

## A WEAK VERSION OF PATH-DEPENDENT FUNCTIONAL ITÔ CALCULUS

BY DORIVAL LEÃO\*, ALBERTO OHASHI<sup>†,1</sup> AND ALEXANDRE B. SIMAS<sup>†</sup>

*Universidade de São Paulo\* and Universidade Federal da Paraíba<sup>†</sup>*

We introduce a variational theory for processes adapted to the multidimensional Brownian motion filtration that provides a differential structure allowing to describe infinitesimal evolution of Wiener functionals at very small scales. The main novel idea is to compute the “sensitivities” of processes, namely derivatives of martingale components and a weak notion of infinitesimal generator, via a finite-dimensional approximation procedure based on controlled inter-arrival times and approximating martingales. The theory comes with convergence results that allow to interpret a large class of Wiener functionals beyond semimartingales as limiting objects of differential forms which can be computed path wisely over finite-dimensional spaces. The theory reveals that solutions of BSDEs are minimizers of energy functionals w.r.t. Brownian motion driving noise.

**1. Introduction.** Let  $(\Omega, \mathcal{F}, \mathbb{P})$  be a probability space equipped with a filtration  $\mathbb{F} = (\mathcal{F}_t)_{t \geq 0}$  generated by a multidimensional noise process  $W$ . The goal of this work is to present a systematic approach to concretely analyze the infinitesimal variation of a given  $\mathbb{F}$ -adapted process w.r.t.  $W$ . A similar type of question has been studied over the last four decades by means of Malliavin calculus and White Noise analysis. In this context, infinitesimal variation of smooth random variables (in the sense of Malliavin) and stochastic distributions (in the sense of Hida) are studied without taking into account an underlying filtration, so that, loosely speaking, these approaches are anticipative in nature. In the present paper, we are interested in providing a nonanticipative calculus in order to characterize variational properties of adapted processes and, more importantly, we aim to provide concrete tools to obtain “maximizers” of variational problems w.r.t.  $W$ . We also make a special effort to cover the largest possible class of processes adapted to a given  $W$ .

Recently, a new branch of stochastic calculus has appeared, known as functional Itô calculus, which results to be an extension of classical Itô calculus to nonanticipative functionals depending on the whole path of a noise  $W$  and not only on its current value; see, for example, Dupire [19], Cont and Fournié [7, 8], Cosso and Russo [14, 15], Peng and Song [41], Buckdahn, Ma, and Zhang [4], Keller and Zhang [27], Ohashi, Shamarova and Shamarov [38] and Oberhauser [37]. Inspired

---

Received November 2015; revised November 2017.

<sup>1</sup>Supported by ENSTA ParisTech.

*MSC2010 subject classifications.* Primary 60H07; secondary 60H25.

*Key words and phrases.* Stochastic calculus of variations, functional Itô calculus.

by Peng [40], the issue of providing a suitable definition of path-dependent PDEs has attracted a great interest; see, for example, Peng and Wang [42], Ekren, Keller, Touzi and Zhang [20], Ekren, Touzi and Zhang [21, 22], Ekren and Zhang [23], Cosso and Russo [13] and Flandoli and Zanco [24]. In the present work, we develop a weak version of the functional Itô calculus inspired by the discretization scheme introduced by Leão and Ohashi [30, 31].

1.1. *Main setup and contributions.* In this paper, we are interested in developing a general differentiation theory for  $\mathbb{F}$ -adapted processes where  $\mathbb{F}$  is generated by a state noise  $W$  which drives the randomness of the system. We are interested in developing a general methodology to compute infinitesimal variations of  $X$  w.r.t. the underlying state noise  $W$  in three fundamental cases:

- (i) A priori,  $X$  does not possess enough regularity w.r.t.  $W$ .
- (ii)  $X$  depends on the whole path of  $W$ .
- (iii) Explicit functional nonanticipative representations of  $X$  are not available.

Cases (i) and (iii) occur very frequently in stochastic control problems and it has been one of the motivations for the use of viscosity methods in PDEs in the Markovian setup. When situation (ii) takes place, we cannot rely on classical approaches. In order to deal with cases (i), (ii) and (iii), we develop a theory based on suitable discretizations on the level of the noise rather than  $X$  itself. The methodology can be interpreted as a weak functional stochastic calculus for path-dependent systems which can be also interpreted path wisely over finite-dimensional spaces by means of a suitable regularization procedure. In contrast to the classical approximation schemes in the literature based on deterministic discretizations on the time scale, we develop a type of space-filtration discretization procedure  $(\mathbb{F}^k)_{k \geq 1}$  which allows us to drastically reduce the dimension of the problem.

We choose the underlying state noise  $W$  as a  $d$ -dimensional Brownian motion  $B$  and the fundamental objects to be analyzed will be  $\mathbb{F}$ -adapted processes  $X$  that we call as *Wiener functionals*. The theory is designed to analyze the sensitivity of  $X$  w.r.t.  $B$  under rather general assumptions and much beyond the standard literature on smooth pathwise functional calculus and other Sobolev-type formulations. The methodology is based on suitable approximating structures

$$\mathcal{Y} = ((X^k)_{k \geq 1}, \mathcal{D})$$

equipped with a continuous-time random walk approximation  $\mathcal{D} = \{\mathcal{T}, A^{k,j}; j = 1, \dots, d, k \geq 1\}$  driven by a suitable class of waiting times  $\mathcal{T} = \{T_n^k; k \geq 1, n \geq 0\}$  [see (2.1), (2.2) and (2.5)] which encodes the evolution of the Brownian motion at small scales. The sequence of processes  $(X^k)_{k \geq 1}$  has to be interpreted as a model simplification naturally defined on finite-dimensional spaces which allows us to concretely approach variational properties of  $X$  w.r.t.  $B$  by means of suitable derivative operators  $(\mathcal{D}^{\mathcal{Y},k} X, U^{\mathcal{Y},k} X)$ , where

(1.1)  $\mathcal{D}^{\mathcal{Y},k} X$  encodes the variation of  $X$  w.r.t.  $B$

and

$$(1.2) \quad U^{\mathcal{Y},k} X \quad \text{encodes variations of } X \text{ “orthogonal” to } B.$$

The variational operators (1.1) and (1.2) are constructed via optional stochastic integration and  $\mathbb{F}^k$ -dual predictable projections, respectively, in the spirit of Del-lacherie and Meyer [17] at the level of a given probability measure. Due to the nice structure coming from  $\mathcal{D}$  which allows us to reduce the dimension, operators (1.1) and (1.2) can also be computed pathwisely in a very concrete way.

Conceptually, the approach developed in this article for analyzing the infinitesimal variation of Wiener functionals w.r.t. Brownian state at very small scales consists of three steps;

1. Based on the available information given by a Wiener functional  $X$  at hand, one designs a discrete-type structure  $\mathcal{Y} = ((X^k)_{k \geq 1}, \mathcal{D})$  where  $X^k = X^k(\mathcal{D})$  is a pure-jump process driven by a discrete-type skeleton  $\mathcal{D} = \{\mathcal{T}, A^{k,j}; j = 1, \dots, d, k \geq 1\}$ . We call  $\mathcal{Y} = ((X^k)_{k \geq 1}, \mathcal{D})$  as an *imbedded discrete structure* associated with  $X$  (see Definition 3.3).

2. At the level of imbedded discrete structures, one has concrete functionals defined on finite-dimensional spaces where one is able to compute path wisely the variational operators  $(\mathcal{D}^{\mathcal{Y},k} X, U^{\mathcal{Y},k} X)$  of  $(X^k)_{k \geq 1}$  w.r.t.  $\mathcal{D}$  freely. At this point, one tries to obtain as much information as possible of  $X$  by computing  $(\mathcal{D}^{\mathcal{Y},k} X, U^{\mathcal{Y},k} X)$ .

3. In a final step, one has to prove the information obtained in step 2 is consistent with  $X$ . At this stage, one has to prove the  $X^k \rightarrow X$  as the level of discretization  $k \rightarrow +\infty$  and, eventually if  $X$  admits enough regularity, the limits of  $(\mathcal{D}^{\mathcal{Y},k} X, U^{\mathcal{Y},k} X)$  might be used to get limiting representations and a more refined information on  $X$ .

It is important to emphasize that it is shown that each Wiener functional  $X$  is equipped with a canonical imbedded discrete structure  $\mathcal{Y} = ((X^k)_{k \geq 1}, \mathcal{D})$  [see (3.1), Lemma 3.1 and Corollary 4.3 for the particular case of Dirichlet processes]. However, our methodology requires effort on the part of the “user” in order to specify the “good” structure that is suitable for analyzing a given problem at hand. The theory generalizes other weak formulations [4, 9, 41] and also the pathwise approaches [7, 8, 14, 15, 19, 38] restricted to Brownian states.

The philosophy of this work is not to propose representations for Wiener functionals (although we present some of them), but rather a concrete way to depict the infinitesimal variation of processes at very small scales via  $(\mathcal{D}^{\mathcal{Y},k} X, U^{\mathcal{Y},k} X)$  for a given choice of structure  $\mathcal{Y}$  associated with  $X$ . In this direction, the article reveals that a large class of Wiener functionals (including Dirichlet processes) can be viewed as limiting objects of differential forms attached to imbedded discrete structures which have to be carefully designed case by case.

In Theorem 4.1, we construct a differential structure for a Wiener functional (see Definition 4.4) of the form

$$(1.3) \quad X(t) = X(0) + \sum_{j=1}^d \int_0^t H_j dB^j + V(t); \quad 0 \leq t \leq T,$$

where  $dB^j$  is the Itô integral,  $H = (H_1, \dots, H_d)$  is an  $\mathbb{F}$ -adapted square-integrable process and  $V$  has continuous paths without a priori regularity conditions. A nonanticipative process  $\mathcal{D}^{\mathcal{Y}}X$  is constructed based on limits of the variations

$$\mathcal{D}^{\mathcal{Y},k,j}X(T_n^{k,j}) = \frac{\Delta X^k(T_n^{k,j})}{\Delta A^{k,j}(T_n^{k,j})}; \quad n \geq 1, 1 \leq j \leq d$$

for a given imbedded discrete structure  $\mathcal{Y} = ((X^k)_{k \geq 1}, \mathcal{D})$ , where  $T_n^{k,j}$  is given by (2.1). It turns out  $\mathcal{D}^{\mathcal{Y}}X = H$  for every structure  $\mathcal{Y}$  (up to a stability property) associated with  $X$ . Then we shall define  $\mathcal{D}X = \mathcal{D}^{\mathcal{Y}}X$  for every stable imbedded discrete structure  $\mathcal{Y}$  [see Theorem 4.1 and (4.9)].

The differential structure of  $X$  heavily depends on possibly orthogonal variations (1.2) of  $X$  w.r.t. noise and this is encoded by the drift term  $V$  in (1.3) as demonstrated by Proposition 4.2. Theorem 4.3 and Section 4.2 show that distinct classes of processes are clearly distinguished by (1.2) as  $k \rightarrow +\infty$  which allows us to investigate variational properties of Wiener functionals very concretely in applications to control theory and much beyond semimartingales. Theorem 4.2 and Section 4.2 show the methodology developed in this article applies to very irregular drifts of unbounded variation and it covers, in particular, drifts of finite  $p$ -variation ( $1 \leq p < 2$ ).

In [34], the authors present a universal variational characterization of the drifts of weakly differentiable processes. It is revealed the asymptotic behavior of (1.2) is always encoded by suitable limits of integral functionals of horizontal-type perturbations and first-order variation w.r.t. noise having a two-parameter occupation time process  $(s, x) \mapsto \mathbb{L}^{k,x}(s)$  for  $\mathcal{D}$  as integrators. The connection between weak differentiability and Brownian local-times are established under finite  $(p, q)$ -variation regularity in the sense of Young.

1.2. *Applications of the theory.* As a test of the relevance of our methodology, in [3, 32, 35], we apply the theory of this article to develop a concrete and systematic method of obtaining near-stochastic optimal controls in a fully non-Markovian setting

$$(1.4) \quad \arg \max_{\eta \in U_0^T} \mathbb{E}[\xi(X^\eta)],$$

where  $\xi : C([0, T]; \mathbb{R}^n) \rightarrow \mathbb{R}$  is a payoff functional and  $\{X^\eta; \eta \in U_0^T\}$  is a family of abstract Wiener functionals parameterized by a set of possibly mutually singular measures  $U_0^T$ . In [32, 35], the authors present a concrete method for computing near-stochastic optimal controls for non-Markovian systems of path-dependent

SDEs driven by Gaussian noises (including fractional Brownian motion), where in [35], both drift and diffusion components are controlled. Monte Carlo methods are developed in [3] for the particular case of optimal stopping problems. Rather than developing representation results for value functionals, the variational theory developed in this article permits to extract near-optimal controls in (1.4) without requiring a priori regularity assumptions on value processes driven by controlled Wiener functionals  $\{X^\eta; \eta \in U_0^T\}$ . The analysis is made via a maximization procedure based on  $U^{\mathcal{Y},k}X$  (or its nonlinear version) for a choice of imbedded structure  $\mathcal{Y}$  of the controlled state. In [32], the differential operator (1.2) (a nonlinear version in [35]) plays the role of a generalized Hamiltonian which yields a feasible construction of near-optimal controls beyond the Markovian case.

In the present article, we also exhibit a novel characterization of solutions of backward SDEs (BSDEs) as solutions of a variational problem w.r.t. Brownian motion state. In Theorem 4.4, it is shown that an Itô process  $Y$  is a solution of a BSDE (4.35) with terminal condition  $\xi$  and driver  $g$  if, and only if,

$$(1.5) \quad DY \text{ minimizes energy in the sense of (4.38) and } Y(T) = \xi.$$

Although the connection between BSDEs and control problems of the form (1.4) is well known (see, e.g., [6]), to our best knowledge, Theorem 4.4 is the first result connecting BSDEs to variational problems w.r.t. the Brownian motion driving noise.

The setup of this article is based on a fixed probability measure. The fully nonlinear case is partially treated in [35], where the authors aggregate imbedded discrete structures parameterized by mutually singular measures arising from non-Markovian stochastic control problems of the form (1.4). However, a full-fledged theory for abstract functionals defined up to polar sets is postponed to further investigations.

The remainder of this article is organized as follows. The next section summarizes some useful notation used in this work. Section 2 provides the discrete structure that lays the foundation of this work. Section 3 presents the differential system associated with the skeleton. Section 4 develops the asymptotic limits of the differential operators presented in Section 3 and some examples are discussed.

1.3. *Notation.* Throughout this article, we are going to fix a filtered probability space  $(\Omega, \mathbb{F}, \mathbb{P})$  which supports a  $d$ -dimensional Brownian motion  $B = (B^1, \dots, B^d)$  where  $\mathbb{F} := (\mathcal{F}_t)_{t \geq 0}$  is the usual  $\mathbb{P}$ -augmentation of the filtration generated by  $B$  under a fixed probability measure  $\mathbb{P}$ . For a given terminal time  $0 < T < \infty$ , let  $\mathbf{B}^p(\mathbb{F})$  be the Banach space of all  $\mathbb{F}$ -adapted real-valued càdlàg processes  $X$  such that

$$\mathbb{E} \sup_{0 \leq t \leq T} |X(t)|^p < \infty,$$

where  $1 \leq p < \infty$ . We denote  $\mathbf{H}^p(\mathbb{F})$  as the subset of  $\mathbf{B}^p(\mathbb{F})$  which consists of all  $\mathbb{F}$ -martingales starting from zero. We shall also equip  $\mathbf{B}^p(\mathbb{F})$  with the weak topology  $\sigma(\mathbf{B}^p, M^q)$  where  $1 \leq p, q < \infty$  with  $\frac{1}{p} + \frac{1}{q} = 1$ . Recall that the topological dual  $M^q(\mathbb{F})$  of  $\mathbf{B}^p(\mathbb{F})$  is the space of processes  $V = (V^{pr}, V^{pd})$  such that:

(i)  $V^{pr}$  and  $V^{pd}$  are right-continuous of bounded variation such that  $V^{pr}$  is  $\mathbb{F}$ -predictable with  $V_0^{pr} = 0$  and  $V^{pd}$  is  $\mathbb{F}$ -optional and purely discontinuous.

(ii)  $\text{Var}(V^{pd}) + \text{Var}(V^{pr}) \in L^q(\mathbb{P}); \frac{1}{p} + \frac{1}{q} = 1,$

where  $\text{Var}(\cdot)$  denotes the total variation of a bounded variation process over the interval  $[0, T]$ . The space  $M^q(\mathbb{F})$  has the strong topology given by

$$\|V\|_{M^q} := \|\text{Var}(V^{pr})\|_{L^q} + \|\text{Var}(V^{pd})\|_{L^q}.$$

The duality pair is given by

$$(V, X) := \mathbb{E} \int_0^T X(s-) dV^{pr}(s) + \mathbb{E} \int_0^T X(s) dV^{pd}(s); \quad X \in \mathbf{B}^p(\mathbb{F}),$$

where the following estimate holds:

$$|(V, X)| \leq \|V\|_{M^q} \|X\|_{\mathbf{B}^p}$$

for every  $V \in M^q(\mathbb{F}), X \in \mathbf{B}^p(\mathbb{F})$  such that  $1 \leq p < \infty$  and  $\frac{1}{p} + \frac{1}{q} = 1$ . We denote  $\sigma(\mathbf{B}^p, M^q)$  the weak topology of  $\mathbf{B}^p(\mathbb{F})$ . We refer the reader to, for example, [17, 30] for more details on this topology. We will see that the  $\sigma(\mathbf{B}^p, M^q)$ -topology will be quite natural for the asymptotic limits of this article.

Inequalities between random variables are understood in the  $\mathbb{P}$ -a.s. sense. Inequalities between processes are understood up to evanescent sets, otherwise, it will be up to  $Leb \times \mathbb{P}$ -null sets, where  $Leb$  is the Lebesgue measure on  $\mathbb{R}_+$ . The space of  $\mathbb{F}$ -adapted processes such that

$$\mathbb{E} \int_0^T |X(t)|^2 dt < \infty$$

is denoted by  $L_a^2(\mathbb{P} \times Leb)$ . If  $E$  is a Borel set, then we denote  $\mathcal{B}(E)$  as the Borel sigma algebra of  $E$ . The usual jump of a process is denoted by  $\Delta Y(t) := Y(t) - Y(t-)$  where  $Y(t-)$  is the left-hand limit of a càdlàg process  $Y$ . Moreover, for any two stopping times  $S$  and  $J$ , we denote the stochastic intervals  $[[S, J[[ := \{(\omega, t); S(\omega) \leq t < J(\omega)\}, [[S, S]] := \{(\omega, t); S(\omega) = t\}$  and so on.

In order to make clear the information encoded by a path  $x \in D([0, t]; \mathbb{R}^d)$  up to a given time  $0 \leq r \leq t$ , we denote  $x_r := \{x(s) : 0 \leq s \leq r\}$  and the value of  $x$  at time  $0 \leq u \leq t$  is denoted by  $x(u)$ . This notation is naturally extended to processes. Sometimes we need to stress that we are working with a functional representation in the spirit of functional calculus (see, e.g., [8, 19]). We denote  $D([0, t]; \mathbb{R}^d)$  as the linear space of  $\mathbb{R}^d$ -valued càdlàg paths on  $[0, t]$  and we set  $\Lambda := \{(t, \omega_t); (t, \omega) \in [0, T] \times D([0, T]; \mathbb{R}^d)\}$ . Similarly, we denote by  $C([0, t]; \mathbb{R}^d)$  the linear space of continuous paths on  $[0, t]$  and we set  $\hat{\Lambda} := \{(t, \omega_t); (t, \omega) \in [0, T] \times C([0, T]; \mathbb{R}^d)\}$ .

**2. The underlying discrete skeleton.** The weak functional stochastic calculus developed in this paper will be constructed from a class of pure jump processes driven by suitable waiting times which describe the local behavior of the Brownian motion: We set  $T_0^{k,j} := 0$  and

$$(2.1) \quad T_n^{k,j} := \inf\{T_{n-1}^{k,j} < t < \infty; |B^j(t) - B^j(T_{n-1}^{k,j})| = \epsilon_k\}, \quad n \geq 1,$$

where  $\sum_{k \geq 1} \epsilon_k^2 < \infty$ . For each  $j \in \{1, \dots, d\}$ , the strong Markov property yields the family  $(T_n^{k,j})_{n \geq 0}$  is a sequence of  $\mathbb{F}$ -stopping times where the increments  $\{\Delta T_n^{k,j}; n \geq 1\}$  is an i.i.d. sequence with the same distribution as  $T_1^{k,j}$ , where we set

$$\Delta T_n^{k,j} := T_n^{k,j} - T_{n-1}^{k,j}.$$

By the Brownian scaling property,  $\Delta T_1^{k,j} = \epsilon_k^2 \tau$  (in law) where  $\tau$  is an absolutely continuous variable with mean equals one and with all finite moments (see, e.g., [5]). Then we define

$$(2.2) \quad A^{k,j}(t) := \sum_{n=1}^{\infty} \epsilon_k \sigma_n^{k,j} \mathbb{1}_{\{T_n^{k,j} \leq t\}}; \quad t \geq 0, j = 1, \dots, d, k \geq 1,$$

where the size of the jumps  $\{\sigma_n^{k,j}; n \geq 1\}$  is given by

$$\sigma_n^{k,j} := \begin{cases} 1; & \text{if } \Delta A^{k,j}(T_n^{k,j}) > 0, \\ -1; & \text{if } \Delta A^{k,j}(T_n^{k,j}) < 0. \end{cases}$$

One can easily check that  $\{\sigma_n^{k,j}; n \geq 1\}$  is an i.i.d. sequence of  $\frac{1}{2}$ -Bernoulli random variables for each  $k \geq 1$  and  $j = 1, \dots, d$ . Moreover,  $\{\Delta A^{k,j}(T_n^{k,j}); n \geq 1\}$  is independent from  $\{\Delta T_n^{k,j}; n \geq 1\}$  for each  $j = 1, \dots, d$  and  $k \geq 1$ . By construction,

$$(2.3) \quad \sup_{t \geq 0} |A^{k,j}(t) - B^j(t)| \leq \epsilon_k \quad \text{a.s.}$$

for every  $k \geq 1$ .

Let  $\mathbb{F}^{k,j} := \{\mathcal{F}_t^{k,j}; t \geq 0\}$  be the natural filtration generated by  $\{A^{k,j}(t); t \geq 0\}$ . One should notice that  $\mathbb{F}^{k,j}$  is a filtration of discrete type (see Section 4 (Chapter 11) and Section 5 (Chapter 5) in [25]) in the sense that

$$\mathcal{F}_t^{k,j} = \left\{ \bigcup_{\ell=0}^{\infty} D_\ell \cap \{T_\ell^{k,j} \leq t < T_{\ell+1}^{k,j}\}; D_\ell \in \mathcal{F}_{T_\ell^{k,j}}^{k,j} \text{ for } \ell \geq 0 \right\}, \quad t \geq 0,$$

where

$$\mathcal{F}_{T_m^{k,j}}^{k,j} = \sigma(T_1^{k,j}, \dots, T_m^{k,j}, \Delta A^{k,j}(T_1^{k,j}), \dots, \Delta A^{k,j}(T_m^{k,j}))$$

and  $\mathcal{F}_0^{k,j} = \{\Omega, \emptyset\}$  and for  $m \geq 1$  and  $j = 1, \dots, d$ . From Theorem 5.56 in [25], we know that

$$\mathcal{F}_{T_m^{k,j}}^{k,j} \cap \{T_m^{k,j} \leq t < T_{m+1}^{k,j}\} = \mathcal{F}_t^{k,j} \cap \{T_m^{k,j} \leq t < T_{m+1}^{k,j}\}$$

for each  $m \geq 0$  and  $j = 1, \dots, d$ . In this case,  $\mathbb{F}^{k,j}$  is a jumping filtration in the sense of Jacod and Skorohod [26]. From Theorem 5.52 in [25],  $\mathbb{F}^{k,j}$  is right-continuous. In the sequel, we denote  $F_k$  as the distribution function of  $\Delta T_1^{k,1}$  and  $f_k := F'_k$ .

LEMMA 2.1. *The process  $A^{k,j}$  is a square-integrable  $\mathbb{F}^{k,j}$ -martingale on  $[0, T]$  and its angle bracket is*

$$(2.4) \quad \langle A^{k,j}, A^{k,j} \rangle(t) = \epsilon_k^2 \int_0^t h^{k,j}(s) ds; \quad 0 \leq t \leq T,$$

where

$$h^{k,j}(s) := \sum_{n=0}^{\infty} \frac{f_k(s - T_n^{k,j})}{1 - F_k(s - T_n^{k,j})} \mathbb{1}_{\{T_n^{k,j} < s \leq T_{n+1}^{k,j}\}}; \quad 0 \leq s \leq T$$

for  $j = 1, \dots, d$ .

PROOF. See the Supplementary Material [33].  $\square$

The multidimensional filtration generated by  $(A^{k,1}, \dots, A^{k,d})$  is naturally characterized as follows. Let  $\mathbb{F}^k := \{\mathcal{F}_t^k; 0 \leq t < \infty\}$  be the product filtration given by  $\mathcal{F}_t^k := \mathcal{F}_t^{k,1} \otimes \mathcal{F}_t^{k,2} \otimes \dots \otimes \mathcal{F}_t^{k,d}$  for  $t \geq 0$ . Let  $\mathcal{T} := \{T_m^k; m \geq 0, k \geq 1\}$  be the order statistics obtained from the family of random variables  $\{T_\ell^{k,j}; \ell \geq 0; j = 1, \dots, d\}$ . That is, we set  $T_0^k := 0$ ,

$$(2.5) \quad T_1^k := \inf_{1 \leq j \leq d} \{T_1^{k,j}\}, \quad T_n^k := \inf_{\substack{1 \leq j \leq d \\ m \geq 1}} \{T_m^{k,j}; T_m^{k,j} \geq T_{n-1}^k\}$$

for  $n \geq 1$ . We observe that the independence among the Brownian motions allows us to state that the elements of  $\{T_n^k; n \geq 0\}$  are almost surely distinct for every  $k \geq 1$ .

In the sequel, we write  $\bigvee_{i=1}^m a_i := \max\{a_1, \dots, a_m\}$  for  $(a_i)_{i=1}^m \in \mathbb{R}^m; m \geq 1$ .

LEMMA 2.2. *For every  $q \geq 1$  and  $\alpha \in (0, 1)$ , there exists a constant  $C$  which depends on  $q \geq 1$  and  $\alpha$  such that*

$$\mathbb{E} \left| \bigvee_{n \geq 1} \Delta T_n^k \right|^q \mathbb{1}_{\{T_n^k \leq T\}} \leq C(\epsilon_k^{2q} \lceil \epsilon_k^{-2} T \rceil^{(1-\alpha)})$$

for every  $k \geq 1$ .

PROOF. In the sequel, we denote  $\tau := \inf\{t \geq 0; |W(t)| = 1\}$  for a real-valued Brownian motion  $W$  and let  $f_\tau$  be the density of  $\tau$ . From Lemma 3 in [5], we know that  $f_\tau(x) = o(e^{-(\gamma-\varepsilon)x})$  as  $x \rightarrow +\infty$  for  $\gamma = \frac{\pi^2}{8}$  and any  $\varepsilon < \gamma$ . In this case, one can check there exists  $\lambda > 0$  such that the Cramer condition holds

$$\varphi(\lambda) := \mathbb{E} \exp(\lambda \tau) < \infty.$$

Let  $\psi(\lambda) := \ln \varphi(\lambda)$  defined on  $\{\lambda \in \mathbb{R}; \varphi(\lambda) < \infty\}$  and the Cramer transform is defined by  $H(a) := \sup_{\lambda < 0} [\lambda a - \psi(\lambda)]$ ;  $a < 1 = \mathbb{E}\tau$ . We recall that  $H(a) > 0$  for every  $a < 1$ . To keep notation simple, we set  $\gamma_k(t) = \lceil \epsilon_k^{-2} t \rceil$ . Let  $N^{k,j}(t) := \max\{n; T_n^{k,j} \leq t\}$ ;  $t \geq 0$  be the clock process associated with  $A^{k,j}$ , let  $\tau_n := \inf\{t > \tau_{n-1}; |W(t) - W(\tau_{n-1})| = 1\}$ ;  $n \geq 1$ ,  $\tau_0 := 0$  and  $\tau = \tau_1$ . Below  $C$  is a positive constant which may defer from line to line. By the very definition,

$$\begin{aligned} \bigvee_{n=1}^{N^{k,j}(T)} \Delta T_n^{k,j} &= \left( \bigvee_{n=1}^{N^{k,j}(T)} \Delta T_n^{k,j} \right) \mathbb{1}_{\{N^{k,j}(T) < 2\gamma_k(T)\}} \\ &\quad + \left( \bigvee_{n=1}^{N^{k,j}(T)} \Delta T_n^{k,j} \right) \mathbb{1}_{\{N^{k,j}(T) \geq 2\gamma_k(T)\}}. \end{aligned}$$

Hence,

$$\begin{aligned} \mathbb{E} \left| \bigvee_{n=1}^{N^{k,j}(T)} \Delta T_n^{k,j} \right|^q &\leq 2C \mathbb{E} \left| \bigvee_{n=1}^{2\gamma_k(T)} \Delta T_n^{k,j} \right|^q + CT^q \mathbb{P}\{N^{k,j}(T) \geq 2\gamma_k(T)\} \\ &=: I_1^k + I_2^k. \end{aligned}$$

Let us fix  $\alpha \in (0, 1)$ . Take  $n = 2\gamma_k(T)$  in Lemma 1.1 in [33] and notice that

$$(\mathbb{E}[(\Delta T_n^{k,j})^{q/(1-\alpha)}])^{1-\alpha} = \epsilon_k^{2q} (\mathbb{E}[\tau^{q/(1-\alpha)}])^{1-\alpha} =: C \epsilon_k^{2q},$$

where  $C$  is a constant depending on  $\alpha$  and  $q$ . Therefore, by applying Lemma 1.1 in [33], we have

$$I_1^k \leq 2^{1-\alpha} C \epsilon_k^{2q} \gamma_k(T)^{(1-\alpha)}; \quad k \geq 1.$$

By the scale invariance, we know that  $\Delta T_n^{k,j}$  has the same law of  $\epsilon_k^2(\tau_n - \tau_{n-1})$  for every  $n \geq 1$ . In this case,

$$\mathbb{P}\{N^{k,j}(T) \geq 2\gamma_k(T)\} = \mathbb{P}\{T_{2\gamma_k(T)}^{k,j} \leq T\} = \mathbb{P}\{\epsilon_k^2 \tau_{2\gamma_k(T)} \leq T\}.$$

In order to evaluate the second term, we notice we shall write  $\tau_{2\gamma_k(T)} = \sum_{n=1}^{2\gamma_k(T)} (\tau_n - \tau_{n-1})$  as a sum of an i.i.d. sequence with expectation equals to one. By writing,

$$\epsilon_k^2 \tau_{2\gamma_k(T)} = \epsilon_k^2 2\gamma_k(T) \frac{\tau_{2\gamma_k(T)}}{2\gamma_k(T)},$$

and noticing that  $2\epsilon_k^2 \gamma_k(T) \geq 2T$ , we shall apply classical large deviation theory to get

$$\begin{aligned} \mathbb{P}\{\epsilon_k^2 \tau_{2\gamma_k(T)} \leq T\} &= \mathbb{P}\left\{\frac{\tau_{2\gamma_k(T)}}{2\gamma_k(T)} \leq \frac{T}{\epsilon_k^2 2\gamma_k(T)}\right\} \leq \mathbb{P}\left\{\frac{\tau_{2\gamma_k(T)}}{2\gamma_k(T)} \leq \frac{1}{2}\right\} \\ &\leq \exp(-2\gamma_k(T)H(1/2)) \end{aligned}$$

for every  $k \geq 1$ . By noticing that  $\bigvee_{n=1}^\infty \Delta T_n^k \mathbb{1}_{\{T_n^k \leq T\}} \leq \bigvee_{n=1}^\infty \Delta T_n^{k,j} \mathbb{1}_{\{T_n^{k,j} \leq T\}}$  a.s. for  $j = 1, \dots, d$  and summing up the above estimates, we arrive at

$$\mathbb{E}\left|\max_{n \geq 1} \Delta T_n^k\right|^q \mathbb{1}_{\{T_n^k \leq T\}} \leq C(\epsilon^{2q} \gamma_k(T)^{(1-\alpha)} + \exp(-2\gamma_k(T)H(1/2)))$$

for every  $k \geq 1$ , where  $C$  is a constant which depends on  $\mathbb{E}|\tau|^q$ . This concludes the proof.  $\square$

LEMMA 2.3. *For each  $k \geq 1$ ,  $(T_n^k)_{n \geq 0}$  is a sequence of  $\mathbb{F}^k$ -stopping times such that  $T_n^k < \infty$  a.s. for every  $k, n \geq 1$  and  $T_n^k \uparrow +\infty$  a.s. as  $n \rightarrow \infty$ . Moreover, the filtration  $\mathbb{F}^k$  satisfies*

$$(2.6) \quad \mathcal{F}_t^k \cap \{T_n^k \leq t < T_{n+1}^k\} = \mathcal{F}_{T_n^k}^k \cap \{T_n^k \leq t < T_{n+1}^k\}; \quad t \geq 0,$$

where  $\mathcal{F}_{T_n^k}^k = \sigma(A^{k,j}(s \wedge T_n^k); s \geq 0, 1 \leq j \leq d)$  for each integer  $n \geq 0$ .

PROOF. See the Supplementary Material [33].  $\square$

Let  $\mathcal{F}_\infty^k$  be the completion of  $\sigma(A^{k,j}(s); s \geq 0; j = 1, \dots, d)$  and let  $\mathcal{N}_k$  be the  $\sigma$ -algebra generated by all  $\mathbb{P}$ -null sets in  $\mathcal{F}_\infty^k$ . With a slight abuse of notation, we write  $\mathbb{F}^k = (\mathcal{F}_t^k)_{t \geq 0}$ , where  $\mathcal{F}_t^k$  is the usual  $\mathbb{P}$ -augmentation (based on  $\mathcal{N}_k$ ) satisfying the usual conditions. From (2.3) and Lemma 2.1 in [30], we do have

$$(2.7) \quad \lim_{k \rightarrow \infty} \mathbb{F}^k = \mathbb{F}$$

weakly (in the sense of [11]) over  $[0, T]$ . Moreover, since  $\sum_{k \geq 1} \epsilon_k^2 < \infty$ , then we shall repeat the same arguments given in the proof of Lemma 2.2 in [28] to state that

$$(2.8) \quad \lim_{k \rightarrow +\infty} \sup_{0 \leq t \leq T} |T_{[\epsilon_k^{-2}t]}^{k,j} - t| = 0$$

almost surely and in  $L^2(\mathbb{P})$  for each  $j = 1, \dots, d$ .

DEFINITION 2.1. The structure  $\mathcal{D} = \{\mathcal{T}, A^{k,j}; k \geq 1, 1 \leq j \leq d\}$  is called a *discrete-type skeleton* for the Brownian motion.

For a given choice of discrete-type skeleton  $\mathcal{D}$ , we will construct a differential theory based on functionals written on  $\mathcal{D}$ . Before we proceed, it is important to point out that there exists a pathwise description of the dynamics generated by a discrete-type skeleton.

2.1. *Pathwise dynamics of the skeleton.* Let us define

$$\mathbb{I}_k := \left\{ (i_1^k, \dots, i_d^k); i_\ell^k \in \{-1, 0, 1\} \forall \ell \in \{1, \dots, d\} \text{ and } \sum_{j=1}^d |i_j^k| = 1 \right\}$$

and  $\mathbb{S}_k := (0, +\infty) \times \mathbb{I}_k$ . Let us define  $\mathfrak{N} : \mathbb{I}_k \rightarrow \{1, 2, \dots, d\} \times \{-1, 1\}$  by

$$(2.9) \quad \mathfrak{N}(\tilde{i}^k) := (\mathfrak{N}_1(\tilde{i}^k), \mathfrak{N}_2(\tilde{i}^k)) := (j, r),$$

where  $j \in \{1, \dots, d\}$  is the coordinate of  $\tilde{i}^k \in \mathbb{I}_k$  which is different from zero and  $r \in \{-1, 1\}$  is the sign of  $\tilde{i}^k$  in the coordinate  $j$ .

The  $n$ -fold Cartesian product of  $\mathbb{S}_k$  is denoted by  $\mathbb{S}_k^n$  and a generic element of  $\mathbb{S}_k^n$  will be denoted by  $\mathbf{b}_n^k := (s_1^k, \tilde{i}_1^k, \dots, s_n^k, \tilde{i}_n^k) \in \mathbb{S}_k^n$  where  $(s_r^k, \tilde{i}_r^k) \in (0, +\infty) \times \mathbb{I}_k$  for  $1 \leq r \leq n$ . Let us define  $\eta_n^k := (\eta_n^{k,1}, \dots, \eta_n^{k,d})$ , where

$$\eta_n^{k,j} := \begin{cases} 1; & \text{if } \Delta A^{k,j}(T_n^k) > 0, \\ -1; & \text{if } \Delta A^{k,j}(T_n^k) < 0, \\ 0; & \text{if } \Delta A^{k,j}(T_n^k) = 0. \end{cases}$$

Let us define

$$\mathcal{A}_n^k := (\Delta T_1^k, \eta_1^k, \dots, \Delta T_n^k, \eta_n^k) \in \mathbb{S}_k^n \quad \text{a.s.}$$

One should notice that

$$\mathcal{F}_{T_n^k}^k = (\mathcal{A}_n^k)^{-1}(\mathcal{B}(\mathbb{S}_k^n)),$$

up to null sets in  $\mathcal{F}_\infty^k$ , where  $\mathcal{B}(\mathbb{S}_k^n)$  is the Borel sigma algebra generated by  $\mathbb{S}_k^n$ ;  $n \geq 1$ .

The law of the system will evolve according to the following probability measure defined by

$$\mathbb{P}_r^k(E) := \mathbb{P}\{\mathcal{A}_r^k \in E\}; \quad E \in \mathcal{B}(\mathbb{S}_k^r)$$

for  $k, r \geq 1$ . By the very definition,

$$\mathbb{P}_n^k(\cdot) = \mathbb{P}_r^k(\cdot \times \mathbb{S}_k^{r-n})$$

for any  $r > n \geq 1$ . By construction,  $\mathbb{P}_r^k(\mathbb{S}_k^n \times \cdot)$  is a regular measure and  $\mathcal{B}(\mathbb{S}_k)$  is countably generated; then it is known (see, e.g., Chapter III, pages 70–73 in [16]) there exists ( $\mathbb{P}_n^k$ -a.s. unique) a disintegration  $\nu_{n,r}^k : \mathcal{B}(\mathbb{S}_k^{r-n}) \times \mathbb{S}_k^n \rightarrow [0, 1]$  which realizes

$$\mathbb{P}_r^k(D) = \int_{\mathbb{S}_k^n} \int_{\mathbb{S}_k^{r-n}} \mathbb{1}_D(\mathbf{b}_n^k, q_{n,r}^k) \nu_{n,r}^k(dq_{n,r}^k | \mathbf{b}_n^k) \mathbb{P}_n^k(d\mathbf{b}_n^k)$$

for every  $D \in \mathcal{B}(\mathbb{S}_k^r)$ , where  $q_{n,r}^k$  is the projection of  $\mathbf{b}_r^k = (s_1^k, \tilde{i}_1^k, \dots, s_r^k, \tilde{i}_r^k) \in \mathbb{S}_k^r$  onto the last  $(r - n)$  components, that is,  $q_{n,r}^k = (s_{n+1}^k, \tilde{i}_{n+1}^k, \dots, s_r^k, \tilde{i}_r^k)$ . If  $r = n + 1$ , we denote  $\nu_{n+1}^k := \nu_{n,n+1}^k$ . By the very definition, for each  $E \in \mathcal{B}(\mathbb{S}_k)$  and  $\mathbf{b}_n^k \in \mathbb{S}_k^n$ , we have

$$\nu_{n+1}^k(E | \mathbf{b}_n^k) = \mathbb{P}\{(\Delta T_{n+1}^k, \eta_{n+1}^k) \in E | \mathcal{A}_n^k = \mathbf{b}_n^k\}; \quad n \geq 1.$$

**3. Abstract differential skeleton.** In this section, we present a differential structure imbedded into the Brownian motion state variable based on a discrete-type structure  $\mathcal{D}$ .

DEFINITION 3.1. A *Wiener functional* is an  $\mathbb{F}$ -adapted continuous process which belongs to  $\mathbf{B}^2(\mathbb{F})$ .

DEFINITION 3.2. We say that a pure jump  $\mathbb{F}^k$ -adapted process of the form

$$X^k(t) = \sum_{n=0}^{\infty} X^k(T_n^k) \mathbb{1}_{\{T_n^k \leq t < T_{n+1}^k\}}; \quad 0 \leq t \leq T,$$

is a *good approximating sequence* (henceforth abbreviated by GAS) w.r.t.  $X$  if  $\mathbb{E}[X^k, X^k | (T) < \infty$  for every  $k \geq 1$  and

$$\lim_{k \rightarrow +\infty} X^k = X \quad \text{weakly in } \mathbf{B}^2(\mathbb{F}).$$

DEFINITION 3.3. An *imbedded discrete structure*  $\mathcal{Y} = ((X^k)_{k \geq 1}, \mathcal{D})$  for a Wiener functional  $X$  consists of the following elements:

- A discrete-type skeleton  $\mathcal{D} = \{\mathcal{T}, A^{k,j}; k \geq 1, 1 \leq j \leq d\}$  for the Brownian state  $B$ .
- A GAS  $\{X^k; k \geq 1\}$  w.r.t.  $X$  associated with the above discrete-type skeleton.

Next, we show there exists a canonical way to attach an imbedded discrete structure to an arbitrary Wiener functional  $X$ .

*Canonical imbedded discrete structure:* One typical example of an imbedded discrete structure  $\mathcal{Y} = ((\delta^k X)_{k \geq 1}, \mathcal{D})$  for a Wiener functional  $X$  is given by

$$(3.1) \quad \delta^k X(t) := \sum_{n=0}^{\infty} \mathbb{E}[X(T_n^k) | \mathcal{F}_{T_n^k}^k] \mathbb{1}_{\{T_n^k \leq t < T_{n+1}^k\}}; \quad 0 \leq t \leq T.$$

LEMMA 3.1. *If  $X$  is a Wiener functional, then  $\mathcal{Y} = ((\delta^k X)_{k \geq 1}, \mathcal{D})$  is an imbedded discrete structure for  $X$ .*

PROOF. Let us denote  $X^k := \sum_{n=0}^{\infty} X(T_n^k) \mathbb{1}_{[[T_n^k, T_{n+1}^k[ [$ . Triangle inequality yields

$$(3.2) \quad \begin{aligned} |\delta^k X(t) - X(t)| &\leq |\delta^k X(t) - X^k(t)| + |X^k(t) - X(t)| \\ &\leq \sup_{n \geq 1} |\mathbb{E}[X(T_n^k) | \mathcal{F}_{T_n^k}^k] - X(T_n^k)| \mathbb{1}_{\{T_n^k \leq t\}} \\ &\quad + \sup_{0 \leq t \leq T} |X^k(t) - X(t)| \\ &=: J_1^k + J_2^k; \quad 0 \leq t \leq T. \end{aligned}$$

By construction  $\mathbb{F}^k \subset \mathbb{F}; k \geq 1$  so that we may apply Theorem 1 in [11] to safely state that  $\lim_{k \rightarrow \infty} J_1^k = 0$  in probability. By the very definition, for a given  $\epsilon > 0$ ,

$$\left\{ \sup_{0 \leq t \leq T} |X^k(t) - X(t)| > \epsilon \right\} = \left\{ \sup_{0 \leq t \leq T} \max_{n \geq 1} |X^k(T_n^k) - X(t)| \mathbb{1}_{\{T_n^k \leq t < T_{n+1}^k\}} > \epsilon \right\}.$$

Lemma 2.2 and the path continuity of  $X$  allow us to conclude  $\lim_{k \rightarrow \infty} J_2^k = 0$  in probability. From (3.2), we shall apply Doob’s maximal inequality on the discrete-time martingale  $\mathbb{E}[\sup_{0 \leq t \leq T} |X(t)| | \mathcal{F}_{T_n^k}^k]; n \geq 1$  to get

$$\sup_{k \geq 1} \mathbb{E} \sup_{0 \leq t \leq T} |\delta^k X(t) - X(t)|^2 \leq C \mathbb{E} \sup_{0 \leq t \leq T} |X(t)|^2 < \infty$$

and hence  $\lim_{k \rightarrow \infty} \delta^k X = X$  strongly in  $\mathbf{B}^1(\mathbb{F})$ . Lastly,  $\{\delta^k X; k \geq 1\}$  is bounded in  $L^2(\Omega; E)$  where  $E$  is the Banach space of càdlàg functions from  $[0, T]$  to  $\mathbb{R}$  equipped with the sup norm. Since  $L^2(\Omega; \mathbb{R})$  is reflexive, then we shall apply Theorem 2.1 and Corollary 3.3 in [18] to state that  $\{\delta^k X; k \geq 1\} \subset \mathbf{B}^2 \subset L^2(\Omega; E)$  is weakly-relatively compact w.r.t.  $L^2(\Omega; E)$ -topology. Since  $\mathbf{B}^2 \subset L^2(\Omega; E)$  ( $\mathbf{B}^2$  been closed) and  $\lim_{k \rightarrow \infty} \delta^k X = X$  strongly in  $\mathbf{B}^1$ , then all  $\mathbf{B}^2$ -weak limit points of  $\{\delta^k X; k \geq 1\}$  are equal, and hence we do have weak convergence in  $\mathbf{B}^2$ .  $\square$

*Functional imbedded discrete structures:* Another example of an imbedded discrete structure can be constructed starting with a fixed nonanticipative functional representation. In the sequel, we make use of the following notation:

$$\omega_t := \omega(t \wedge \cdot); \quad \omega \in D([0, T]; \mathbb{R}^d).$$

This notation is naturally extended to processes. We say that  $F$  is a *nonanticipative* functional if it is a Borel mapping and

$$F_t(\omega) = F_t(\omega_t); \quad (t, \omega) \in [0, T] \times D([0, T]; \mathbb{R}^d).$$

We recall the set  $\Lambda = \{(t, \omega_t); t \in [0, T]; \omega \in D([0, T]; \mathbb{R}^d)\}$ . Let us endow  $\Lambda$  with the metric

$$\mathbf{d}((t, \omega); (t', \omega')) := \sup_{0 \leq u \leq T} \|\omega(u \wedge t) - \omega'(u \wedge t')\|_{\mathbb{R}^d} + |t - t'|.$$

Let  $X$  be a Wiener functional and let  $\mathcal{Y} = ((X^k)_{k \geq 1}, \mathcal{D})$  be an imbedded structure for  $X$ . By Doob–Dynkin lemma, there exists a functional  $\hat{F}$  defined on  $\hat{\Lambda} = \{(t, \omega_t); t \in [0, T]; \omega \in C([0, T]; \mathbb{R}^d)\}$  such that

$$(3.3) \quad \hat{F}_t(B_t) = X(t); \quad 0 \leq t \leq T.$$

When we write  $X = F(B)$  for a given nonanticipative functional  $F$  defined on  $\Lambda$ , it is implicitly assumed that we are fixing a functional  $F$  which is consistent to  $\hat{F}$  in the sense that  $F_t(x_t) = \hat{F}_t(x_t)$  for every  $x \in C([0, T]; \mathbb{R})$ . Let  $X$  be a Wiener functional

$$X(t) = F_t(B_t); \quad 0 \leq t \leq T,$$

where a  $F$  is a nonanticipative functional  $F$  defined on  $\Lambda$ . Then we shall define the following structure  $\mathcal{F} := ((\mathbf{F}^k)_{k \geq 1}, \mathcal{D})$ :

$$(3.4) \quad \mathbf{F}^k(t) := \sum_{\ell=0}^{\infty} F_{T_\ell^k}(A_{T_\ell^k}^k) \mathbb{1}_{\{T_\ell^k \leq t < T_{\ell+1}^k\}}; \quad 0 \leq t \leq T.$$

The reader should not confuse  $F(A^k)$  with  $\mathbf{F}^k$  because  $\{\mathbf{F}^k(t); 0 \leq t \leq T\}$  is a pure jump process while  $\{F_t(A_t^k); 0 \leq t \leq T\}$  does not necessarily has this property. Under continuity assumptions in the sense of pathwise functional calculus (see [8, 19]), one can easily check that  $\lim_{k \rightarrow \infty} \mathbf{F}^k = F(B)$  weakly in  $\mathbf{B}^2(\mathbb{F})$  so that  $\mathcal{F} = ((\mathbf{F}^k)_{k \geq 1}, \mathcal{D})$  is an imbedded discrete structure for the Wiener functional  $F(B)$ .

Concrete examples of imbedded discrete structures arise in many contexts:

- Discretization of value processes arising from path-dependent optimal stochastic control problems; see the works [3, 32, 35].
- Euler–Maruyama schemes arising from path-dependent stochastic differential equations driven by Gaussian noises; see the works [32, 35].
- Functional imbedded discrete structures associated with path-dependent functionals under  $(p, q)$ -variation regularity; see [34].

In principle, one can always construct an imbedded discrete structure to a Wiener functional  $X$  by only observing the basic probabilistic structure of  $X$ . The use of functional structures  $\mathcal{F}$  is only indicated when one has some a priori information on a functional which realizes  $X = F(B)$ . In general, this is not the case and other types of imbedded structures must be considered; see [32, 35] for concrete examples of applications to control theory.

*Differential operators on imbedded discrete structures:* The reader should really think an imbedded discrete structure as a model simplification for a given Wiener functional  $X$  where we are able to compute freely the sensitivities of  $X$  w.r.t. the Brownian state, that is, without any regularity assumptions. In the sequel, we provide a detailed explanation on this point. For a given Wiener functional  $X$ , let us choose an imbedded discrete structure  $\mathcal{Y} = ((X^k)_{k \geq 1}, \mathcal{D})$  associated with  $X$ .

In the sequel,  $(Y)^{p,k}$  denotes the  $\mathbb{F}^k$ -dual predictable projection of an  $\mathbb{F}^k$ -adapted process  $Y$  with locally integrable variation (see Chapter 5 in [25]). At first, we observe that the  $\mathbb{F}^k$ -dual predictable projection  $(X^k - X^k(0))^{p,k}$  of the process  $X^k - X^k(0)$  is well defined and it is the unique  $\mathbb{F}^k$ -predictable bounded variation process such that

$$X^k - X^k(0) - (X^k - X^k(0))^{p,k} \text{ is an } \mathbb{F}^k\text{-local martingale.}$$

Then one can write

$$(3.5) \quad X^k(t) = X^k(0) + \sum_{0 < s \leq t} \Delta X^k(s) = X^k(0) + \sum_{j=1}^d \int_0^t \mathcal{D}^{\mathcal{Y},k,j} X(u) dA^{k,j}(u),$$

where

$$\mathcal{D}^{\mathcal{Y},k,j} X(u) := \sum_{\ell=1}^{\infty} \frac{\Delta X^k(T_\ell^{k,j})}{\Delta A^{k,j}(T_\ell^{k,j})} \mathbb{1}_{\{T_\ell^{k,j}=u\}}; \quad 0 \leq u \leq T, \mathcal{Y} = ((X^k)_{k \geq 1}, \mathcal{D}),$$

and the integral in (3.5) is interpreted in the Lebesgue–Stieltjes’ sense. In the sequel,  $\mu_{[A^{k,j}]}$  is the Doléans’ measure (see, e.g., Chapter 5 in [25]) generated by the point process  $[A^{k,j}, A^{k,j}]; 1 \leq j \leq d, k \geq 1$ , that is,

$$\mu_{[A^{k,j}]}(H) := \mathbb{E} \int_0^T \mathbb{1}_H(s) d[A^{k,j}, A^{k,j}](s); \quad H \in \mathcal{F}_T^k \times \mathcal{B}([0, T]).$$

From the square integrability of the martingale,  $A^{k,j}, \mu_{[A^{k,j}]}$  is a finite measure. In the sequel,  $\mathcal{P}^k$  is the  $\mathbb{F}^k$ -predictable sigma-algebra of  $[0, T] \times \Omega$ . We observe that

$$\int_0^\cdot \frac{\mathcal{D}^{\mathcal{Y},k,j} X}{\Delta A^{k,j}} d[A^{k,j}, A^{k,j}]$$

is a process with locally integrable variation. Then there exists a unique (up to sets of  $\mu_{[A^{k,j}]}$ -measure zero)  $\mathbb{F}^k$ -predictable process  $\mathbb{E}_{\mu_{[A^{k,j}]}}[\frac{\mathcal{D}^{\mathcal{Y},k,j} X}{\Delta A^{k,j}} | \mathcal{P}^k]$  such that

$$\left( \int_0^\cdot \frac{\mathcal{D}^{\mathcal{Y},k,j} X}{\Delta A^{k,j}} d[A^{k,j}, A^{k,j}] \right)^{p,k} = \int_0^\cdot \mathbb{E}_{\mu_{[A^{k,j}]}} \left[ \frac{\mathcal{D}^{\mathcal{Y},k,j} X}{\Delta A^{k,j}} | \mathcal{P}^k \right] d[A^{k,j}, A^{k,j}].$$

See, for example, Theorem 5.25 and remark in [25]. We then denote

$$(3.6) \quad U^{\mathcal{Y},k,j} X(u) := \mathbb{E}_{\mu_{[A^{k,j}]}} \left[ \frac{\mathcal{D}^{\mathcal{Y},k,j} X}{\Delta A^{k,j}} | \mathcal{P}^k \right](u); \quad 0 \leq u \leq T, k \geq 1, 1 \leq j \leq d,$$

where it is understood that the stochastic process  $\mathcal{D}^{\mathcal{Y},k,j} X / \Delta A^{k,j}$  is null on the complement of the union of stochastic intervals  $\bigcup_{n=1}^\infty [[T_n^{k,j}, T_n^{k,j}]]$ . We call

$$(3.7) \quad \sum_{j=1}^d U^{\mathcal{Y},k,j} X$$

as the  $\mathbb{F}^k$ -weak infinitesimal generator of the imbedded discrete structure  $\mathcal{Y}$  w.r.t.  $X$ . We then denote  $\mathcal{D}^{\mathcal{Y},k} X = (\mathcal{D}^{\mathcal{Y},k,1} X, \dots, \mathcal{D}^{\mathcal{Y},k,d} X)$  and  $U^{\mathcal{Y},k} X = (U^{\mathcal{Y},k,1} X, \dots, U^{\mathcal{Y},k,d} X)$ .

LEMMA 3.2. *Let  $\mathcal{Y} = ((X^k)_{k \geq 1}, \mathcal{D})$  be an imbedded discrete structure w.r.t.  $X$ . The  $\mathbb{F}^k$ -dual predictable projection of  $X^k - X^k(0)$  is given by the absolutely continuous process*

$$\sum_{j=1}^d \int_0^t U^{\mathcal{Y},k,j} X(s) d[A^{k,j}, A^{k,j}](s); \quad 0 \leq t \leq T.$$

Moreover,

$$(3.8) \quad \sum_{j=1}^d U^{\mathcal{Y},k,j} X(T_{n+1}^k) = \mathbb{E} \left[ \frac{\Delta X^k(T_{n+1}^k)}{\epsilon_k^2} \middle| \mathcal{F}_{T_{n+1}^k} \right] \quad a.s.$$

for each  $n \geq 0$  and  $k \geq 1$ .

PROOF. See the Supplementary Material [33].  $\square$

For a given embedded discrete structure  $\mathcal{Y} = ((X^k)_{k \geq 1}, \mathcal{D})$  w.r.t.  $X$ , let us denote

$$\begin{aligned} \sum_{j=1}^d \oint_0^t \mathcal{D}^{\mathcal{Y},k,j} X(s) dA^{k,j}(s) &:= \sum_{j=1}^d \int_0^t \mathcal{D}^{\mathcal{Y},k,j} X(s) dA^{k,j}(s) \\ &\quad - \left( \sum_{j=1}^d \int_0^t \mathcal{D}^{\mathcal{Y},k,j} X(s) dA^{k,j}(s) \right)^{p,k}(t), \end{aligned}$$

where  $\oint$  is the  $\mathbb{F}^k$ -optional integral as introduced by Dellacherie and Meyer (see Chapter 8, Section 2 in [17]) for optional integrands. Let us denote

$$(3.9) \quad \begin{aligned} \mathbb{D}^{\mathcal{Y},k,j} X(t) &:= \sum_{\ell=1}^{\infty} \mathcal{D}^{\mathcal{Y},k,j} X(t) \mathbb{1}_{\{T_{\ell}^k \leq t < T_{\ell+1}^k\}}, \\ \mathbb{U}^{\mathcal{Y},k,j} X(t) &:= U^{\mathcal{Y},k,j} X(t) \frac{d\langle A^{k,j}, A^{k,j} \rangle}{dt} \end{aligned}$$

for  $0 \leq t \leq T$ . We then arrive at the following result.

PROPOSITION 3.1. *Let  $\mathcal{Y} = ((X^k)_{k \geq 1}, \mathcal{D})$  be an imbedded discrete structure for a Wiener functional  $X$ . Then the  $\mathbb{F}^k$ -special semimartingale decomposition of  $X^k$  is given by*

$$(3.10) \quad \begin{aligned} X^k(t) &= X^k(0) + \sum_{j=1}^d \oint_0^t \mathbb{D}^{\mathcal{Y},k,j} X(s) dA^{k,j}(s) \\ &\quad + \sum_{j=1}^d \int_0^t \mathbb{U}^{\mathcal{Y},k,j} X(s) ds; \quad 0 \leq t \leq T. \end{aligned}$$

At this point, we stress that in many cases playing with the variational operators  $(\mathcal{D}^{\mathcal{Y},k,j} X, U^{\mathcal{Y},k,j} X)$  attached to an imbedded discrete structure is good enough to infer nontrivial information on  $X$  without computing the limit of  $(\mathbb{D}^{\mathcal{Y},k,j} X, \mathbb{U}^{\mathcal{Y},k,j} X)$ , which in many cases, it can be problematic due to lack of smoothness of  $X$  w.r.t. state; see [32, 35] for details. The differential form (3.10)

describes an imbedded discrete structure for a Wiener functional  $X$  and it will be the starting point to analyze the sensitivities of  $X$  w.r.t. Brownian state under rather weak regularity conditions as demonstrated in Section 4.

*Pathwise description of the variational operators attached to  $\mathcal{Y}$ :* In order to compute the preliminary variational operators  $(\mathcal{D}^{\mathcal{Y},k,j} X, U^{\mathcal{Y},k,j} X)$ , we can actually proceed path wisely because all the objects are path-dependent functionals of the discrete-type skeleton  $\mathcal{D}$ . In the sequel, for any  $(r, n)$  such that  $1 \leq r \leq n$  and  $\mathbf{b}_n^k = (s_1^k, \tilde{i}_1^k, \dots, s_n^k, \tilde{i}_n^k)$ , we denote

$$\pi_r(\mathbf{b}_n^k) := (s_1^k, \tilde{i}_1^k, \dots, s_r^k, \tilde{i}_r^k).$$

Let  $F_n^k : \mathbb{S}_k^n \rightarrow \mathbb{R}$  be a sequence of Borel functions. Let us define

$$\nabla_j F^k(\mathbf{b}_n^k) := \frac{F_n^k(\mathbf{b}_n^k) - F_{n-1}^k(\pi_{n-1}(\mathbf{b}_n^k))}{\epsilon_k \mathfrak{N}_2(\tilde{i}_n^k)} \mathbb{1}_{\{\mathfrak{N}_1(\mathbf{b}_n^k)=j\}}; \quad \mathbf{b}_n^k \in \mathbb{S}_k^n, n \geq 1$$

for  $j = 1, \dots, d$ , and

$$(3.11) \quad \mathcal{U} F^k(\mathbf{b}_n^k) := \int_{\mathbb{S}_k} \frac{F_{n+1}^k(\mathbf{b}_n^k, s_{n+1}^k, \tilde{i}_{n+1}^k) - F_n^k(\mathbf{b}_n^k)}{\epsilon_k^2} \nu_{n+1}^k(ds_{n+1}^k d\tilde{i}_{n+1}^k | \mathbf{b}_n^k)$$

for  $\mathbf{b}_n^k \in \mathbb{S}_k^n, n \geq 0$ . We observe that for a sequence of Borel maps  $F_n^k : \mathbb{S}_k^n \rightarrow \mathbb{R}$  satisfying

$$X^k(T_n^k) = F_n^k(\mathcal{A}_n^k) \quad \text{a.s.}$$

we will get

$$\nabla_j F^k(\mathcal{A}_n^k) = \mathcal{D}^{\mathcal{Y},k,j} X(T_n^k) \mathbb{1}_{\{\mathfrak{N}_1(\eta_n^k)=j\}} \quad \text{a.s., } n \geq 0, j = 1, \dots, d,$$

and

$$(3.12) \quad \mathcal{U} F^k(\mathcal{A}_n^k) = \mathbb{E} \left[ \sum_{j=1}^d U^{\mathcal{Y},k,j} X(T_{n+1}^k) \middle| \mathcal{F}_{T_n^k}^k \right] \quad \text{a.s.}$$

for each  $n \geq 0$ . The right-hand side of (3.12) will be called the  $\mathbb{F}^k$ -conditional weak infinitesimal generator of  $X$  based on a structure  $\mathcal{Y}$ . It turns out that the operator (3.11) play the role of the Hamiltonian in a non-Markovian optimal stopping problem (see [32] and Example 3.2). In the sequel, we present two examples related to martingales and stochastic control based on a generic Wiener functional.

**EXAMPLE 3.1.** Let  $X \in \mathbf{H}^2(\mathbb{F})$  be a martingale with terminal condition  $\xi \in L^2(\mathcal{F}_T)$  and  $\mathbb{F}$  is the filtration generated by a one-dimensional Brownian motion. A simple imbedded structure associated with  $X$  is the following: Let  $\xi^k := G_{\gamma(k,T)}^k(\mathcal{A}_{\gamma(k,T)}^k)$  be an approximation for  $\xi$ , that is,  $\lim_{k \rightarrow +\infty} \xi^k = \xi$  in  $L^2(\mathbb{P})$ , where  $\gamma(k, T) = \lceil \epsilon_k^{-2} T \rceil$ . Let us define a good approximation sequence by

$$X^k(T_n^k) := \mathbb{E}[\xi^k | \mathcal{F}_{T_n^k}^k]; \quad 0 \leq n \leq \gamma(k, T).$$

Due to the path continuity of  $X$ , the weak convergence  $\mathbb{F}^k \rightarrow \mathbb{F}$ , Lemma 2.2 and (2.8), one can easily check that  $((X^k)_{k \geq 1}, \mathcal{D})$  is an imbedded discrete structure. A simple computation reveals that  $X^k(T_n^k) = F_n^k(\mathcal{A}_n^k)$  a.s. where

$$F_r^k(\mathbf{b}_r^k) := G_r^k(\mathbf{b}_r^k); \quad \mathbf{b}_r^k \in \mathbb{S}_k^r,$$

$$F_n^k(\mathbf{b}_n^k) = \int_{\mathbb{S}^{r-n}} G_r^k(\mathbf{b}_n^k, q_{n,r}^k) v_{n,r}^k(dq_{n,r}^k | \mathbf{b}_n^k); \quad \mathbf{b}_n^k \in \mathbb{S}_k^n, 0 \leq n < r,$$

where  $r$  is a shorthand notation for  $\gamma(k, T)$ . Then

$$\nabla F_n^k(\mathbf{b}_n^k) = \frac{F_n^k(\mathbf{b}_n^k) - F_{n-1}^k(\pi_{n-1}(\mathbf{b}_n^k))}{\epsilon_k \tilde{i}_n^k}; \quad \mathbf{b}_n^k \in \mathbb{S}_k^n, 1 \leq n \leq r,$$

and

$$\mathcal{U} F_n^k(\mathbf{b}_n^k) = \frac{1}{2} \int_0^\infty \frac{F_{n+1}^k(\mathbf{b}_n^k, x, 1) - F_n^k(\mathbf{b}_n^k)}{\epsilon_k^2} f_k(x) dx$$

$$+ \frac{1}{2} \int_0^\infty \frac{F_{n+1}^k(\mathbf{b}_n^k, x, -1) - F_n^k(\mathbf{b}_n^k)}{\epsilon_k^2} f_k(x) dx$$

for  $\mathbf{b}_n^k \in \mathbb{S}_k^n, 0 \leq n \leq r$ .

EXAMPLE 3.2. Let us illustrate the case of the optimal stopping problem. In [3, 32], the authors propose a systematic way in concretely solving optimal stopping problems based on generic Wiener functionals beyond the Markovian case. Let  $S$  be the Snell supermartingale process

$$(3.13) \quad S(t) := \operatorname{ess\,sup}_{\tau \geq t} \mathbb{E}[Z(\tau) | \mathcal{F}_t], \quad 0 \leq t \leq T,$$

where  $\operatorname{ess\,sup}$  is computed over the class of all  $\mathbb{F}$ -stopping times located on  $[t, T]$ . It is shown that value processes of the form (3.13) based on reward continuous functionals  $Z$  applied to path-dependent SDEs driven by fractional Brownian motion admit imbedded discrete structures  $\mathcal{Y} = ((S^k)_{k \geq 1}, \mathcal{D})$  with pathwise representations  $S^k(T_n^k) = \mathbb{V}_n^k(\mathcal{A}_n^k)$  a.s.;  $0 \leq n < r$  where  $\mathbb{V}^k$  solves the nonlinear equation

$$\max\{\mathcal{U} \mathbb{V}_i^k(\mathbf{b}_i^k); \gamma_{i,r}^k(\mathbf{b}_i^k) - \mathbb{V}_i^k(\mathbf{b}_i^k)\} = 0; \quad i = r - 1, \dots, 0,$$

$$\mathbb{V}_r^k(\mathbf{b}_r^k) = \gamma_{r,r}^k(\mathbf{b}_r^k); \quad \mathbf{b}_r^k \in \mathbb{S}_k^r,$$

where  $\gamma_{n,r}^k$  in an explicit pathwise representation for a simple imbedded discrete structure  $\mathcal{Z} = ((Z^k)_{k \geq 1}, \mathcal{D})$  associated with  $Z$ .

**4. Differential structure of Wiener functionals.** In this section, we present asymptotic results for the sensitivities associated with imbedded discrete structures as defined in the previous section.

**DEFINITION 4.1.** Let  $\mathcal{Y} = ((X^k)_{k \geq 1}; \mathcal{D})$  be an imbedded discrete structure for a Wiener functional  $X$ . We say that  $\mathcal{Y}$  has *finite energy* if

$$\mathcal{E}^{2,\mathcal{Y}}(X) := \sup_{k \geq 1} \mathbb{E} \sum_{n \geq 1} |\Delta X^k(T_n^k)|^2 \mathbb{1}_{\{T_n^k \leq T\}} < \infty.$$

**DEFINITION 4.2.** Let  $\mathcal{Y} = ((X^k)_{k \geq 1}; \mathcal{D})$  be an imbedded discrete structure for a Wiener functional  $X$ . We say that  $X$  admits the  $\mathcal{Y}$ -covariation w.r.t. to  $j$ th component of the Brownian motion  $B$  if the limit

$$\langle X, B^j \rangle^{\mathcal{Y}}(t) := \lim_{k \rightarrow \infty} [X^k, A^{k,j}](t)$$

exists weakly in  $L^1(\mathbb{P})$  for each  $t \in [0, T]$ .

The idea behind Definition 4.1 is compactness for an imbedded discrete structure which will allow us to extract convergent subsequences on the components of the special semimartingale decomposition given in Proposition 3.1. The role of the  $\mathcal{Y}$ -covariation is to bring stability to the semimartingale decomposition in Proposition 3.1.

It is important to point out that the above properties are only important to get convergence of the underlying differential structure  $(\mathbb{D}^{\mathcal{Y},k,j} X, \mathbb{U}^{\mathcal{Y},k,j} X; j = 1, \dots, d)$  but in typical applications, we do not really need the existence of the limiting differential form to solve concrete problems in, for example, non-Markovian stochastic control; see [32, 35] for all the details.

From Proposition 3.1, we know that each imbedded discrete structure  $\mathcal{Y} = ((X^k)_{k \geq 1}; \mathcal{D})$  carries a sequence of  $\mathbb{F}^k$ -special semimartingale decompositions

$$X^k(t) = X^k(0) + M^{\mathcal{Y},k}(t) + N^{\mathcal{Y},k}(t); \quad 0 \leq t \leq T,$$

where  $M^{\mathcal{Y},k}$  is an  $\mathbb{F}^k$ -square-integrable martingale and  $N^{\mathcal{Y},k}$  is an  $\mathbb{F}^k$ -predictable absolutely continuous process.

**DEFINITION 4.3.** Let  $X = X(0) + M + N$  be an  $\mathbb{F}$ -adapted process such that  $M \in \mathbf{H}^2(\mathbb{F})$  and  $N \in \mathbf{B}^2(\mathbb{F})$  has continuous paths. An imbedded discrete structure  $\mathcal{Y} = ((X^k)_{k \geq 1}; \mathcal{D})$  for  $X$  is said to be *stable* if  $M^{\mathcal{Y},k} \rightarrow M$  weakly in  $\mathbf{B}^2(\mathbb{F})$  as  $k \rightarrow +\infty$ .

**REMARK 4.1.** By the Burkholder–Davis–Gundy’s inequality and the fact that  $\Delta X^k = \Delta M^{\mathcal{Y},k}$ , we have  $\mathcal{E}^{2,\mathcal{Y}}(X) < \infty$  for any stable discrete structure  $\mathcal{Y}$  w.r.t.  $X \in \mathbf{B}^2(\mathbb{F})$ .

Going to the literature on convergence of stochastic processes, we can find some compactness conditions to ensure that a given imbedded structure is stable.

LEMMA 4.1. *Let  $X = X(0) + M + N$  be a square-integrable continuous semimartingale where  $M \in \mathbf{H}^2(\mathbb{F})$  and  $N$  is  $\mathbb{F}$ -adapted with bounded variation. If  $\mathcal{Y} = ((X^k)_{k \geq 1}, \mathcal{D})$  is an imbedded discrete structure w.r.t.  $X$  such that  $\mathcal{E}^{2,\mathcal{Y}}(X) < \infty$ ,  $\lim_{k \rightarrow +\infty} X^k = X$  in  $\mathbf{B}^2(\mathbb{F})$  and*

$$\text{Var}(N^{\mathcal{Y},k}) \text{ is tight in } \mathbb{R},$$

*then  $\mathcal{Y}$  is stable, where  $\text{Var}$  denotes the first variation of a process over  $[0, T]$ . Let  $X$  be a continuous strong Dirichlet process (see Section 4.1) with canonical decomposition  $X = X(0) + M + N$ ;  $M \in \mathbf{H}^2(\mathbb{F})$  and  $N$  has 2-null variation. Let  $\mathcal{Y} = ((X^k)_{k \geq 1}, \mathcal{D})$  be a structure associated with  $X$ . If for every  $\epsilon > 0$*

$$\lim_{k \rightarrow +\infty} \sup_{\ell \geq 1} \mathbb{P} \left( \sum_{r=1}^{\infty} |N^{\mathcal{Y},\ell}(T_r^k) - N^{\mathcal{Y},\ell}(T_{r-1}^k)|^2 \mathbb{1}_{\{T_r^k \leq T\}} > \epsilon \right) = 0,$$

*$\sup_{0 \leq t \leq T} |N^{\mathcal{Y},k}(t)|$  is bounded in probability,  $\lim_{k \rightarrow +\infty} X^k = X$  in  $\mathbf{B}^2(\mathbb{F})$  and finite energy  $\mathcal{E}^{2,\mathcal{Y}}(X) < \infty$  holds, then  $\mathcal{Y}$  is stable.*

PROOF. For the semimartingale case, we just need to apply Theorem 11 in [36] to conclude  $M^{\mathcal{Y},k} \rightarrow M$  uniformly in probability as  $k \rightarrow +\infty$  and this implies  $M^{\mathcal{Y},k} \rightarrow M$  weakly in  $\mathbf{B}^2(\mathbb{F})$  as  $k \rightarrow +\infty$  so that  $\mathcal{Y}$  is stable. Theorem 2 in [12] allows us to conclude stability for the Dirichlet case up to the fact that our partition is random, but one can easily check that all arguments in the proof of Theorem 2 in [12] apply to our case as well.  $\square$

Stability of an imbedded discrete structure is a natural property since we are interested in analyzing Wiener functionals with at least a nonnull martingale component; see also Remark 4.2 for further details. Let us now devote our attention to the study of the asymptotic properties of

$$\mathbb{D}^{\mathcal{Y},k,j} X; \quad j = 1, \dots, d$$

for an imbedded discrete structure  $\mathcal{Y} = ((X^k)_{k \geq 1}, \mathcal{D})$  w.r.t.  $X$ . We set

$$(4.1) \quad \mathcal{D}_j^{\mathcal{Y}} X := \lim_{k \rightarrow +\infty} \mathbb{D}^{\mathcal{Y},k,j} X \quad \text{weakly in } L_a^2(\mathbb{P} \times \text{Leb})$$

whenever the right-hand side of (4.1) exists for a given finite-energy embedded structure  $\mathcal{Y}$  and, in this case, we write

$$\mathcal{D}^{\mathcal{Y}} X := (\mathcal{D}_1^{\mathcal{Y}} X, \dots, \mathcal{D}_d^{\mathcal{Y}} X).$$

We start the analysis with the following result.

**THEOREM 4.1.** *Let  $X = X(0) + \sum_{j=1}^d \int H_j dB^j + V$  be an  $\mathbb{F}$ -adapted process such that  $H = (H_1, \dots, H_d)$ ,  $H_j \in L^2_a(\mathbb{P} \times \text{Leb})$ ;  $j = 1, \dots, d$  and  $V \in \mathbf{B}^2(\mathbb{F})$  has continuous paths. Then*

$$\mathcal{D}^{\mathcal{Y}} X = H$$

for every stable imbedded discrete structure  $\mathcal{Y} = ((X^k)_{k \geq 1}, \mathcal{D})$ .

At first, we observe Lemma 3.4 in [30] holds for the  $\mathbb{F}^k$ -martingale  $A^{k,j}$ . Then we have the following result.

**LEMMA 4.2.** *Let  $H = \mathbb{E}[\mathbb{1}_G | \mathcal{F}_\cdot]$  and  $H^k = \mathbb{E}[\mathbb{1}_G | \mathcal{F}^k_\cdot]$  be positive and uniformly integrable martingales w.r.t. filtrations  $\mathbb{F}$  and  $\mathbb{F}^k$ , respectively, where  $G \in \mathcal{F}_T$ . Then, for each  $1 \leq j \leq d$ ,*

$$\left\| \int_0^\cdot H(s) dB_s^j - \int_0^\cdot H^k(s) dA^{k,j}(s) \right\|_{\mathbf{B}^2} \rightarrow 0 \quad \text{as } k \rightarrow \infty.$$

**PROOF.** Since  $A^{k,j}$  is a pure jump martingale and  $\mathbb{E} \sup_{0 \leq t \leq T} |B^j(t)|^p < \infty$  for every  $p > 2$ , then we shall apply Lemma 3.4 in [30] to conclude the proof.  $\square$

**LEMMA 4.3.** *Let  $\{Y^k; k \geq 1\}$  be a sequence of  $\mathbb{F}^k$ -square-integrable martingales such that  $\lim_{k \rightarrow \infty} Y^k = Z$  weakly in  $\mathbf{B}^2(\mathbb{F})$ , where  $Z \in \mathbf{H}^2(\mathbb{F})$ . Then, for each  $1 \leq j \leq d$ ,*

$$(4.2) \quad \lim_{k \rightarrow \infty} [Y^k, A^{k,j}](t) = [Z, B^j](t) \quad \text{weakly in } L^1(\mathbb{P})$$

for every  $t \in [0, T]$ .

**PROOF.** In the notation of the proof of Lemma 3.5 in [30], we observe that since for every BMO  $\mathbb{F}$ -martingale  $U$ , we have  $\lim_{k \rightarrow \infty} [Z^{k,X}, U](t) = [Z, U](t)$  weakly in  $L^1(\mathbb{P})$  for every  $t \in [0, T]$ , then we shall take  $W = B$ . By using Lemma 4.2, the proof of Lemma 3.5 in [30] works perfectly for the pure-jump sequence  $\{Y^k; k \geq 1\}$  which allows us to conclude that (4.2) holds true.  $\square$

An immediate consequence of Lemma 4.3 is the following result.

**COROLLARY 4.1.** *Let  $\mathcal{Y} = ((X^k)_{k \geq 1}; \mathcal{D})$  be an imbedded discrete structure for a Wiener functional  $X = X(0) + Y + V$  where  $Y \in \mathbf{H}^2(\mathbb{F})$  and  $V \in \mathbf{B}^2(\mathbb{F})$  is a continuous process. If  $\mathcal{Y}$  is stable, then  $\langle X, B^j \rangle^{\mathcal{Y}} = [Y, B^j]$  exists for  $j = 1, \dots, d$  and  $\mathcal{E}^{2,\mathcal{Y}}(X) < \infty$ .*

**PROOF.** Let  $X^k = X^k(0) + M^{\mathcal{Y},k} + N^{\mathcal{Y},k}$  be the  $\mathbb{F}^k$ -special semimartingale decomposition of  $X^k$ . If  $\lim_{k \rightarrow +\infty} M^{\mathcal{Y},k}$  converges weakly in  $\mathbf{B}^2(\mathbb{F})$ , then it is bounded in the strong norm of  $\mathbf{B}^2(\mathbb{F})$  and Burkholder–Davis–Gundy’s inequality

yields that  $\sup_{k \geq 1} \mathbb{E}[M^{\mathcal{Y},k}, M^{\mathcal{Y},k}](T) = \mathcal{E}^{2,\mathcal{Y}}(X) < \infty$ . By applying Lemma 4.3 jointly with the predictable martingale representation theorem of the Brownian motion, we conclude  $\langle X, B^j \rangle^{\mathcal{Y}} = [Y, B^j]$  for  $j = 1, \dots, d$ .  $\square$

We are now ready to prove Theorem 4.1.

**PROOF OF THEOREM 4.1.** Throughout this proof,  $C$  is a constant which may defer from line to line. In the sequel, we fix an arbitrary stable imbedded discrete structure  $\mathcal{Y} = ((X^k)_{k \geq 1}, \mathcal{D})$  associated with  $X \in \mathbf{B}^2(\mathbb{F})$  having a representation

$$X = X(0) + M + N,$$

where  $M = \sum_{j=1}^d \int H_j dB^j \in \mathbf{H}^2(\mathbb{F})$  and  $N \in \mathbf{B}^2(\mathbb{F})$  is a continuous process. We claim that

$$\mathcal{D}^{\mathcal{Y}} X = H.$$

We start by observing that the stability of  $\mathcal{Y}$  implies that

$$M^{\mathcal{Y},k} = \sum_{j=1}^d \oint_0^{\cdot} \mathbb{D}^{\mathcal{Y},k,j} X dA^{k,j}(s); \quad k \geq 1$$

is a bounded in  $\mathbf{B}^2(\mathbb{F})$  which is equivalent to  $\sup_{k \geq 1} \mathbb{E}[M^{\mathcal{Y},k}, M^{\mathcal{Y},k}](T) < \infty$ . Moreover,

$$\begin{aligned} & \mathbb{E} \int_0^T \|\mathbb{D}^{\mathcal{Y},k} X(s)\|_{\mathbb{R}^d}^2 ds \\ (4.3) \quad &= \mathbb{E}[M^{\mathcal{Y},k}, M^{\mathcal{Y},k}](T) \\ & - \mathbb{E} \sum_{j=1}^d \sum_{n=1}^{\infty} |\mathbb{D}^{\mathcal{Y},k,j} X(T_n^{k,j})|^2 (T_{n+1}^{k,j} - T_n^{k,j}) \mathbb{1}_{\{T_n^{k,j} \leq T < T_{n+1}^{k,j}\}} \end{aligned}$$

for every  $k \geq 1$ . Therefore,  $\sup_{k \geq 1} \mathbb{E} \int_0^T \|\mathbb{D}^{\mathcal{Y},k} X(s)\|_{\mathbb{R}^d}^2 ds < \infty$ .

Let us fix  $g \in L^\infty$ ,  $t \in [0, T]$  and  $j = 1, \dots, d$ . By the very definition, we have

$$\begin{aligned} \mathbb{E} g \int_0^t \mathbb{D}^{\mathcal{Y},k,j} X(s) ds &= \mathbb{E} g \sum_{n=1}^{\infty} \mathbb{D}^{\mathcal{Y},k,j} X(T_{n-1}^{k,j}) \Delta T_n^{k,j} \mathbb{1}_{\{T_{n-1}^{k,j} \leq t\}} \\ & - \mathbb{E} g \sum_{n=1}^{\infty} \mathbb{D}^{\mathcal{Y},k,j} X(T_{n-1}^{k,j}) (T_n^{k,j} - t) \mathbb{1}_{\{T_{n-1}^{k,j} < t \leq T_n^{k,j}\}} \\ &=: I^{k,j,1}(t) + I^{k,j,2}(t). \end{aligned}$$

By the strong Markov property,  $\Delta T_n^{k,j}$  is independent from  $\mathcal{F}_{T_{n-1}^{k,j}}^k$  and the Brownian scaling yields  $\mathbb{E} \Delta T_n^{k,j} = \epsilon_k^2$  for every  $n \geq 1$ . Then we shall estimate

$$|I^{k,j,2}(t)| \leq C \epsilon_k \sum_{n=1}^{\infty} \mathbb{E} |\Delta X^k(T_{n-1}^{k,j})| \mathbb{1}_{\{T_{n-1}^{k,j} \leq t < T_n^{k,j}\}} \rightarrow 0$$

as  $k \rightarrow \infty$ . We claim

$$\begin{aligned}
 & \mathbb{E} \left[ g \sum_{n=1}^{\infty} \mathbb{D}^{\mathcal{Y},k,j} X(T_{n-1}^{k,j}) \Delta T_n^{k,j} \mathbb{1}_{\{T_{n-1}^{k,j} \leq t\}} \right] \\
 (4.4) \quad &= \mathbb{E} \left[ \sum_{n=2}^{\infty} g_n^{k,j} \mathbb{D}^{\mathcal{Y},k,j} X(T_{n-1}^{k,j}) \Delta T_n^{k,j} \mathbb{1}_{\{T_{n-1}^{k,j} \leq t\}} \right] \\
 &+ \mathbb{E} \left[ g \sum_{n=1}^{\infty} \Delta X^k(T_{n-1}^{k,j}) \Delta A^{k,j}(T_{n-1}^{k,j}) \mathbb{1}_{\{T_{n-1}^{k,j} \leq t\}} \right],
 \end{aligned}$$

where  $g_n^{k,j} := \mathbb{E}[g | \mathcal{F}_{T_n^{k,j}}^k] - \mathbb{E}[g | \mathcal{F}_{T_{n-1}^{k,j}}^k]$ ;  $n \geq 1$ . Indeed, for each  $n, k \geq 1$

$$\begin{aligned}
 & \int_{\{T_{n-1}^{k,j} \leq t\}} g \mathbb{D}^{\mathcal{Y},k,j} X(T_{n-1}^{k,j}) \Delta T_n^{k,j} d\mathbb{P} \\
 &= \int_{\{T_{n-1}^{k,j} \leq t\}} \mathbb{E}[g | \mathcal{F}_{T_n^{k,j}}^k] \mathbb{D}^{\mathcal{Y},k,j} X(T_{n-1}^{k,j}) \Delta T_n^{k,j} d\mathbb{P} \\
 &\quad - \int_{\{T_{n-1}^{k,j} \leq t\}} \mathbb{E}[g | \mathcal{F}_{T_{n-1}^{k,j}}^k] \mathbb{D}^{\mathcal{Y},k,j} X(T_{n-1}^{k,j}) \Delta T_n^{k,j} d\mathbb{P} \\
 &\quad + \int_{\{T_{n-1}^{k,j} \leq t\}} \mathbb{E}[g | \mathcal{F}_{T_{n-1}^{k,j}}^k] \mathbb{D}^{\mathcal{Y},k,j} X(T_{n-1}^{k,j}) \Delta T_n^{k,j} d\mathbb{P}.
 \end{aligned}$$

Since  $|\Delta A^{k,j}(T_{n-1}^{k,j})|^2 = \epsilon_k^2$  a.s. for every  $n \geq 2$ , we have

$$\begin{aligned}
 & \int_{\{T_{n-1}^{k,j} \leq t\}} \mathbb{E}[g | \mathcal{F}_{T_{n-1}^{k,j}}^k] \mathbb{D}^{\mathcal{Y},k,j} X(T_{n-1}^{k,j}) \Delta T_n^{k,j} d\mathbb{P} \\
 &= \int_{\{T_{n-1}^{k,j} \leq t\}} \mathbb{E}[g | \mathcal{F}_{T_{n-1}^{k,j}}^k] \mathbb{D}^{\mathcal{Y},k,j} X(T_{n-1}^{k,j}) \epsilon_k^2 d\mathbb{P} \\
 &= \int_{\{T_{n-1}^{k,j} \leq t\}} \mathbb{E}[g | \mathcal{F}_{T_{n-1}^{k,j}}^k] \Delta X^k(T_{n-1}^{k,j}) \Delta A^{k,j}(T_{n-1}^{k,j}) d\mathbb{P} \\
 &= \int_{\{T_{n-1}^{k,j} \leq t\}} g \Delta X^k(T_{n-1}^{k,j}) \Delta A^{k,j}(T_{n-1}^{k,j}) d\mathbb{P}.
 \end{aligned}$$

Then

$$\begin{aligned}
 \int_{\{T_{n-1}^{k,j} \leq t\}} g \mathbb{D}^{\mathcal{Y},k,j} X(T_{n-1}^{k,j}) \Delta T_n^{k,j} d\mathbb{P} &= \int_{\{T_{n-1}^{k,j} \leq t\}} g_n^{k,j} \mathbb{D}^{\mathcal{Y},k,j} X(T_{n-1}^{k,j}) \Delta T_n^{k,j} d\mathbb{P} \\
 &\quad + \int_{\{T_{n-1}^{k,j} \leq t\}} g \Delta X^k(T_{n-1}^{k,j}) \Delta A^{k,j}(T_{n-1}^{k,j}) d\mathbb{P}.
 \end{aligned}$$

This shows that (4.4) holds. We shall write

$$\begin{aligned}
 I^{k,j,1}(t) &= \mathbb{E} \sum_{n=1}^{\infty} g_n^{k,j} \mathbb{D}^{\mathcal{Y},k,j} X(T_{n-1}^{k,j}) \Delta T_n^{k,j} \mathbb{1}_{\{T_{n-1}^{k,j} \leq t\}} \\
 &\quad + \mathbb{E} g \sum_{n=1}^{\infty} \Delta X^k(T_{n-1}^{k,j}) \Delta A^{k,j}(T_{n-1}^{k,j}) \mathbb{1}_{\{T_{n-1}^{k,j} \leq t\}} \\
 &=: I_1^{k,j,1}(t) + I_2^{k,j,1}(t).
 \end{aligned}$$

Let us write  $\Delta T_n^{k,j} = (\Delta T_n^{k,j})^{1/2} (\Delta T_n^{k,j})^{1/2}$ ;  $n \geq 1$ . Then Cauchy–Schwarz inequality, the finite energy property  $\mathcal{E}^{2,\mathcal{Y}}(X) < \infty$  of  $\mathcal{Y}$  and Lemma 3.1 in [33] yield

$$\begin{aligned}
 |I_1^{k,j,1}(t)| &\leq \left( \mathbb{E} \sum_{n=1}^{\infty} |g_n^{k,j}|^2 \Delta T_n^{k,j} \mathbb{1}_{\{T_{n-1}^{k,j} \leq T\}} \right)^{1/2} \\
 &\quad \times \left( \mathbb{E} \sum_{n=1}^{\infty} |\mathbb{D}^{\mathcal{Y},k,j} X(T_{n-1}^{k,j})|^2 \Delta T_n^{k,j} \mathbb{1}_{\{T_{n-1}^{k,j} \leq T\}} \right)^{1/2} \\
 &\leq \left( T \mathbb{E} \sup_{n \geq 1} |g_n^{k,j}|^2 \mathbb{1}_{\{T_{n-1}^{k,j} \leq T\}} \right)^{1/2} (\mathcal{E}^{2,\mathcal{Y}}(X))^{1/2} \rightarrow 0
 \end{aligned}$$

as  $k \rightarrow \infty$ . Therefore, we arrive at the following conclusion:

$$(4.5) \quad \lim_{k \rightarrow \infty} \mathbb{E} g \int_0^t \mathbb{D}^{\mathcal{Y},k,j} X(s) ds \text{ exists} \iff \lim_{k \rightarrow \infty} \mathbb{E} g[X^k, A^{k,j}](t)$$

exists for  $j = 1, \dots, d$  and, in this case,

$$(4.6) \quad \lim_{k \rightarrow \infty} \mathbb{E} g \int_0^t \mathbb{D}^{\mathcal{Y},k,j} X(s) ds = \lim_{k \rightarrow \infty} \mathbb{E} g[X^k, A^{k,j}](t); \quad j = 1, \dots, d.$$

By Corollary 4.1 and the stability of  $\mathcal{Y}$ , we have

$$\begin{aligned}
 (4.7) \quad \lim_{k \rightarrow \infty} \mathbb{E} g \int_0^t \mathbb{D}^{\mathcal{Y},k,j} X(s) ds &= \lim_{k \rightarrow \infty} \mathbb{E} g[X^k, A^{k,j}](t) \\
 &= \mathbb{E} g[M, B^j](t) = \mathbb{E} g \int_0^t H_j(s) dB^j(s)
 \end{aligned}$$

for  $j = 1, \dots, d$ . Since  $\mathcal{Y}$ ,  $t \in [0, T]$  and  $g \in L^\infty(\mathbb{P})$  are arbitrary and  $\{\mathbb{D}^{\mathcal{Y},k,j} X; k \geq 1\}$  is  $L_a^2(\mathbb{P} \times Leb)$ -weakly relatively compact for each  $j = 1, \dots, d$  and  $\mathcal{Y}$ , we then conclude that

$$\mathcal{D}^{\mathcal{Y}} X = H$$

for every stable discrete structure  $\mathcal{Y}$  associated with  $X$ . This concludes the proof. □

A closer look at the proof of Theorem 4.1 yields the following result.

COROLLARY 4.2. Let  $X = X(0) + \sum_{j=1}^d \int H_j dB^j + V$  be an  $\mathbb{F}$ -adapted process such that  $H_j \in L_a^2(\mathbb{P} \times Leb)$ ;  $j = 1, \dots, d$  and  $V \in \mathbf{B}^2(\mathbb{F})$  has continuous paths. Then  $\mathcal{Y}$  is a stable imbedded discrete structure for  $X$  if, and only if,  $\mathcal{Y}$  has finite energy and  $\mathcal{D}^{\mathcal{Y}}X$  exists.

PROOF. Just observe (4.5), (4.6), (4.7) and (4.3).  $\square$

In view of Theorem 4.1 and Corollary 4.2, it is natural to arrive at the following definition.

DEFINITION 4.4. Let  $X = X(0) + \sum_{j=1}^d \int H_j dB^j + V$  be an  $\mathbb{F}$ -adapted process such that  $H_j \in L_a^2(\mathbb{P} \times Leb)$ ;  $j = 1, \dots, d$  and  $V \in \mathbf{B}^2(\mathbb{F})$  has continuous paths. We say that  $X$  is *weakly differentiable* if there exists a finite energy imbedded discrete structure  $\mathcal{Y} = ((X^k)_{k \geq 1}, \mathcal{D})$  such that  $\mathcal{D}^{\mathcal{Y}}X$  exists. The space of weakly differentiable processes will be denoted by  $\mathcal{W}(\mathbb{F})$ .

Let  $X \in \mathcal{W}(\mathbb{F})$  with a decomposition

$$(4.8) \quad X = X(0) + \sum_{j=1}^d \int H_j dB^j + V,$$

where  $H = (H_1, \dots, H_d)$ ,  $H_j \in L_a^2(\mathbb{P} \times Leb)$ ;  $j = 1, \dots, d$  and  $V \in \mathbf{B}^2(\mathbb{F})$  is a continuous process. Then Theorem 4.1 allows us to define

$$(4.9) \quad \mathcal{D}X := \mathcal{D}^{\mathcal{Y}}X$$

for every stable imbedded discrete structure  $\mathcal{Y} = ((X^k)_{k \geq 1}, \mathcal{D})$  w.r.t.  $X$  and, in this case,  $\mathcal{D}X = H$ . The weak differentiability notion requires existence of  $\mathcal{D}^{\mathcal{Y}}X$  for a finite energy imbedded discrete structure  $\mathcal{Y} = ((X^k)_{k \geq 1}, \mathcal{D})$  and, from Theorem 4.1, this concept of derivative does not depend on the choice of the stable structure  $\mathcal{Y}$ .

REMARK 4.2. It is important to observe that for a given  $X$  of the form (4.8), we cannot expect that  $\mathcal{D}^{\mathcal{Y}}X = H$  holds for every  $\mathcal{Y}$  because there are imbedded discrete structures

$$X^k = X^k(0) + M^{\mathcal{Y},k} + N^{\mathcal{Y},k}$$

such that  $X^k \rightarrow X$  weakly in  $\mathbf{B}^2(\mathbb{F})$  but  $\{M^{\mathcal{Y},k}; k \geq 1\}$  fails to converge to the martingale component of  $X$ . In this case, because of  $[X^k, A^{k,j}] = [M^{\mathcal{Y},k}, A^{k,j}]$ , Lemma 4.3 and (4.5),  $\mathcal{D}^{\mathcal{Y}}X$  may not even exist or it will not coincide with  $H$ . This type of phenomena is well known in time-deterministic discretizations of filtrations; see, for example, [10, 36] and other references therein. This is the reason why we restrict the computation of the weak derivative  $\mathcal{D}X$  to stable structures and this is the best one might expect.

REMARK 4.3. Perhaps the simplest class of examples which are not in  $\mathcal{W}(\mathbb{F})$  is given by Brownian motion transformations with low regularity

$$f(B), \quad f \in H^{1,q}(\mathbb{R}),$$

where  $H^{1,q}(\mathbb{R})$  is the Sobolev space with degree of integrability  $1 \leq q < 2$ .

A closer look at the proof of Theorem 4.1 gives the following useful criteria to compute the derivative.

PROPOSITION 4.1. *If  $X \in \mathcal{W}(\mathbb{F})$  is associated with a stable imbedded discrete structure  $\mathcal{Y}$ , then  $X$  has  $\mathcal{Y}$ -covariations  $\langle X, B^j \rangle^{\mathcal{Y}}$ ;  $j = 1, \dots, d$  as absolutely continuous processes and*

$$\mathcal{D}_j X = \frac{d\langle X, B^j \rangle^{\mathcal{Y}}}{dt}; \quad j = 1, \dots, d.$$

PROOF. Just observe (4.5), (4.6), (4.7) and use Lemma 4.3.  $\square$

PROPOSITION 4.2. *If  $X \in \mathcal{W}(\mathbb{F})$  is a weakly differentiable process with a decomposition*

$$X(t) = X(0) + \sum_{j=1}^d \int_0^t \mathcal{D}_j X(s) dB^j(s) + V_X(t); \quad 0 \leq t \leq T,$$

then  $V_X$  can be described by

$$(4.10) \quad V_X(\cdot) = \lim_{k \rightarrow +\infty} \sum_{j=1}^d \int_0^\cdot \mathbb{U}^{\mathcal{Y},k,j} X(s) ds$$

weakly in  $\mathbf{B}^2(\mathbb{F})$  as  $k \rightarrow +\infty$  for every stable imbedded discrete structure  $\mathcal{Y}$  associated with  $X$ .

PROOF. The proof is an immediate consequence of Proposition 3.1 and Theorem 4.1.  $\square$

It is important to stress that any  $X \in \mathcal{W}(\mathbb{F})$  can be decomposed into a martingale and a nonmartingale component  $V_X$  which encodes all the possible orthogonal infinitesimal variations of  $X$  w.r.t. the Brownian motion noise and, in particular, it may have unbounded variation paths; See Section 4.2 in this direction. Indeed, the process (4.10) encodes all the possible orthogonal infinitesimal variations of a weakly differentiable process w.r.t. the Brownian motion noise; see Section 4.4 for the smooth case and [34] for a universal variational characterization of  $V_X$  as a functional of local-times.

The name weak derivative is justified by the following remark. Under regularity conditions, there exists a local description of  $\mathcal{D}_j X$  which justifies the name weak derivative. For a given  $t \geq 0$  and  $j \in \{1, \dots, d\}$ , we set

$$(4.11) \quad T^{t,\epsilon,j} := \inf\{s \geq 0; |B^j(t+s) - B^j(t)| = \epsilon\},$$

where  $\epsilon > 0$ . The stopping time  $T^{t,\epsilon,j}$  localizes the  $j$ th Brownian motion around the point  $t$ . Let  $X$  be a real-valued square-integrable  $\mathbb{F}$ -semimartingale of the form

$$X(t) = X(0) + \sum_{j=1}^d \int_0^t H_j(s) dB^j(s) + \int_0^t Z(s) ds; \quad 0 \leq t \leq T.$$

Assume that  $H_j$  has càdlàg paths for  $j = 1, \dots, d$ . Then a routine computation shows

$$(4.12) \quad \lim_{\epsilon \rightarrow 0+} \frac{X(t + T^{t,\epsilon,j}) - X(t)}{B^j(t + T^{t,\epsilon,j}) - B^j(t)} = H_j(t) \quad \text{in probability}$$

for every  $t \geq 0$  and  $1 \leq j \leq d$ . At first glance, one may say that regularity condition on  $H = (H_1, \dots, H_d)$  is not too strong, but the strong regularity comes from the drift term. Indeed, even in the semimartingale case, when the drift is a bounded variation process without absolutely continuous paths, the limit (4.12) may not exist.

4.1. *Differentiable processes.* The goal of this section is to show that the differentiability notion presented in Definition 4.4 covers a wide range class of Wiener functionals including non-semimartingales. It turns out that any strong Dirichlet process in the sense of Bertoin [1] will be weakly differentiable, and hence all square-integrable continuous  $\mathbb{F}$ -semimartingales will be weakly differentiable as well; see also Section 4.2 for examples concerning finite  $p$ -variation Wiener functionals. Let us recall the notion of a strong Dirichlet process: For a given strictly increasing sequence of  $\mathbb{F}$ -stopping times  $\tau = (S_0, \dots, S_n, \dots)$  such that  $S_0 = 0$  a.s. and  $S_n \uparrow +\infty$  a.s. as  $n \rightarrow +\infty$ , we write

$$|\tau| := \mathbb{E} \sup_{n \geq 0} |S_{n+1} - S_n| \mathbb{1}_{\{S_n \leq T\}}.$$

We denote by  $\mathbb{T}$  as the set of all random partitions of the above form. If  $Y$  is a Wiener functional and  $\tau = (S_0, \dots, S_n, \dots)$  is a random mesh, then we write

$$Q_\tau^\alpha(Y) := |Y(0)|^\alpha + \sum_{n=0}^\infty |Y(S_{n+1}) - Y(S_n)|^\alpha$$

for  $1 \leq \alpha < +\infty$ .

DEFINITION 4.5. A Wiener functional  $Y$  is said to be of finite  $\alpha$ -variation if

$$\|Y\|_{Q^\alpha} := \sup_{\tau \in \mathbb{T}} [\mathbb{E}(Q_\tau^\alpha(Y))]^{1/\alpha} < \infty.$$

Let us denote  $\mathbf{Q}^\alpha(\mathbb{F})$  as the set of Wiener functionals with  $\alpha$ -finite variation. One can easily check that  $(\mathbf{Q}^\alpha(\mathbb{F}), \|\cdot\|_{\mathbf{Q}^\alpha})$  is a Banach space. Let  $\mathbf{Q}_0^\alpha(\mathbb{F})$  be the linear subspace of  $\mathbf{Q}^\alpha(\mathbb{F})$  constituted of the elements  $Y \in \mathbf{Q}_0^\alpha(\mathbb{F})$  with  $\alpha$ -null variation, that is,

$$\lim_{|\tau| \rightarrow 0} \mathbb{E}(\mathcal{Q}_\tau^\alpha(Y)) = 0.$$

One can easily check that  $\mathbf{Q}_0^2(\mathbb{F})$  is a Banach space. We refer to [1] for more details. Finally, the space of strong Dirichlet processes is given by

$$\mathcal{R}(\mathbb{F}) := \mathbf{H}^2(\mathbb{F}) \oplus \mathbf{Q}_0^2(\mathbb{F}).$$

In other words, any  $X \in \mathcal{R}(\mathbb{F})$  admits a unique decomposition

$$X = X(0) + M + N,$$

where  $M \in \mathbf{H}^2(\mathbb{F})$  and  $N \in \mathbf{Q}_0^2(\mathbb{F})$ .

**THEOREM 4.2.** *Let  $X = X(0) + M + N$  be a strong Dirichlet process where  $(M, N)$  is the canonical decomposition where  $M = \sum_{j=1}^d \int H_j dB^j \in \mathbf{H}^2(\mathbb{F})$  and  $N \in \mathbf{Q}_0^2(\mathbb{F})$ . Then  $X \in \mathcal{W}(\mathbb{F})$  is weakly differentiable, where*

$$(4.13) \quad \mathcal{D}X = H$$

and

$$(4.14) \quad N = \lim_{k \rightarrow +\infty} \sum_{j=1}^d \int_0^\cdot \mathbb{U}^{\mathcal{Y},k,j} X(s) ds$$

weakly in  $\mathbf{B}^2(\mathbb{F})$  for every stable imbedded discrete structure  $\mathcal{Y}$  w.r.t.  $X$ .

**REMARK 4.4.** Theorem 4.2 justifies the use of the variational operators  $(\mathbb{D}^{\mathcal{Y},k,j} X, \mathbb{U}^{\mathcal{Y},k,j} X)$  based on an imbedded discrete structure  $\mathcal{Y} = ((X^k)_{k \geq 1}, \mathcal{D})$  associated with a large class of Wiener functionals  $X$ . However, we stress that in many examples the computation of  $(\mathcal{D}^{\mathcal{Y},k,j} X, \mathbb{U}^{\mathcal{Y},k,j} X)$  is good enough to extract nontrivial information from  $X$  without needing to go to the limits (4.13) and (4.14); see the works [3, 32, 35].

**4.2. Differentiability and finite  $p$ -variation Wiener functionals.** Before we proceed with the proof of Theorem 4.2, let us make a few remarks concerning finite  $p$ -variation processes associated with Theorem 4.2. Let us recall that  $\mathcal{R}(\mathbb{F})$  contains a large class of Wiener functionals beyond semimartingales. For instance, it is known that (see Lemma 1.1 in [2]) that

$$\mathbf{Q}^\alpha(\mathbb{F}) \subset \mathbf{Q}_0^\beta(\mathbb{F}) \text{ up to localization}$$

for every  $1 \leq \alpha < \beta$ . Let  $\mathcal{V}_\alpha(\mathbb{F})$  be the space of all Wiener functionals of the form

$$X = X(0) + M + N,$$

where  $M \in \mathbf{H}^2(\mathbb{F})$  and  $N \in \mathbf{Q}^\alpha(\mathbb{F})$  for  $1 \leq \alpha < 2$ . Then  $\mathcal{V}_\alpha(\mathbb{F}) \subset \mathcal{R}(\mathbb{F})$ ;  $1 \leq \alpha < 2$ , and hence Theorem 4.2 applies to elements of  $\mathcal{V}_\alpha(\mathbb{F})$ ;  $1 \leq \alpha < 2$  up to localization. More concrete examples arise by making use of the pathwise Young integral. One typical example of a class of processes in  $\mathcal{V}_\alpha(\mathbb{F})$  ( $1 \leq \alpha < 2$ ) is given by

$$X(t) = X(0) + \sum_{j=1}^d \int_0^t \mathcal{D}_j X(s) dB^j(s) + \int_0^t Y(s) dW(s),$$

where  $(Y, W)$  is a pair of Wiener functionals satisfying the constraint

$$Y \in \mathbf{Q}^\beta(\mathbb{F}), \quad W \in \mathbf{Q}^\alpha(\mathbb{F}); \quad \frac{1}{\alpha} + \frac{1}{\beta} > 1,$$

where

$$t \mapsto \int_0^t Y(s) dW(s) \in \mathbf{Q}^\alpha(\mathbb{F})$$

is interpreted in Young sense.

For a typical example, let us present the fractional Brownian motion case. The following result is known for deterministic partitions but it also holds for random partitions based on Garsia–Rodemich–Rumsey’s inequality. For sake of completeness, we give the details of the proof.

LEMMA 4.4. *If  $\frac{1}{2} < H < 1$ , then*

$$B_H \in \mathbf{Q}^p(\mathbb{F}); \quad p > \frac{1}{H}.$$

PROOF. See the Supplementary Material [33].  $\square$

Then  $\mathcal{W}(\mathbb{F})$  contains the class of processes of the form

$$X(t) = X(0) + \sum_{j=1}^d \int_0^t \mathcal{D}_j X(s) dB^j(s) + \int_0^t Y(s) dB_H(s),$$

where  $Y \in \mathbf{Q}^\beta(\mathbb{F})$ ,  $\frac{1}{\beta} + \frac{1}{H-\epsilon} > 1$  for  $\epsilon \in (0, H)$  and  $\frac{1}{2} < H < 1$ .

4.3. *Proof of Theorem 4.2.* The proof of Theorem 4.2 will be divided into several steps. Throughout this section, we fix an arbitrary strong Dirichlet process

$$X = X(0) + \sum_{j=1}^d \int H_j dB^j + N,$$

where  $N \in \mathbf{Q}_0^2(\mathbb{F})$ . In view of Theorem 4.1, we need to show the existence of  $\mathcal{D}^{\mathcal{Y}}X$  for a stable imbedded discrete structure  $\mathcal{Y}$  associated with  $X$ . Since  $X \in \mathcal{R}(\mathbb{F})$  is an arbitrary Dirichlet process, it is natural to guess that the use of the canonical structure  $\mathcal{Y} = ((\delta^k X)_{k \geq 1}, \mathcal{D})$  will be a natural choice.

**PROPOSITION 4.3.** *If  $X \in \mathcal{R}(\mathbb{F})$  and  $\mathcal{Y} = ((\delta^k X)_{k \geq 1}, \mathcal{D})$ , then  $\mathcal{E}^{2,\mathcal{Y}}(X) < +\infty$ . In particular, there exists a constant  $C$  which does not depend on  $k \geq 1$  such that*

$$(4.15) \quad \mathbb{E}[\delta^k X, \delta^k X](T) \leq C \mathbb{E} \sum_{\ell=1}^{\infty} |X(T_{\ell}^k) - X(T_{\ell-1}^k)|^2 \mathbb{1}_{\{T_{\ell}^k \leq T\}}$$

for every  $k \geq 1$ .

**PROOF.** Throughout this proof,  $C$  is a constant which may defer from line to line. The proof is significantly more tricky than the one-dimensional case treated in Lemma 3.1 in [30]. We divide the proof into three parts.

*STEP 1:* Let  $\tau_n^{k,j} := \min\{T_m^{k,r}; T_m^{k,r} > T_n^{k,j}\}$ . It is easy to see that  $\tau_n^{k,j}$  is an  $\mathbb{F}^k$ -stopping time. Indeed, for any  $s \geq 0$ , we have

$$\begin{aligned} \{\tau_n^{k,j} \leq s\} &= \{\tau_n^{k,j} \leq s\} \cap \{s \leq T_n^{k,j}\} \cup \{\tau_n^{k,j} \leq s\} \cap \{T_n^{k,j} < s\} \\ &= \{\tau_n^{k,j} \leq s\} \cap \{T_n^{k,j} < s\} \\ &= \bigcup_{m=1}^d \bigcup_{r \geq 1} (\{T_n^{k,j} < T_r^{k,m}\} \cap \{T_r^{k,m} \leq s\}), \end{aligned}$$

where  $\{T_n^{k,j} < T_r^{k,m}\} \in \mathcal{F}_{T_r^{k,m}-}^k$  so that  $\{T_n^{k,j} < T_r^{k,m}\} \cap \{T_r^{k,m} \leq s\} \in \mathcal{F}_s^k$  for each  $j, m = 1, \dots, d$  and  $n \geq 0$  and  $r \geq 1$ . Therefore,  $\{\tau_n^{k,j} \leq s\} \in \mathcal{F}_s^k$ . This shows that  $\tau_n^{k,j}$  is an  $\mathbb{F}^k$ -stopping time for each  $k \geq 1, n \geq 0$  and  $1 \leq j \leq d$ . Now, we notice that

$$\{\Delta \delta^k X \neq 0\} \subset \bigcup_{j=1}^{\infty} \bigcup_{n=0}^{\infty} [[\tau_n^{k,j}, \tau_n^{k,j}]]$$

and

$$\begin{aligned} \Delta \delta^k X(\tau_n^{k,j}) &= \delta^k X(\tau_n^{k,j}) - \delta^k X(T_n^{k,j}) \\ &\text{a.s. for each } k \geq 1, n \geq 0 \text{ and } 1 \leq j \leq d. \end{aligned}$$

In the sequel, we use the convention that  $T_{-1}^{k,j} = T_0^{k,j} = 0$ . We shall write

$$(4.16) \quad [\delta^k X, \delta^k X](T) = \sum_{j=1}^d \sum_{\ell=0}^{\infty} |\Delta \delta^k X(\tau_{\ell}^{k,j})|^2 \mathbb{1}_{\{\tau_{\ell}^{k,j} \leq T\}},$$

where

$$\begin{aligned} \Delta\delta^k X(\tau_\ell^{k,j}) &= \mathbb{E}[X(\tau_\ell^{k,j})|\mathcal{F}_{\tau_\ell^{k,j}}^k] - \mathbb{E}[X(T_\ell^{k,j})|\mathcal{F}_{T_\ell^{k,j}}^k] \\ &= \mathbb{E}[X(\tau_\ell^{k,j})|\mathcal{F}_{\tau_\ell^{k,j}}^k] - \mathbb{E}[X(T_{\ell-1}^{k,j})|\mathcal{F}_{\tau_\ell^{k,j}}^k] \\ &\quad + \mathbb{E}[X(T_{\ell-1}^{k,j})|\mathcal{F}_{\tau_\ell^{k,j}}^k] - \mathbb{E}[X(T_\ell^{k,j})|\mathcal{F}_{T_\ell^{k,j}}^k]. \end{aligned}$$

We notice that

$$\mathcal{F}_{\tau_\ell^{k,j}}^k = \mathcal{F}_{T_\ell^{k,j}}^k \vee \mathcal{H}_\ell^{k,j},$$

where  $\mathcal{H}_\ell^{k,j} = \sigma(\tau_\ell^{k,j} - T_\ell^{k,j}, \Delta A^{k,m}(\tau_\ell^{k,j}); 1 \leq m \leq d)$  for every  $j = 1, \dots, d$  and  $\ell \geq 0$ . Let us denote

$$\varphi_\ell^{k,j} := \mathbb{E}[X(T_{\ell-1}^{k,j})|\mathcal{F}_{\tau_\ell^{k,j}}^k] - \mathbb{E}[X(T_\ell^{k,j})|\mathcal{F}_{T_\ell^{k,j}}^k].$$

By the very definition,  $\varphi_0^{k,j} = 0, \varphi_1^{k,j} = \mathbb{E}[X(0) - X(T_1^{k,j})|\mathcal{F}_{T_1^{k,j}}^k]$ . For  $\ell \geq 2$ , we need to work a little more.

*STEP 2:* Let us fix  $\ell \geq 2$  and  $1 \leq j \leq d$ . We claim for every  $Z_1 \in L^2(\mathcal{F}_{T_{\ell-1}^{k,j}})$  and  $Z_2 \in L^2(\mathcal{H}_\ell^{k,j})$ , we have

$$(4.17) \quad \mathbb{E}[Z_1 Z_2 | \mathcal{F}_{T_\ell^{k,j}}^k] = \mathbb{E}[Z_1 | \mathcal{F}_{T_\ell^{k,j}}^k] \mathbb{E}[Z_2 | \mathcal{F}_{T_\ell^{k,j}}^k].$$

In other words,  $\mathcal{F}_{T_{\ell-1}^{k,j}}$  and  $\mathcal{H}_\ell^{k,j}$  are conditionally independent given  $\mathcal{F}_{T_\ell^{k,j}}^k$ . Let us fix  $\ell \geq 2$  and  $j = 1, \dots, d$ . Let  $\mathcal{D}_\ell^{k,j}$  the sigma-algebra such that

$$\mathcal{F}_{T_\ell^{k,j}}^k = \mathcal{F}_{T_{\ell-1}^{k,j}}^k \vee \mathcal{D}_\ell^{k,j}.$$

It is easy to check that

$$\mathcal{D}_\ell^{k,j} = \sigma(W^{k,m,\ell,j}(s \wedge (T_\ell^{k,j} - T_{\ell-1}^{k,j})); s \geq 0, 1 \leq m \leq d),$$

where  $W^{k,m,\ell,j}(s) := A^{k,m}(s + T_{\ell-1}^{k,j}) - A^{k,m}(T_{\ell-1}^{k,j}); s \geq 0$  for  $\ell \geq 2$  and  $1 \leq j, m \leq d$ . We notice that

$$\begin{aligned} \mathbb{E}[Z_1 Z_2 | \mathcal{F}_{T_\ell^{k,j}}^k] &= \mathbb{E}[\mathbb{E}[Z_1 Z_2 | \mathcal{F}_{T_{\ell-1}^{k,j}}^k \vee \mathcal{D}_\ell^{k,j}] | \mathcal{F}_{T_\ell^{k,j}}^k] \\ (4.18) \quad &= \mathbb{E}[Z_1 \mathbb{E}[Z_2 | \mathcal{F}_{T_{\ell-1}^{k,j}}^k \vee \mathcal{D}_\ell^{k,j}] | \mathcal{F}_{T_\ell^{k,j}}^k]. \end{aligned}$$

Since  $Z_2$  is  $\mathcal{H}_\ell^{k,j}$ -measurable, then there exists  $f : \mathbb{R}_+ \times \{-\epsilon_k, 0, \epsilon_k\}^d \rightarrow \mathbb{R}$  such that

$$Z_2 = f(\tau_\ell^{k,j} - T_\ell^{k,j}, \Delta A^{k,m}(\tau_\ell^{k,j}); 1 \leq m \leq d) \quad \text{a.s.}$$

More importantly,  $Z_2$  is a functional of  $A^{k,m}(T_n^{k,r}), T_n^{k,m}; n \geq 0, 1 \leq r, m \leq d$ . In this case,

$$(4.19) \quad \mathbb{E}[Z_2 | \mathcal{F}_{T_\ell^{k,j}}^k] = \mathbb{E}[Z_2 | \mathcal{F}_{T_{\ell-1}^{k,j}}^k \vee \mathcal{D}_\ell^{k,j}] = \mathbb{E}[Z_2 | \mathcal{F}_{T_{\ell-1}^{k,j}}^k \vee \mathcal{H}_\ell^{k,j}] \quad \text{a.s.}$$

Therefore, (4.19) and (4.18) yield (4.17). The conditional independence (4.17) yields

$$\begin{aligned} \mathbb{E}[X(T_{\ell-1}^{k,j}) | \mathcal{F}_{\tau_\ell^{k,j}}^k] &= \mathbb{E}[X(T_{\ell-1}^{k,j}) | \mathcal{F}_{T_\ell^{k,j}}^k \vee \mathcal{H}_\ell^{k,j}] \\ &= \mathbb{E}[X(T_{\ell-1}^{k,j}) | \mathcal{F}_{T_\ell^{k,j}}^k], \end{aligned}$$

and hence  $\varphi_\ell^{k,j} = \mathbb{E}[X(T_{\ell-1}^{k,j}) - X(T_\ell^{k,j}) | \mathcal{F}_{T_\ell^{k,j}}^k]; \ell \geq 2$ .

STEP 3: As a consequence of STEP 2, we may apply Jensen’s inequality to find a positive constant  $C$  (which does not depend on  $k, j, n$ ) such that

$$(4.20) \quad \begin{aligned} |\Delta \delta^k X(\tau_n^{k,j})|^2 &\leq C(\mathbb{E}[|X(\tau_n^{k,j}) - X(T_n^{k,j})|^2 | \mathcal{F}_{\tau_n^{k,j}}^k] \\ &\quad + \mathbb{E}[|X(T_n^{k,j}) - X(T_{n-1}^{k,j})|^2 | \mathcal{F}_{T_n^{k,j}}^k] \\ &\quad + \mathbb{E}[|X(T_{\ell-1}^{k,j}) - X(T_\ell^{k,j})|^2 | \mathcal{F}_{T_\ell^{k,j}}^k]) \quad \text{a.s.} \end{aligned}$$

Moreover, we notice that  $\sum_{j=1}^d \sum_{\ell=0}^\infty \mathbb{E}[|X(T_{\ell-1}^{k,j}) - X(T_\ell^{k,j})|^2 | \mathcal{F}_{T_\ell^{k,j}}^k] \mathbb{1}_{\{\tau_\ell^{k,j} \leq T\}}$  equals to

$$(4.21) \quad \sum_{j=1}^d \sum_{\ell=0}^\infty \mathbb{E}[|X(T_{\ell-1}^{k,j}) - X(T_\ell^{k,j})|^2 | \mathcal{F}_{T_\ell^{k,j}}^k] \mathbb{1}_{\{T_\ell^{k,j} \leq T\}} \quad \text{a.s.}$$

Then (4.16), (4.20) and (4.21) yield

$$\begin{aligned} \mathbb{E}[\delta^k X, \delta^k X](T) &\leq C \mathbb{E} \sum_{j=1}^d \sum_{\ell=0}^\infty |X(\tau_\ell^{k,j}) - X(T_\ell^{k,j})|^2 \mathbb{1}_{\{\tau_\ell^{k,j} \leq T\}} \\ &\quad + C \mathbb{E} \sum_{j=1}^d \sum_{\ell=0}^\infty |X(T_\ell^{k,j}) - X(T_{\ell-1}^{k,j})|^2 \mathbb{1}_{\{\tau_\ell^{k,j} \leq T\}} \\ &\quad + C \mathbb{E} \sum_{j=1}^d \sum_{\ell=0}^\infty |X(T_{\ell-1}^{k,j}) - X(T_\ell^{k,j})|^2 \mathbb{1}_{\{T_\ell^{k,j} \leq T\}} \\ &\leq 3C \mathbb{E} \sum_{\ell=1}^\infty |X(T_\ell^k) - X(T_{\ell-1}^k)|^2 \mathbb{1}_{\{T_\ell^k \leq T\}} \\ &\leq 3C \|X\|_{\mathbf{Q}^2}^2 < \infty. \end{aligned}$$

This shows  $\mathcal{E}^{2,\mathcal{Y}}(X) < +\infty$  and (4.15). This concludes the proof.  $\square$

Before we proceed with the proof, let us recall some basics of semimartingale theory which will also help us to fix notation. Let  $\mathbf{H}^2(\mathbb{F}^k)$  be the space of all square-integrable  $\mathbb{F}^k$ -martingales starting at zero. From [26], we know that any square-integrable  $\mathbb{F}^k$ -martingale has bounded variation paths and it is purely discontinuous whose jumps are exhausted by  $\bigcup_{n \geq 1} [[T_n^k, T_n^k]]$ . In this case, any  $Y^k \in \mathbf{H}^2(\mathbb{F}^k)$  can be uniquely written as

$$(4.22) \quad Y^k(t) = Y^{k,\text{pj}}(t) - N^{k,Y^k}(t); \quad t \geq 0,$$

where  $N^{k,Y^k}$  is an  $\mathbb{F}^k$ -predictable continuous bounded variation process,

$$Y^{k,\text{pj}}(t) := \sum_{0 < s \leq t} \Delta Y^k(s); \quad t \geq 0$$

and  $Y^{k,\text{pj}}(0) = N^{k,Y^k}(0) = 0$ . From Theorems 1 and 2 in [26], we can always write

$$Y^{k,\text{pj}}(t) = \sum_{n=1}^{\infty} \Delta Y^k(T_n^k) \mathbb{1}_{\{T_n^k \leq t\}}; \quad t \geq 0.$$

At first, we observe that Proposition 3.1 in [30] holds for any sequence  $\{Y^k; k \geq 1\}$  of the form (4.22).

LEMMA 4.5. *Let  $\{Y^k; k \geq 1\}$  be a sequence of square-integrable martingales  $Y^k \in \mathbf{H}^2(\mathbb{F}^k); k \geq 1$ . If  $\sup_{k \geq 1} \mathbb{E}[Y^k, Y^k](T) < \infty$ , then  $\{Y^k; k \geq 1\}$  is  $\mathbf{B}^2$ -weakly relatively sequentially compact where all limit points are  $\mathbb{F}$ -square-integrable martingales.*

PROOF. Let us consider  $R^k(t) := \mathbb{E}[Y^k(T) | \mathcal{F}_t]; 0 \leq t \leq T$ , we can apply exactly the same arguments given in the proof of Proposition 3.1 in [30] to show that both  $\{R^k; k \geq 1\}$  and  $\{Y^k; k \geq 1\}$  are  $\mathbf{B}^2$ -weakly relatively compact and all limit points are  $\mathbb{F}$ -square-integrable martingales over  $[0, T]$ .  $\square$

In the sequel, we fix  $W \in \mathbf{H}^2(\mathbb{F})$  starting at zero and for simplicity of notation we assume that  $W$  is defined on the whole positive line  $[0, +\infty)$  and write  $W^k(t) := \mathbb{E}[W(T) | \mathcal{F}_t^k]; t \geq 0$ . Let,  $W^k(t) = W^{k,\text{pj}}(t) - N^{k,W^k}(t); t \geq 0$ , be the  $\mathbb{F}^k$ -special semimartingale decomposition given in (4.22). Let  $\delta^k W = M^{k,W} + N^{k,W}$  be the special semimartingale decomposition given by (2.10) in [30]. Since  $W \in \mathbf{H}^2(\mathbb{F})$  and  $\mathbb{F}^k \subset \mathbb{F}$  for every  $k \geq 1$ , then

$$\mathbb{E}[W(T) | \mathcal{F}_t^k] = \mathbb{E}[\mathbb{E}[W(\infty) | \mathcal{F}_T] | \mathcal{F}_t^k] = \mathbb{E}[W(\infty) | \mathcal{F}_t^k]; \quad 0 \leq t \leq T$$

so that  $\mathbb{E}[W(T) | \mathcal{F}_{T_n^k}^k] = \mathbb{E}[W(\infty) | \mathcal{F}_{T_n^k}^k] = \mathbb{E}[W(T_n^k) | \mathcal{F}_{T_n^k}^k]$  on  $\{T_n^k \leq T\}$ . In other words,

$$(4.23) \quad W^k(T_n^k) = \delta^k W(T_n^k) \text{ on } \{T_n^k \leq T\}; \quad k \geq 1.$$

Let us denote  $Z^k := W^k - M^{k,W}; k \geq 1$ . Since  $Z^k$  is a purely discontinuous martingale, then it has a decomposition of the form (4.22).

LEMMA 4.6. *The sequence  $\{Z^k; k \geq 1\}$  satisfies  $\sup_{k \geq 1} \mathbb{E}[Z^k, Z^k](T) < \infty$  and*

$$\Delta Z^k(T_n^k) \mathbb{1}_{\{T_n^k \leq t\}} = (N^{k,W^k}(T_n^k) - N^{k,W^k}(T_{n-1}^k)) \mathbb{1}_{\{T_n^k \leq t\}}; \quad n \geq 1.$$

Therefore,  $\lim_{k \rightarrow \infty} Z^k = 0$  weakly in  $\mathbf{B}^2(\mathbb{F})$ , if and only if

$$(4.24) \quad [Z^k, A^{k,j}](t) = \sum_{n=1}^{\infty} (N^{k,W^k}(T_n^k) - N^{k,W^k}(T_{n-1}^k)) \Delta A^{k,j}(T_n^k) \mathbb{1}_{\{T_n^k \leq t\}} \rightarrow 0$$

weakly in  $L^1(\mathbb{P})$  as  $k \rightarrow \infty$ , for every  $t \in [0, T]$ .

PROOF. By applying Proposition 4.3 and the fact that  $W \in \mathbf{H}^2(\mathbb{F})$ , we observe that  $\sup_{k \geq 1} \mathbb{E}[\delta^k W, \delta^k W](T) < \infty$ . Moreover, Burkholder–Davis–Gundy inequality yields the bound  $\sup_{k \geq 1} \mathbb{E}[W^k, W^k](T) < \infty$ , and hence,  $\sup_{k \geq 1} \mathbb{E}[Z^k, Z^k](T) < \infty$ . For a given  $t \in [0, T]$ , we have

$$(4.25) \quad \begin{aligned} \Delta Z^k(T_n^k) &= (\Delta W^k(T_n^k) - \Delta M^{k,W}(T_n^k)) \\ &= (W^k(T_n^k) - W^k(T_n^k-) - \delta^k W(T_n^k) + \delta^k W(T_{n-1}^k)) \\ &= (W^k(T_n^k) - W^k(T_n^k-) - W^k(T_n^k) + W^k(T_{n-1}^k)) \end{aligned}$$

$$(4.26) \quad \begin{aligned} &= (-W^k(T_n^k-) + W^k(T_{n-1}^k)) \\ &= (N^{k,W^k}(T_n^k) - N^{k,W^k}(T_{n-1}^k)) \end{aligned}$$

on  $\{T_n^k \leq t\}$  for  $n \geq 1$ , where in (4.25) and (4.26), we have used identity (4.23) and the fact that  $N^{k,W^k}$  has continuous paths, respectively. The last statement (4.24) is a simple application of Lemmas 4.5, 4.3 and the predictable martingale representation of the Brownian motion.  $\square$

LEMMA 4.7. *Let  $\delta^k W = M^{k,W} + N^{k,W}$  be the canonical  $\mathbb{F}^k$ -semimartingale decomposition for a Brownian martingale  $W \in \mathbf{H}^2(\mathbb{F})$  starting at zero. Then*

$$M^{k,W} \rightarrow W$$

weakly in  $\mathbf{B}^2(\mathbb{F})$  as  $k \rightarrow \infty$ . Moreover,  $\lim_{k \rightarrow \infty} [\delta^k W, A^{k,j}](t) = [W, B^j](t)$  weakly in  $L^1(\mathbb{P})$  for every  $t \in [0, T]$ ,  $j = 1, \dots, d$ .

PROOF. Lemma 4.6 and the predictability of  $N^{k,W^k}$  yield  $\Delta Z^k(T_n^k) \mathbb{1}_{\{T_n^k \leq t\}}$  is  $\mathcal{F}_{T_n^k-}^k$ -measurable for each  $n \geq 1$  and  $t \geq 0$ . By construction,  $|\Delta A^{k,j}(T_n^k)| > 0$  only on the set  $\{\mathfrak{N}_1(A_n^k) = j\}$ . Then the strong Markov property and the fact that  $\sigma_n^{k,j}$  is independent from  $\Delta T_n^{k,j}$  imply that

$$\mathbb{E}[\Delta A^{k,j}(T_n^k) | \mathcal{F}_{T_n^k-}^k] = 0 \quad \text{a.s., } j = 1, \dots, d.$$

Therefore,

$$\begin{aligned} \mathbb{E}[\Delta Z^k(T_n^k) \Delta A^{k,j}(T_n^k) | \mathcal{F}_{T_n^k-}^k] &= \Delta Z^k(T_n^k) \mathbb{E}[\Delta A^{k,j}(T_n^k) | \mathcal{F}_{T_n^k-}^k] \\ &= 0 \quad \text{a.s.} \end{aligned}$$

on  $\{T_n^k \leq t\}$  for each  $n \geq 1$  and  $t \geq 0$  and  $1 \leq j \leq d$ . By applying Proposition 1.1 in [29] on the pure jump process  $[Z^k, A^{k,j}]$  given by (4.24), we can safely state that this process is an  $\mathbb{F}^k$ -martingale for every  $k \geq 1$ . Lemma 4.6 yields

$$\sup_{k \geq 1} \mathbb{E}[Z^k, Z^k](T) = \sup_{k \geq 1} \mathbb{E} \sum_{n=1}^{\infty} (N^{k,W^k}(T_n^k) - N^{k,W^k}(T_{n-1}^k))^2 \mathbb{1}_{\{T_n^k \leq T\}} < \infty,$$

so that

$$\mathbb{E}[[Z^k, A^{k,j}], [Z^k, A^{k,j}]](T) \leq \epsilon_k^2 \mathbb{E}[Z^k, Z^k](T) \leq \epsilon_k^2 \sup_{r \geq 1} \mathbb{E}[Z^r, Z^r](T) \rightarrow 0$$

as  $k \rightarrow \infty$ . Therefore,  $\lim_{k \rightarrow \infty} [Z^k, A^{k,j}] = 0$  strongly in  $\mathbf{B}^2(\mathbb{F})$  so that Lemma 4.6 yields  $\lim_{k \rightarrow \infty} Z^k = \lim_{k \rightarrow \infty} (W^k - M^{k,W}) = 0$  weakly in  $\mathbf{B}^2(\mathbb{F})$ . The set  $\{M^{k,W}; k \geq 1\}$  is  $\mathbf{B}^2(\mathbb{F})$ -weakly relatively sequentially compact where all limits points are square-integrable  $\mathbb{F}$ -martingales over  $[0, T]$ . The weak convergence  $\lim_{k \rightarrow \infty} \mathbb{F}^k = \mathbb{F}$  yields  $\lim_{k \rightarrow \infty} W^k = X$  strongly in  $\mathbf{B}^1(\mathbb{F})$ . This allows us to conclude  $\lim_{k \rightarrow \infty} M^{k,W} = W$  weakly in  $\mathbf{B}^2(\mathbb{F})$ . As a consequence, we apply Lemma 4.3 to state that  $\lim_{k \rightarrow \infty} [M^{k,W}, A^{k,j}](t) = [W, B^j](t)$  weakly in  $L^1(\mathbb{P})$  for each  $t \in [0, T]$  and for every  $j = 1, \dots, d$ .  $\square$

At this point, we are finally able to complete the proof of Theorem 4.2.

**COROLLARY 4.3.** *Under conditions of Theorem 4.2, the canonical imbedded discrete structure  $\mathcal{Y} = ((\delta^k X)_{k \geq 1}, \mathcal{D})$  is stable for every  $X \in \mathcal{R}(\mathbb{F})$  so that  $\mathcal{R}(\mathbb{F}) \subset \mathcal{W}(\mathbb{F})$ . Moreover,*

$$(X, B^j)^{\mathcal{Y}}(t) = [X, B^j](t) \quad \text{a.s., } 0 \leq t \leq T, j = 1, \dots, d.$$

**PROOF.** Let  $X = X(0) + M + N$  be the Dirichlet decomposition of  $X$  where by Brownian motion predictable representation, we can select  $M = \sum_{j=1}^d M^j$  where  $M^j = \int H_j dB^j; j = 1, \dots, d$  for  $H_j \in L_a^2(\mathbb{P} \times \text{Leb})$ . By definition,  $\delta^k X = X(0) + \delta^k M + \delta^k N$  and from Lemma 3.1, we know that  $\lim_{k \rightarrow \infty} \delta^k X = X$  weakly in  $\mathbf{B}^2(\mathbb{F})$ . By applying Lemma 4.7 for the martingale  $M$ , for each  $j = 1, \dots, d$ , we have

$$\begin{aligned} (X, B^j)^{\mathcal{Y}}(t) &= \lim_{k \rightarrow \infty} [\delta^k M, A^{k,j}](t) + \lim_{k \rightarrow \infty} [\delta^k N, A^{k,j}](t) \\ (4.27) \quad &= [M^j, B^j](t) + \lim_{k \rightarrow \infty} [\delta^k N, A^{k,j}](t); \quad 0 \leq t \leq T, \end{aligned}$$

weakly in  $L^1(\mathbb{P})$  as long as the second component in the right-hand side of (4.27) converges. Let us check the  $L^1$ -weak convergence of  $[\delta^k N, A^{k,j}](t)$ . By Proposition 4.3, we have

$$\mathcal{E}^{2,\mathcal{Y}}(X) = \sup_{k \geq 1} \mathbb{E} \sum_{n=1}^{\infty} (\Delta \delta^k X(T_n^k))^2 \mathbb{1}_{\{T_n^k \leq T\}} < \infty.$$

By applying Kunita–Watanabe inequality, we get for each  $t \in [0, T]$ :

$$(4.28) \quad \begin{aligned} \mathbb{E} |[\delta^k N, A^{k,j}](t)| &\leq (\mathbb{E}[\delta^k N, \delta^k N](t))^{1/2} \times (\mathbb{E}[A^{k,j}, A^{k,j}](t))^{1/2} \\ &\leq C (\mathbb{E}[\delta^k N, \delta^k N](t))^{1/2} \rightarrow 0 \end{aligned}$$

as  $k \rightarrow \infty$ , where  $C = (\max_{1 \leq j \leq p} \sup_{k \geq 1} \mathbb{E}[A^{k,j}, A^{k,j}](T))^{1/2} < \infty$ . In (4.28), we observe  $\lim_{k \rightarrow \infty} \mathbb{E}[\delta^k N, \delta^k N](t) = 0$  due to (4.15) and the fact that

$$\lim_{|\tau| \rightarrow 0} \mathbb{E}(Q_\tau^2(N)) = 0.$$

From (4.27) and (4.28), we have  $\langle X, B^j \rangle^{\mathcal{Y}} = [M^j, B^j] = [X, B^j]$ ;  $j = 1, \dots, d$ . □

4.4. *Itô processes and BSDEs.* In this section, we examine the connection to Itô processes and backward SDEs (henceforth abbreviated to BSDEs). The key point is the existence of the limit

$$\lim_{k \rightarrow +\infty} \mathbb{U}^{\mathcal{Y},k,j} X; \quad j = 1, \dots, d.$$

This type of question is related to some regularity properties of the non-martingale component associated with a weakly differentiable process  $X$  in the sense of Definition 4.4; see also Lemma 4.1.

**THEOREM 4.3.** *If there exists an imbedded discrete structure  $\mathcal{Y}$  for  $X \in \mathbf{B}^2(\mathbb{F})$  such that*

$$(4.29) \quad \mathcal{E}^{2,\mathcal{Y}}(X) < \infty$$

and

$$(4.30) \quad \{ \mathbb{U}^{\mathcal{Y},k,j} X; 1 \leq j \leq d, k \geq 1 \} \text{ is uniformly integrable in } L_a^1(\text{Leb} \times \mathbb{P}),$$

then  $X$  is an Itô process. In particular, any square-integrable Itô process  $X$  is uniquely written in the following differential form:

$$(4.31) \quad X(t) = X(0) + \sum_{j=1}^d \int_0^t \mathcal{D}_j X(s) dB^j(s) + \int_0^t \mathcal{U}X(s) ds; \quad 0 \leq t \leq T,$$

where

$$(4.32) \quad UX := \mathcal{U}^{\mathcal{Y}} X = \lim_{k \rightarrow \infty} \sum_{j=1}^d \mathbb{U}^{\mathcal{Y},k,j} X \quad \text{weakly in } L_a^1(\mathbb{P} \times \text{Leb})$$

for every stable discrete structure  $\mathcal{Y} = ((X^k)_{k \geq 1}, \mathcal{D})$  associated with  $X$  satisfying (4.29) and (4.30).

PROOF. Throughout this proof,  $C$  is a generic constant which may defer from line to line. Let  $X \in \mathbf{B}^2(\mathbb{F})$  be a Wiener functional admitting an imbedded discrete structure  $\mathcal{Y} = ((X^k)_{k \geq 1}, \mathcal{D})$  satisfying (4.29) and (4.30). Then  $\mathbb{D}^{\mathcal{Y},k,j} X; j = 1, \dots, d$  and  $\sum_{i=1}^d \mathbb{U}^{\mathcal{Y},k,i} X$  are weakly relatively sequentially compact sequences in  $L_a^2(\mathbb{P} \times \text{Leb})$  and  $L_a^1(\mathbb{P} \times \text{Leb})$ , respectively. Then we shall extract common weakly convergent subsequences. With a slight abuse of notation, we still denote them by  $\mathbb{D}^{\mathcal{Y},k,j} X; j = 1, \dots, d$  and  $\sum_{i=1}^d \mathbb{U}^{\mathcal{Y},k,i} X$ . By applying the same argument used in the proof of Theorem 4.2 along the convergent subsequence  $\{\mathbb{D}^{\mathcal{Y},k,j} X; k \geq 1\}$ , there exists a vector of adapted processes  $H_j \in L_a^2(\mathbb{P} \times \text{Leb}); j = 1, \dots, d$  and there exists  $N \in \mathbf{B}^2(\mathbb{F})$  such that

$$X(t) = X(0) + \sum_{j=1}^d \int_0^t H_j(s) dB^j(s) + N(t); \quad 0 \leq t \leq T.$$

We claim that  $N(t) = \int_0^t \gamma(s) ds; 0 \leq t \leq T$  where  $\gamma := \lim_{k \rightarrow \infty} \sum_{i=1}^d \mathbb{U}^{\mathcal{Y},k,i} X$  weakly in  $L_a^1(\mathbb{P} \times \text{Leb})$ . From Proposition 3.1, we already know that

$$\lim_{k \rightarrow \infty} \sum_{j=1}^d \int_0^t \mathbb{U}^{\mathcal{Y},k,j} X(s) d\langle A^{k,j}, A^{k,j} \rangle(s) = N(\cdot)$$

weakly in  $\mathbf{B}^2(\mathbb{F})$ . By construction,  $N$  is  $\mathbb{F}$ -adapted and it has continuous paths. Hence, in order to show that  $N$  and  $\int \gamma(s) ds$  are indistinguishable, one only has to check they are modifications from each other. It is sufficient to check for a given  $g \in L^\infty(\mathcal{F}_T - \text{measurable})$  and  $0 \leq t \leq T$ ,

$$(4.33) \quad \mathbb{E}gN(t) = \mathbb{E}g \int_0^t \gamma(s) ds.$$

But this is obvious. Indeed, by the very definition and the uniqueness of the weak limit we have

$$\begin{aligned} \mathbb{E}g \int_0^t \sum_{j=1}^d \mathbb{U}^{\mathcal{Y},k,j} X(s) d\langle A^{k,j}, A^{k,j} \rangle(s) &= \mathbb{E}g \int_0^t \sum_{j=1}^d \mathbb{U}^{\mathcal{Y},k,j} X(s) ds \\ &\rightarrow \mathbb{E}g \int_0^t \gamma(s) ds = \mathbb{E}gN(t). \end{aligned}$$

This proves (4.33). Now let us check the second part of the theorem. Let us assume that  $X$  is a square-integrable Itô process of the form

$$X(t) = X(0) + \sum_{j=1}^d \int_0^t H_j(s) dB^j(s) + \int_0^t V(s) ds; \quad 0 \leq t \leq T.$$

Let us now check that  $(\mathcal{D}X, \mathcal{U}X)$  exists. As a strong Dirichlet process, from Theorem 4.2, we already know that  $X \in \mathcal{R}(\mathbb{F})$  is weakly differentiable and  $\mathcal{D}X = (H_1, \dots, H_d)$  in  $L^2_a(\mathbb{F} \times Leb)$ .

CLAIM.  $\mathcal{U}X = V$ .

Let us consider an arbitrary stable imbedded discrete structure  $\mathcal{Y} = ((X^k)_{k \geq 1}, \mathcal{D})$  associated with  $X$ . To shorten notation, let us denote  $M := \sum_{j=1}^d \int H_j dB^j$  and  $Y := \int V(s) ds$ . The  $\mathbb{F}^k$ -semimartingale decomposition based on  $\mathcal{Y}$  is

$$\begin{aligned} X^k(t) &= X^k(0) + \sum_{j=1}^d \oint_0^t \mathbb{D}^{\mathcal{Y},k,j} X(s) dA^{k,j}(s) \\ &\quad + \sum_{j=1}^d \int_0^t U^{\mathcal{Y},k,j} X(s) d\langle A^{k,j}, A^{k,j} \rangle(s) \end{aligned}$$

for  $0 \leq t \leq T$ . Because  $\mathcal{Y}$  is stable, then we shall use Proposition 3.1 to state that

$$\sum_{j=1}^d \int_0^t U^{\mathcal{Y},k,j} X(s) d\langle A^{k,j}, A^{k,j} \rangle(s) \rightarrow Y$$

weakly in  $\mathbf{B}^2(\mathbb{F})$  as  $k \rightarrow \infty$ . In particular, by taking  $g \in L^\infty(\mathcal{F}_T)$  and  $t \in [0, T]$ , we shall consider the bounded linear functional  $S = g \mathbb{1}_{[0,t]} \in \mathbf{M}^2(\mathbb{F})$  to get

$$\begin{aligned} \mathbb{E}g \int_0^t \sum_{j=1}^d \mathbb{U}^{\mathcal{Y},k,j} X(s) ds &= \mathbb{E}g \int_0^t \sum_{j=1}^d U^{\mathcal{Y},k,j} X(s) d\langle A^{k,j}, A^{k,j} \rangle(s) \\ &\rightarrow \mathbb{E}g \int_0^t V(s) ds \end{aligned}$$

as  $k \rightarrow \infty$  for each for  $g \in L^\infty(\mathcal{F}_T)$  and  $t \in [0, T]$ . Hence, (4.31) holds true and we conclude the proof.  $\square$

Of course, it is already known that any Itô process

$$X = X(0) + \sum_{j=1}^d \int H_j dB^j + \int Z ds$$

is completely characterized by  $(H_1, \dots, H_d, Z)$ . The main message of Theorem 4.3 is that

$$(H_1, \dots, H_d, Z) = (\mathcal{D}_1 X, \dots, \mathcal{D}_d X, \mathcal{U}X),$$

where  $(\mathcal{D}_1 X, \dots, \mathcal{D}_d X, \mathcal{U}X)$  can be intrinsically constructed by means of any stable imbedded discrete structure  $\mathcal{Y}$  satisfying (4.29) and (4.30).

The differential operator  $\mathcal{U}X$  basically describes the mean of any square-integrable Itô process in an infinitesimal time interval

$$\mathbb{E}X(t) \sim \mathbb{E}X(0) + t\mathbb{E}\mathcal{U}X(t) \quad \text{for small } t > 0.$$

Of course, when  $X(\cdot) = g(\cdot, W(\cdot))$  is a smooth transformation of a Markovian diffusion  $W$ , then  $\mathcal{U}X(t) = \partial_t g(t, W(t)) + \mathcal{L}g(t, W(t))$ , where  $\mathcal{L}$  is the infinitesimal generator of  $W$ . This justifies the following definition.

**DEFINITION 4.6.** We say that a Wiener functional  $X \in \mathcal{W}(\mathbb{F})$  admits a stochastic infinitesimal generator if  $\mathcal{U}X$  exists.

We advocate the existence of the infinitesimal generator as a “heat-type operator” requires strong pathwise regularity in the sense of [8, 19].

**REMARK 4.5.** If  $F$  has pathwise  $\mathbb{C}^{1,2}(\Lambda)$ -regularity (in the sense of [19]) and  $X = F(Z)$  for a continuous  $\mathbb{F}$ -semimartingale  $Z$ , then

$$(4.34) \quad \nabla^h F(Z) + \frac{1}{2} \operatorname{tr} \nabla^{v,2} F(Z) \frac{d[Z, Z]}{dt} = \mathcal{U}X = \mathcal{U}^{\mathcal{Y}} X$$

for every stable imbedded discrete structure  $\mathcal{Y}$  w.r.t.  $X = F(Z)$ , where  $\nabla^h$  and  $\nabla^{v,2}$  denote the horizontal and second order vertical derivatives in the sense of pathwise calculus. The existence of the splitting on the left-hand side of (4.34) requires severe regularity either because  $\mathcal{U}X$  may not exist or one of the functionals  $\nabla^h F$  and  $\operatorname{tr} \nabla^{v,2} F$  may not exist. In fact, we advocate the important object is  $\mathcal{U}^{\mathcal{Y},k} X$  where  $\mathcal{Y}$  ranges over all imbedded discrete structures for a given possibly nonsmooth  $X$ ; see [32, 35] for details.

Under strong regularity conditions, we shall provide a local characterization of  $\mathcal{U}X$ ; see Proposition 5.1 in the Supplementary Material [33].

**4.5. Variational representation of BSDEs.** Let us now briefly illustrate the role of  $(DX, \mathcal{U}X)$  in the BSDEs. Let  $\mathbb{H}$  be the Cameron–Martin space associated with the one-dimensional Wiener measure and let  $L_a^2(\Omega; \mathbb{H})$  be the  $\mathbb{P}$ -equivalent class of  $\mathbb{H}$ -valued random variables such that

$$\mathbb{E} \|u\|_{\mathbb{H}}^2 := \mathbb{E} \int_0^T |\dot{u}(s)|^2 ds < \infty,$$

where the Radon–Nikodym derivative  $\dot{u}$  is  $\mathbb{F}$ -adapted. Let us consider a well-posed BSDE (see [39]):

$$(4.35) \quad \begin{aligned} Y(t) = & \xi + \int_t^T g(r, Y(r), Z(r)) dr \\ & - \sum_{j=1}^d \int_t^T Z^j(r) dB^j(r); \quad 0 \leq t \leq T, \end{aligned}$$

where  $g : \Omega \times [0, T] \times \mathbb{R} \times \mathbb{R}^d \rightarrow \mathbb{R}$  is the generator of the BSDE and  $\xi \in L^2(\mathcal{F}_T)$ . A strong solution of the BSDE (4.35) is an  $(\mathbb{R} \times \mathbb{R}^d)$ -valued  $\mathbb{F}$ -adapted process  $(Y, Z)$  which satisfies (4.35) almost surely. For a given  $\xi \in L^2(\mathcal{F}_T)$  and under suitable technical assumptions on  $g$ , it is well known there exists a unique solution  $(Y, Z)$ .

Now, for a given pair  $(g, \xi)$ , let  $Y$  be a square-integrable Itô process such that  $Y(T) = \xi$  a.s. and we set

$$\begin{aligned} \Lambda(Y, g, \xi)(t) & := Y(t) - Y(0) + \int_0^t g(s, \mathcal{D}Y(s), Y(s)) ds, \\ \mathbb{Y}^{(g, \xi)}(t) & := \int_0^t \Lambda(Y, g, \xi)(s) ds; \quad 0 \leq t \leq T. \end{aligned}$$

In the sequel, when we write  $\mathbb{Y}^{(g, \xi)}$  it is implicitly assumed that  $Y$  is a square-integrable Itô process such that  $Y(T) = \xi$  where  $(g, \xi)$  is given.

**THEOREM 4.4.** *Let  $\xi \in L^2(\mathcal{F}_T)$  be a fixed terminal condition. A pair  $(Y, Z)$  is a strong solution of (4.35), if and only if  $Y \in \mathcal{W}(\mathbb{F})$ ,  $Z = \mathcal{D}Y$  and*

$$(4.36) \quad \begin{cases} \mathcal{U}Y(t) + g(t, Y(t), \mathcal{D}Y(t)) = 0; & 0 \leq t < T, \text{ a.s.} \\ Y(T) = \xi & \text{a.s.} \end{cases}$$

*In particular,  $Y(0) = \mathbb{E}[\xi] - \mathbb{E} \int_0^T \mathcal{U}Y(s) ds$ . Moreover, for a given pair  $(g, \xi)$ , a square-integrable Itô process  $Y$  is a solution of (4.36), if and only if*

$$(4.37) \quad \mathbb{Y}^{(g, \xi)} \in \arg \min_{X \in L^2_a(\Omega; \mathbb{H}); X(T) = \mathbb{Y}^{(g, \xi)}(T)} \mathbb{E} \|X\|_{\mathbb{H}}^2,$$

*or, in other words,*

$$(4.38) \quad \sum_{j=1}^d \mathbb{E} \int_0^T \int_0^s |\mathcal{D}_j Y(r)|^2 dr ds = \min_{X \in L^2_a(\Omega; \mathbb{H}); X(T) = \mathbb{Y}^{(g, \xi)}(T)} \mathbb{E} \int_0^T |\dot{X}(s)|^2 ds.$$

**PROOF.** See the Supplementary Material [33].  $\square$

**Acknowledgments.** The authors are grateful to several referees and to an Associate Editor for their careful reading of the first versions of the manuscript which has allowed them to considerably improve the quality of the paper. A. Ohashi would like to thank UMA-ENSTA-ParisTech for the very kind hospitality during the last stage of this project as well as Francesco Russo for stimulating discussions and suggestions that helped shaping the theory.

## SUPPLEMENTARY MATERIAL

**Supplement to “A weak version of path-dependent functional Itô calculus”** (DOI: [10.1214/17-AOP1250SUPP](https://doi.org/10.1214/17-AOP1250SUPP); .pdf). The proofs of Lemmas 2.1, 2.3, 3.2, 4.4 and Theorem 4.4 are provided in the online supplement [33].

## REFERENCES

- [1] BERTOIN, J. (1986). Les processus de Dirichlet en tant qu’espace de Banach. *Stochastics* **18** 155–168. [MR0861819](#)
- [2] BERTOIN, J. (1989). Sur une intégrale pour les processus à  $\alpha$ -variation bornée. *Ann. Probab.* **17** 1521–1535. [MR1048943](#)
- [3] BEZERRA, S. C., OHASHI, A. and RUSSO, F. (2018). Discrete-type approximations for non-Markovian optimal stopping problems: Part II. Available at [arXiv:1707.02520](https://arxiv.org/abs/1707.02520).
- [4] BUCKDAHN, R., MA, J. and ZHANG, J. (2015). Pathwise Taylor expansions for random fields on multiple dimensional paths. *Stochastic Process. Appl.* **125** 2820–2855. [MR3332856](#)
- [5] BURQ, Z. A. and JONES, O. D. (2008). Simulation of Brownian motion at first-passage times. *Math. Comput. Simulation* **77** 64–71. [MR2388251](#)
- [6] COHEN, S. N. and ELLIOTT, R. J. (2015). *Stochastic Calculus and Applications*, 2nd ed. Springer, Cham. [MR3443368](#)
- [7] CONT, R. and FOURNIÉ, D.-A. (2010). Change of variable formulas for non-anticipative functionals on path space. *J. Funct. Anal.* **259** 1043–1072. [MR2652181](#)
- [8] CONT, R. and FOURNIÉ, D.-A. (2013). Functional Itô calculus and stochastic integral representation of martingales. *Ann. Probab.* **41** 109–133. [MR3059194](#)
- [9] CONT, R. and FOURNIÉ, D. A. (2016). Functional Itô calculus and functional Kolmogorov equations. In *Stochastic Integration by Parts and Functional Itô Calculus* 115–207. Birkhäuser/Springer, Cham. [MR3497715](#)
- [10] COQUET, F., MACKEVIČIUS, V. and MÉMIN, J. (1998). Stability in  $\mathbf{D}$  of martingales and backward equations under discretization of filtration. *Stochastic Process. Appl.* **75** 235–248. [MR1632205](#)
- [11] COQUET, F., MÉMIN, J. and SŁOMIŃSKI, L. (2001). On weak convergence of filtrations. In *Séminaire de Probabilités, XXXV. Lecture Notes in Math.* **1755** 306–328. Springer, Berlin. [MR1837294](#)
- [12] COQUET, F. and SŁOMIŃSKI, L. (1999). On the convergence of Dirichlet processes. *Bernoulli* **5** 615–639. [MR1704558](#)
- [13] COSSO, A. and RUSSO, F. (2015). Strong-viscosity solutions: Semilinear parabolic PDEs and path-dependent PDEs. Available at [arXiv:1505.02927](https://arxiv.org/abs/1505.02927).
- [14] COSSO, A. and RUSSO, F. (2016). Functional Itô versus Banach space stochastic calculus and strict solutions of semilinear path-dependent equations. *Infin. Dimens. Anal. Quantum Probab. Relat. Top.* **19** Art. ID 1650024. [MR3584624](#)

- [15] COSSO, A. and RUSSO, F. (2016). Functional and Banach space stochastic calculi: Path-dependent Kolmogorov equations associated with the frame of a Brownian motion. In *Stochastics of Environmental and Financial Economics—Centre of Advanced Study, Oslo, Norway, 2014–2015* (F. E. Benth and G. Di Nunno, eds.). *Springer Proc. Math. Stat.* **138** 27–80. Springer, Cham. [MR3451168](#)
- [16] DELLACHERIE, C. and MEYER, P.-A. (1975). *Probabilités et Potentiel. Publications de l’Institut de Mathématique de l’Université de Strasbourg XV. Actualités Scientifiques et Industrielles* **1372**. Hermann, Paris. Chapitres I à IV, Édition entièrement refondue. [MR0488194](#)
- [17] DELLACHERIE, C. and MEYER, P.-A. (1982). *Probabilities and Potential B: Theory of Martingales. North-Holland Mathematics Studies* **72**. North-Holland, Amsterdam. Translated from the French by J. P. Wilson. [MR0745449](#)
- [18] DIESTEL, J., RUESS, W. M. and SCHACHERMAYER, W. (1993). On weak compactness in  $L^1(\mu, X)$ . *Proc. Amer. Math. Soc.* **118** 447–453. [MR1132408](#)
- [19] DUPIRE, B. Functional Itô calculus. Portfolio Research Paper 2009-04, Bloomberg.
- [20] EKREN, I., KELLER, C., TOUZI, N. and ZHANG, J. (2014). On viscosity solutions of path dependent PDEs. *Ann. Probab.* **42** 204–236. [MR3161485](#)
- [21] EKREN, I., TOUZI, N. and ZHANG, J. (2016). Viscosity solutions of fully nonlinear parabolic path dependent PDEs: Part I. *Ann. Probab.* **44** 1212–1253. [MR3474470](#)
- [22] EKREN, I., TOUZI, N. and ZHANG, J. (2016). Viscosity solutions of fully nonlinear parabolic path dependent PDEs: Part II. *Ann. Probab.* **44** 2507–2553. [MR3531674](#)
- [23] EKREN, I. and ZHANG, J. (2016). Pseudo-Markovian viscosity solutions of fully nonlinear degenerate PPDEs. *Probab. Uncertain. Quant. Risk* **1** Paper No. 6. [MR3583183](#)
- [24] FLANDOLI, F. and ZANCO, G. (2016). An infinite-dimensional approach to path-dependent Kolmogorov equations. *Ann. Probab.* **44** 2643–2693. [MR3531677](#)
- [25] HE, S. W., WANG, J. G. and YAN, J. A. (1992). *Semimartingale Theory and Stochastic Calculus*. Kexue Chubanshe (Science Press), Beijing; CRC Press, Boca Raton, FL. [MR1219534](#)
- [26] JACOD, J. and SKOROHOD, A. V. (1994). Jumping filtrations and martingales with finite variation. In *Séminaire de Probabilités, XXVIII. Lecture Notes in Math.* **1583** 21–35. Springer, Berlin. [MR1329098](#)
- [27] KELLER, C. and ZHANG, J. (2016). Pathwise Itô calculus for rough paths and rough PDEs with path dependent coefficients. *Stochastic Process. Appl.* **126** 735–766. [MR3452811](#)
- [28] KHOSHNEVISAN, D. and LEWIS, T. M. (1999). Stochastic calculus for Brownian motion on a Brownian fracture. *Ann. Appl. Probab.* **9** 629–667. [MR1722276](#)
- [29] LE JAN, Y. (1978). Temps d’arrêt stricts et martingales de sauts. *Z. Wahrsch. Verw. Gebiete* **44** 213–225.
- [30] LEÃO, D. and OHASHI, A. (2013). Weak approximations for Wiener functionals. *Ann. Appl. Probab.* **23** 1660–1691. [MR3098445](#)
- [31] LEÃO, D. and OHASHI, A. (2017). Corrigendum: Weak approximations for Wiener functionals [*Ann. Appl. Probab.* (2013) 23 1660–1691] [[MR3098445](#)]. *Ann. Appl. Probab.* **27** 1294–1295. [MR3655867](#)
- [32] LEÃO, D., OHASHI, A. and RUSSO, F. (2018). Discrete-type approximations for non-Markovian optimal stopping problems: Part I. Available at [arXiv:1707.05234](#).
- [33] LEÃO, D., OHASHI, A. and SIMAS, A. B. (2018). Supplement to “A weak version of path-dependent functional Itô calculus.” DOI:[10.1214/17-AOP1250SUPP](#).
- [34] LEÃO, D., OHASHI, A. and SIMAS, A. B. (2018). Weak differentiability of Wiener functionals and occupation times. *Bull. Sci. Math.* To appear. Available at [arXiv:1711.10895](#).
- [35] LEÃO, D., OHASHI, A. and SOUZA, F. (2018). Stochastic near-optimal controls for path-dependent systems. Available at [arXiv:1707.04976](#).
- [36] MÉMIN, J. (2003). Stability of Doob–Meyer decomposition under extended convergence. *Acta Math. Appl. Sin. Engl. Ser.* **19** 177–190. [MR2011481](#)

- [37] OBERHAUSER, H. (2016). The functional Itô formula under the family of continuous semimartingale measures. *Stoch. Dyn.* **16** Art. ID 1650010. [MR3494685](#)
- [38] OHASHI, A., SHAMAROVA, E. and SHAMAROV, N. N. (2016). Path-dependent Itô formulas under  $(p, q)$ -variations. *ALEA Lat. Am. J. Probab. Math. Stat.* **13** 1–31. [MR3460868](#)
- [39] PARDOUX, É. and PENG, S. G. (1990). Adapted solution of a backward stochastic differential equation. *Systems Control Lett.* **14** 55–61. [MR1037747](#)
- [40] PENG, S. (2010). Backward stochastic differential equation, nonlinear expectation and their applications. In *Proceedings of the International Congress of Mathematicians, Vol. I* 393–432. Hindustan Book Agency, New Delhi. [MR2827899](#)
- [41] PENG, S. and SONG, Y. (2015).  $G$ -expectation weighted Sobolev spaces, backward SDE and path dependent PDE. *J. Math. Soc. Japan* **67** 1725–1757. [MR3417511](#)
- [42] PENG, S. and WANG, F. (2016). BSDE, path-dependent PDE and nonlinear Feynman–Kac formula. *Sci. China Math.* **59** 19–36. [MR3436993](#)

D. LEÃO  
DEPARTAMENTO DE MATEMÁTICA  
APLICADA E ESTATÍSTICA  
UNIVERSIDADE DE SÃO PAULO  
13560-970, SÃO CARLOS—SP  
BRAZIL  
E-MAIL: [leao@estatcamp.com.br](mailto:leao@estatcamp.com.br)

A. OHASHI  
A. B. SIMAS  
DEPARTAMENTO DE MATEMÁTICA  
UNIVERSIDADE FEDERAL DA PARAÍBA  
13560-970, JOÃO PESSOA—PARAÍBA  
BRAZIL  
E-MAIL: [amfohashi@gmail.com](mailto:amfohashi@gmail.com)  
[alexandre@mat.ufpb.br](mailto:alexandre@mat.ufpb.br)