d\langle W \rangle_r$ are of bounded variation.

Therefore, $\mathbb{E}_x(\langle Q \rangle_t^{(s)}) = 0$ μ -a.e. $x \in \mathbb{S} \setminus V_0$. By an argument similar to the proof of Step 2, it can be shown that $x \mapsto \mathbb{E}_x(\langle Q \rangle_t^{(s)})$ is continuous. Therefore, $\mathbb{E}_x(\langle Q \rangle_t^{(s)}) = 0$ for all $x \in \mathbb{S} \setminus V_0$. In particular, by (6.10),

$$\begin{aligned} u(t \wedge \sigma^{(s)} + s, X_{t \wedge \sigma^{(s)}}) &= u(s, X_0) - \int_0^{t \wedge \sigma^{(s)}} g^{(s)}(r, X_r) dr - \int_0^{t \wedge \sigma^{(s)}} f^{(s)}(r, X_r) d\langle W \rangle_r \\ &\quad + \int_0^{t \wedge \sigma^{(s)}} \nabla u(r + s, X_r) dW_r \quad \mathbb{P}_x\text{-a.s. for all } x \in \mathbb{S} \setminus V_0. \end{aligned}$$

This implies that

$$(Y_t^{(s)}, Z_t^{(s)}) = (u(t \wedge \sigma^{(s)} + s, X_{t \wedge \sigma^{(s)}}), \nabla u(t \wedge \sigma^{(s)} + s, X_{t \wedge \sigma^{(s)}}))$$

is the unique solution of the BSDE (3.11) on $(\Omega, \{\mathcal{F}_t\}, \mathbb{P}_x)$ for $x \in \mathbb{S} \setminus V_0$, which is clearly also valid for $x \in V_0$. As a result, we obtain the representation (3.12).

The uniqueness of the solution of (3.9) follows immediately from the representation (3.12) and the uniqueness of solutions of (3.11). \square

Example. As an application of the representation formula given in Theorem 3.19, we solve the following parabolic equation

$$\begin{cases} (\partial_t u + \mathcal{L}u) \cdot \mu = -au \cdot \mu - (bu + c\nabla u) \cdot \nu, & (t, x) \in [0, T) \times \mathbb{S} \setminus V_0, \\ u(t, x) = \varphi(t, x) \text{ on } [0, T) \times V_0, & u(T) = \psi, \end{cases}$$

where $a, b, c \in \mathbb{R}$ are constants.

Let

$$\Psi(t, x) = \begin{cases} \varphi(t, x), & \text{if } (t, x) \in [0, T) \times V_0, \\ \psi(x), & \text{if } (t, x) \in \{T\} \times \mathbb{S} \setminus V_0. \end{cases}$$

By the example in Section 5.3, we see that the solution of the BSDE

$$\begin{cases} dY_t^{(s)} = -aY_t^{(s)} dt - (bY_t^{(s)} + cZ_t^{(s)}) d\langle W \rangle_t + Z_t^{(s)} dW_t, & t \in [0, \sigma^{(s)}), \\ Y_{\sigma^{(s)}}^{(s)} = \Psi(\sigma^{(s)}, X_{\sigma^{(s)}}), \end{cases}$$

on $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq 0}, \mathbb{P}_x)$ is given by

$$Y_t^{(s)} = \Phi_{t \wedge \sigma^{(s)}}^{-1} \mathbb{E}_x[\Phi_{\sigma^{(s)}} \Psi(\sigma^{(s)}, X_{\sigma^{(s)}}) | \mathcal{F}_t], \quad t \in [0, T - s],$$

where

$$\Phi_t = \exp \left[at + \left(b - \frac{c^2}{2} \right) \langle W \rangle_t + cW_t \right], \quad t \geq 0.$$

Therefore

$$u(t, x) = \mathbb{E}_x(Y_0^{(t)}) = \mathbb{E}_x[\Phi_{(T-t) \wedge \sigma_{V_0}} \Psi((T-t) \wedge \sigma_{V_0}, X_{(T-t) \wedge \sigma_{V_0}})], \quad t \in [0, T], \quad x \in \mathbb{S}.$$

In particular, if $\varphi = 0$, then

$$u(t, x) = \mathbb{E}_x[\Phi_{T-t} \psi(X_{T-t}^0)], \quad t \in [0, T], \quad x \in \mathbb{S}.$$

Remark 6.1. It is well known that, for \mathbb{R}^d , solutions of BSDEs correspond to viscosity solutions of corresponding PDEs, which is a very weak formulation of solutions. Moreover, Theorem 3.19 shows that solutions of BSDEs correspond to the solution of the PDE (3.9) whenever a solution exists. These justify to name the functions given by (3.12) the *viscosity solutions* of (3.9). The existence of such very weak solutions is guaranteed by Theorem 3.5. On the other hand, we also note that the existence of solutions of BSDEs does not imply the existence of weak solutions of the corresponding semi-linear parabolic equations. In future work we shall explore the existence of weak solutions of (3.9) on the Sierpinski gasket.

Acknowledgements

The authors want to thank Professor Hino for his helpful suggestions on the first version of the paper. We would also like to express our gratitude to the anonymous referee for the invaluable comments and suggestions (in particular, providing many related papers), which help improve the quality of the presentation.

References

- [1] Bally, V., Pardoux, E., and Stoica, L. (2005). Backward stochastic differential equations associated to a symmetric Markov process. *Potential Analysis*, 22(1):17–60.
- [2] Barlow, M. T. and Kigami, J. (1997). Localized eigenfunctions of the Laplacian on p.c.f self-similar sets. *Journal of the London Mathematical Society*, 56(02):320–332.
- [3] Barlow, M. T. and Perkins, E. A. (1988). Brownian motion on the Sierpinski gasket. *Probability Theory and Related Fields*, 79(4):543–623.
- [4] Baudoin, F. and Kelleher, D. J. (2016). Differential forms on Dirichlet spaces and Bakry-Émery estimates on metric graphs. *preprint arXiv:1604.02520*.
- [5] Bismut, J. M. (1978). An introductory approach to duality in optimal stochastic control. *SIAM review*, 20(1):62–78.
- [6] Cipriani, F. and Sauvageot, J.-L. (2003). Derivations as square roots of Dirichlet forms. *Journal of Functional Analysis*, 201(1):78–120.
- [7] Cipriani, F. and Sauvageot, J.-L. (2009). Fredholm modules on pcf self-similar fractals and their conformal geometry. *Communications in Mathematical Physics*, 286(2):541–558.
- [8] Evans, L. C. (2010). *Partial Differential Equations*, volume 19 of *Graduate Studies in Mathematics*. American Mathematical Society, 2 edition.
- [9] Föllmer, H., Protter, P., and Shiriyayev, A. N. (1995). Quadratic covariation and an extension of Itô’s formula. *JSTOR*, 1:149–169.
- [10] Fukushima, M., Oshima, Y., and Takeda, M. (2010). *Dirichlet Forms and Symmetric Markov Processes*, volume 19 of *De Gruyter Studies in Mathematics*. De Gruyter, 2 edition.

- [11] Fukushima, M. and Shima, T. (1992). On a spectral analysis for the Sierpinski gasket. *Potential Analysis*, 1(1):1–35.
- [12] Goldstein, S. (1987). Random walks and diffusions on fractals. In *Percolation Theory and Ergodic Theory of Infinite Particle Systems*, pages 121–129. Springer.
- [13] Hino, M. (2008). Martingale dimensions for fractals. *Institute of Mathematical Statistics*, 36(3):971–991.
- [14] Hino, M. (2010). Energy measures and indices of Dirichlet forms, with applications to derivatives on some fractals. *Proceedings of the London Mathematical Society*, 100(1):269–302.
- [15] Hinz, M., Röckner, M., and Teplyaev, A. (2013). Vector analysis for Dirichlet forms and quasi-linear PDE and SPDE on metric measure spaces. *Stochastic Processes and Their Applications*, 123(12):4373–4406.
- [16] Hinz, M. and Teplyaev, A. (2013). Dirac and magnetic Schrödinger operators on fractals. *Journal of Functional Analysis*, 265(11):2830–2854.
- [17] Hinz, M. and Teplyaev, A. (2015). Finite energy coordinates and vector analysis on fractals. *Fractal Geometry and Stochastics V*, 70:209–227.
- [18] Ionescu, M., Rogers, L., and Teplyaev, A. (2012). Derivations, Dirichlet forms and spectral analysis. *Journal of Functional Analysis*, 263(8):2141–2169.
- [19] Kigami, J. (1989). A harmonic calculus on the Sierpinski spaces. *Japan Journal of Applied Mathematics*, 6(2):259–290.
- [20] Kigami, J. (1993). Harmonic metric and Dirichlet form on the Sierpinski gasket. *Pitman Research Notes in Mathematics Series*, pages 201–218.
- [21] Kigami, J. (1998). Distributions of localized eigenvalues of Laplacians on post critically finite self-similar sets. *Journal of Functional Analysis*, 156(1):170–198.
- [22] Kigami, J. (2001). *Analysis on Fractals*, volume 143 of *Cambridge Tracts in Mathematics*. Cambridge University Press.
- [23] Kigami, J. and Lapidus, M. L. (1993). Weyl’s problem for the spectral distribution of Laplacians on p.c.f. self-similar fractals. *Communications in Mathematical Physics*, 158(1):93–125.
- [24] Klimsiak, T. and Rozkosz, A. (2013). Dirichlet forms and semilinear elliptic equations with measure data. *Journal of Functional Analysis*, 265(6):890–925.
- [25] Klimsiak, T. and Rozkosz, A. (2016). Semilinear elliptic equations with measure data and quasi-regular Dirichlet forms. *Colloquium Mathematicae*, 145(1):35–67.
- [26] Kusuoka, S. (1987). A diffusion process on a fractal. In Itô, K. and Ikeda, N., editors, *Probabilistic Methods in Mathematical Physics: Proceedings of the Taniguchi International Symposium (Katata and Kyoto, 1985)*, pages 251–274. Academic Press.
- [27] Kusuoka, S. (1989). Dirichlet forms on fractals and products of random matrices. *Publications of the Research Institute for Mathematical Sciences*, 25(4):659–680.
- [28] Lyons, T. and Zheng, W. A. (1990). Diffusion processes with non-smooth diffusion coefficients and their density functions. *Proceedings of the Royal Society of Edinburgh: Section A Mathematics*, 115(3-4):231–242.

- [29] Pardoux, E. and Peng, S. (1990). Adapted solution of a backward stochastic differential equation. *Systems & Control Letters*, 14(1):55–61.
- [30] Peng, S. (1991). Probabilistic interpretation for systems of quasilinear parabolic partial differential equations. *Stochastics and Stochastic Reports*, 37(1-2):61–74.
- [31] Qian, Z. and Ying, J. (2012). Martingale representations for diffusion processes and backward stochastic differential equations. In *Séminaire de Probabilités XLIV*, pages 75–103. Springer.
- [32] Rammal, R. and Toulouse, G. (1983). Random walks on fractal structures and percolation clusters. *Journal de Physique Lettres*, 44(1):13–22.
- [33] Stannat, W. (1999). *The Theory of Generalized Dirichlet Forms and Its Applications in Analysis and Stochastics*, volume 678. American Mathematical Society.
- [34] Stoica, L. (2003). A probabilistic interpretation of the divergence and BSDE's. *Stochastic Processes and Their Applications*, 103(1):31–55.
- [35] Strichartz, R. S. (2000). Taylor approximations on Sierpinski gasket type fractals. *Journal of Functional Analysis*, 174(1):76–127.
- [36] Strichartz, R. S. (2006). *Differential Equations on Fractals: A Tutorial*. Princeton University Press.
- [37] Teplyaev, A. (2000). Gradients on fractals. *Journal of Functional Analysis*, 174:128–154.
- [38] Trutnau, G. (2000). Stochastic calculus of generalized Dirichlet forms and applications to stochastic differential equations in infinite dimensions. *Osaka Journal of Mathematics*, 37(2):315–343.
- [39] Yong, J. and Zhou, X. (1999). *Stochastic Controls: Hamiltonian Systems and HJB Equations*, volume 43. Springer Science & Business Media.
- [40] Zhu, R.-C. (2015). BSDE and generalized Dirichlet forms: the infinite dimensional case. *Forum Mathematicum*, 27(1):201–253.