

# Explicit Form and Path Regularity of Martingale Representations

Jin Ma\*, Philip Protter<sup>†</sup>, and Jianfeng Zhang<sup>‡</sup>

**Abstract.** Let  $X$  be the solution of a stochastic differential equation driven by a Wiener process and a compensated Poisson random measure, such that  $X$  is an  $L^2$  martingale. If  $H = \Phi(X_s; 0 \leq s \leq T)$  is in  $L^2$ , then

$$H = \alpha + \int_0^T \xi_s dX_s + N_T,$$

where  $N$  is an  $L^2$  martingale orthogonal to  $X$  (the Kunita-Watanabe decomposition). We give sufficient conditions on the functional  $\Phi$  such that  $\xi$  has *regular* paths (that is, left continuous with right limits). In Finance this has an interpretation as that the risk minimizing hedging strategy of a contingent claim in an incomplete market has “smooth” regular sample paths. This means the hedging process can be approximated and the resulting approximations will converge, along the sample paths, to the risk minimal (and hence optimal) portfolio.

**Keywords:** Incomplete markets, hedging strategies, Lévy processes, minimal martingale measures, Kunita-Watanabe decomposition, martingale representation, stochastic differential equations, Meyer-Zheng topology.

*1991 Mathematics Subject Classification.* Primary: 60H10, Secondary: 34F05, 90A12

---

\*Department of Mathematics, Purdue University, West Lafayette, IN 47907-1395; email: ma-jin@math.purdue.edu. This author is supported in part by NSF grant #9971720.

<sup>†</sup>Departments of Mathematics and Statistics, Purdue University, West Lafayette, IN 47907-1395; and ORIE, Cornell University, Ithaca, NY 14853; email: protter@orie.cornell.edu. This author is supported in part by NSF grants #9971720, #9401109-INT, and NSA grant #MDA-904-00-1-0035.

<sup>‡</sup>Department of Mathematics, Purdue University, West Lafayette, IN 47907-1395; email: jfzhang@math.purdue.edu. This author is supported in part by NSF grant #9971720 and Purdue Research Foundation Summer Grant 2000.

# 1 Introduction

**A. Background.** The problems addressed in this article are motivated by questions arising in Financial Asset Pricing Theory where the market is not complete. The framework is as follows: let  $X = (X_t)_{t \geq 0}$  be a semimartingale representing the price process of a risky asset. Under the standard assumption of the absence of arbitrage opportunities, there exists a probability measure  $P^*$ , equivalent to the original probability measure  $P$  (the “objective” probability), such that  $X$  is a  $P^*$ -local martingale (technically one requires  $X$  only to be a  $P^*$ -sigma-martingale ; see [5]).  $P^*$  is known as the risk neutral measure.

Let us assume that  $X$  is in fact a  $P^*$ -martingale in  $L^2$ , for  $0 \leq t \leq T$ , as is often the case. For a non-redundant contingent claim  $H \in L^2(\mathcal{F}_T, dP^*)$  we have a unique decomposition:

$$H = \alpha + \int_0^T \xi_s^H dX_s + N_T, \quad (1.1)$$

where  $N$  is an  $L^2(dP^*)$  martingale strongly orthogonal to  $X$ . (The decomposition (1.1) is called the Kunita-Watanabe  $L^2$ -martingale decomposition; see [4] or [13] for background.) Let  $(\eta_t)_{0 \leq t \leq T}$  be an (optional) strategy devoted to the trading of a risk-free savings account, whose price is fixed at 1. The *value* of the portfolio at time  $t$  is then

$$V_t = \xi_t X_t + \eta_t,$$

and the *cost* up to time  $t$  is

$$C_t = V_t - \int_0^t \xi_s dX_s.$$

We require  $V_T = H$ . A strategy  $(\xi, \eta)$  is *self-financing* if  $(C_t)_{t \geq 0}$  is constant, and it is *mean-self-financing* if  $E\{C_T - C_t | \mathcal{F}_t\} = 0$ ; that is, if  $C$  is a martingale. If we wish to minimize the remaining risk after time  $t$ , we then wish to minimize the quantity

$$E\{(C_T - C_t)^2 | \mathcal{F}_t\}, \quad (1.2)$$

interpreting risk in the  $L^2$ , or “squared error”, sense. H. Föllmer and M. Schweizer [6] have shown that the strategy

$$(\xi_t, \eta_t) = (\xi_t^H, V_t - \int_0^t \xi_s^H dX_s)$$

is *optimal* in the sense that it minimizes the risk quantity (1.2).

Therefore, at least in this special case where  $X$  is an  $L^2$ -martingale under the risk-neutral measure and  $H \in L^2(\mathcal{F}_T, dP^*)$ , the hedging strategy  $\xi^H$  of (1.1), while of necessity not perfect replication, is nevertheless optimal under squared error loss.

Several issues arise immediately: (1) when are there formulae to describe  $\xi^H$  analogous to those available in the traditional Black-Scholes paradigm? (2) when formulae are *not* available, what can one infer about  $\xi^H$ ? In particular, when can one be assured of path regularity of  $\xi^H$ ?

Issue (1) above is addressed in [8], where it is shown that if there is an underlying quasi-left continuous strong Markov process  $Y = (\Omega, \mathcal{F}, (\mathcal{F}_t), P_y^*, Y)$  and if  $X$  is an  $L^2$ -martingale under each  $P_y^*$  and if  $H = g(Y_T)$  for an appropriate class of functions  $g$ , then there is an explicit formula for  $\xi$ . (Note that “explicit” in the preceding sentence can mean different things to different people.) The results from [8] are perhaps the most interesting when  $Y = X$  and  $X$  is the solution of a stochastic differential equation driven by a Lévy martingale.

Issue (2) above is the topic of this paper. For most contingent claims it is not possible to obtain explicit formulas for  $\xi^H$ . Instead here we are concerned with when the processes  $\xi^H$ —which are *a priori* assumed to be only predictably measurable — have regular sample paths. In particular by “regular” we mean that it is at least *left continuous with right limits*, known by its French acronym *càglàd*. When  $\xi^H$  can be shown to have *càglàd* paths it is useful for two reasons: (a) approximations of  $\xi^H$  will converge to it in a Skorohod-type topology; and (b) one can approximate  $\int \xi_s^H dX_s$  with Riemann-type sums and have convergence uniformly in probability (or even almost surely if the partition size shrinks quickly). The importance of part (b) in Finance has been emphasized, for example, in [14].

**B. New Results.** We assume that under the risk neutral measure  $P^*$  we have a Wiener process  $W$  and a compensated Poisson random measure  $\tilde{\mu}(drdz) = \mu(dr dz) - drF(dz)$ , where  $F$  is a Lévy measure. We further assume that  $\int_{\mathbb{R}} z^2 F(dz) < \infty$ , so that the process

$$Z_t = W_t + \int_0^t \int_{\mathbb{R}} z \tilde{\mu}(dr dz)$$

is a Lévy process with  $E\{Z_t\} = 0$  and  $E\{Z_t^2\} < \infty$ ,  $0 \leq t \leq T$ . (Thus  $Z$  is an

$L^2(dP^*)$  martingale. Our price process  $X$  satisfies

$$X_t = y + \int_0^t \sigma(r, X_r) dW_r + \int_0^t \int_{\mathbb{R}} b(r, X_{r-}) z \tilde{\mu}(dr dz) \quad (1.3)$$

and thus is also an  $L^2(dP^*)$  martingale with mild hypotheses on  $\sigma$  and  $b$ .

A contingent claim  $H \in \mathcal{F}_T$  can be assumed to be of the form  $H = \Phi(X_s; s \leq T)$ , where  $\Phi$  is a functional mapping  $\mathbb{D}$  to  $\mathbb{R}$ , where  $\mathbb{D}$  is the Skorohod space of càdlàg (right continuous with left limits) functions. We find hypotheses on  $\Phi$  such that  $\xi^H$  is càglàd. An hypothesis, which we call the  $L^1$ -Lipschitz condition (see section 2 for details), was used in a recent paper by Ma and Zhang [11] to study the path regularity problem for the solutions to backward SDEs driven by Brownian motions. This paper in a sense extends the result there to the Lévy case. Some examples of path-dependent options covered by our results include, but are not limited to,

- (i)  $\Phi(X)_T = \frac{1}{T} \int_0^T X_s ds$ ;
- (ii)  $\Phi(X)_T = g(\sup_{0 \leq t \leq T} h(t, X_t))$  (Lookback option);
- (iii)  $\Phi(X)_T = g(\int_0^T h(s, X_{s-}) dX_s)$ ; or
- (iv)  $\Phi(X) = g(\Phi_1(X), \dots, \Phi_n(X))$ , where  $g$  is Lipschitz and  $\Phi_i$ 's are of any of the forms (i)–(iii). (For example, if  $g(x) = (K - x)^+$ , then  $g$  combined with (i) gives an Asian Option.)

We remark that to justify the price equation (1.3) one should note that in almost all of the existing theory of Financial Asset Pricing, the price process is assumed to be Markov under the risk neutral measure. (The price process need not be Markov under the objective measure however.) E. Çinlar and J. Jacod [2] have shown that all “reasonable” strong Markov martingale processes are solutions of equations of the form

$$X_t = y + \int_0^t \sigma(r, X_r) dW_r + \int_0^t b(r, X_{r-}, z) \tilde{\mu}(dr dz). \quad (1.4)$$

Thus our assumption merely restricts the general case by assuming  $b(r, x, z)$  is of the form  $b(r, x)z$ , as well as some restrictions on the integrability of the coefficients.

A simple way in which our model might arise is if the objective price process  $X$  is modeled as a geometric Lévy process:

$$dX_t = \sigma_t X_{t-} dZ_t + b_t X_t dt, \quad (1.5)$$

where  $Z$  is an  $L^2$  Lévy martingale under  $P$ . Since Lévy processes give rise to incomplete markets, we have to choose an equivalent risk neutral measure  $P^*$  in a natural

way. We can do this using the idea of Föllmer and Schweizer [6]: an equivalent risk neutral probability  $P^*$  is *minimal* if any square-integrable  $P$ -martingale  $M$  orthogonal to  $Z$  is also a  $P^*$ -martingale. T. Chan [1] has shown, under some restrictive assumptions, that if  $U$  satisfies

$$dU_t = 1 + U_t \gamma_t dW_t,$$

where  $\gamma$  can be taken to be non-random if  $\sigma$  and  $b$  are non-random, then  $dP^* = U_T dP$ , and thus it follows that under this  $P^*$  the process  $X$  of (1.5) satisfies (1.3). That is, the drift is removed by the canonical risk neutral minimal martingale measure, and the price process has the desired form under  $P^*$ .

## 2 Preliminaries

Throughout this paper we assume that  $(\Omega, \mathcal{F}, P; \{\mathcal{F}_t\}_{0 \leq t \leq T})$  is a complete filtered probability space satisfying the *usual hypotheses* (see, e.g., [13]), and  $T > 0$  is a fixed time duration. We denote  $\mathbf{F} \triangleq \{\mathcal{F}_t\}_{t \geq 0}$  and assume  $\mathbf{F}$  is quasi-left continuous. Let  $W$  be an  $\mathbf{F}$ -adapted Brownian motion, and  $\mu$  be a random measure generated by an  $\mathbf{F}$ -adapted Lévy process (see, e.g., [9]). Let  $\nu(dt dz) = dt F(dz)$  be the (non-random) compensator of  $\mu$ , and denote  $\tilde{\mu} = \mu - \nu$ . We assume that  $F$  integrates  $z^2$ , and by rescaling we may assume without loss of generality that  $\int_{\mathbb{R}} z^2 F(dz) = 1$ .

In what follows we denote  $\mathbb{D} \triangleq \mathbb{D}[0, T]$  to be the space of all càdlàg functions on  $[0, T]$ ; and  $C_b^{0,1}([0, T] \times \mathbb{R})$  to be the space of all continuous functions on  $[0, T] \times \mathbb{R}$  that are continuously differentiable with bounded derivatives in the spatial variable  $x$ . To simplify the presentation, we assume all the processes are 1-dimensional, but all the results in this paper can be extended to higher dimensional cases without substantial difficulties.

For  $(s, y) \in [0, T] \times \mathbb{R}$ , let us consider a local martingale,  $\{X_t^{s,y}\}_{s \leq t \leq T}$ , defined as the (unique) solution to the following stochastic differential equation:

$$X_t = y + \int_s^t \sigma(r, X_{r-}) dW_r + \int_s^t \int_{\mathbb{R}} b(r, X_{r-}) z \tilde{\mu}(dr dz), \quad s \leq t \leq T. \quad (2.6)$$

We assume that the coefficients  $\sigma$  and  $b$  satisfy the following standing assumptions:

**(A1)** (i) The functions  $\sigma, b \in C_b^{0,1}([0, T] \times \mathbb{R})$ , such that  $\sigma^2(t, x) \neq 0$ , for all  $(t, x)$ ; and

(ii) There exists a constant  $K > 0$  such that

$$\begin{cases} \sup_{0 \leq t \leq T} [|\sigma(t, 0)| + |b(t, 0)|] \leq K; \\ |\sigma'_x(t, x)| + |b'_x(t, x)| \leq K. \end{cases} \quad (2.7)$$

Here and in the sequel we denote  $\sigma'_x$  and  $b'_x$  to be the partial derivatives of  $\sigma$  and  $b$ , respectively, with respect to the spatial variable  $x$ .

We first give two lemmas. Since the proofs are more or less standard (see, e.g., [8]), we omit them. The first lemma shows that the assumption (A1), together with the requirement that  $\int z^2 F(dz) < \infty$ , renders  $X$  a martingale rather than a local martingale.

**Lemma 2.1** *Assume (A1). Then for any  $0 \leq s < t \leq T$ , there exists a constant  $C > 0$  depending only on  $K$  and  $T$ , such that*

$$E\left\{ \sup_{s \leq t \leq T} |X_t^{s,y}|^2 \right\} \leq C(1 + |y|^2), \quad (2.8)$$

and

$$E\left\{ \sup_{s \leq r \leq t} |X_r^{s,y} - y|^2 \right\} \leq C(1 + |y|^2)(t - s). \quad (2.9)$$

Consequently, for any  $(s, y) \in [0, T] \times \mathbb{R}$ ,  $X^{s,y}$  is a true martingale defined on  $[s, T]$ . Furthermore, one has

$$d\langle X^{s,y}, X^{s,y} \rangle_t = (\sigma^2 + b^2)(t, X_{t-}^{s,y}) dt, \quad t \in [s, T]. \quad (2.10)$$

■

Throughout the paper, unless otherwise specified, we denote  $C > 0$  to be a generic constant depending only on  $K$  and  $T$ , which may vary from line to line. The following *variational* process  $\nabla X$ , defined in Lemma 2.2, is important in the paper.

**Lemma 2.2** *Assume (A1), and let  $\alpha$  and  $\beta$  be two  $\mathbf{F}$ -predictable processes that are bounded by  $K$ . For each  $0 \leq s < T$ , let  $Y^s$  be the solution to the following (linear) SDE:*

$$Y_t^s = 1 + \int_s^t \alpha_r Y_{r-}^s dW_r + \int_s^t \int_{\mathbb{R}} \beta_r Y_{r-}^s z \tilde{\mu}(dr dz). \quad (2.11)$$

Then it holds that

$$E\left\{ \sup_{s \leq t \leq T} |Y_t^s|^2 \right\} \leq C. \quad (2.12)$$

In particular, if  $\alpha_t = \sigma'_x(t, X_t^{s,y})$  and  $\beta_t = b'_x(t, X_t^{s,y})$ , then we denote the solution to (2.11) by  $Y \triangleq \nabla X$ , which satisfies the following relation:

$$E \left\{ \left| \frac{1}{h} (X_t^{s,x+h} - X_t^{s,x}) - \nabla X_t^{s,x} \right|^2 \right\} \rightarrow 0, \quad \text{as } h \rightarrow 0.$$

■

We note that sometimes the notion of “*difference quotient*” of  $X$  will also be used: for any  $y, h \in \mathbb{R}$  and  $t \in [s, T]$ ,

$$\Delta_h X_t^{s,y} \triangleq \frac{1}{h} [X_t^{s,y+h} - X_t^{s,y}].$$

The following identity is then obvious:

$$\sup_{s \leq t \leq T} |X_t^{s,y_1} - X_t^{s,y_2}| = |y_1 - y_2| \sup_{s \leq t \leq T} |\Delta_{y_1 - y_2} X_t^{s,y_2}|. \quad (2.13)$$

Now let  $\Phi : \mathbb{D} \mapsto \mathbb{R}$  be a functional such that  $E|\Phi(X)|^2 < \infty$ . We consider the following  $\mathbf{F}$ -martingale:

$$M_t \triangleq E\{\Phi(X) | \mathcal{F}_t\}, \quad t \geq 0. \quad (2.14)$$

By martingale theory, there exists an  $\mathbf{F}$ -predictable process,  $\xi$ , such that

$$M_t = \alpha + \int_0^t \xi_s dX_s + N_t, \quad (2.15)$$

where  $\alpha = M_0$  and  $N$  is an  $\mathbf{F}$ -martingale that is orthogonal to  $X$  (i.e.,  $[N, X]$  is an  $\mathbf{F}$ -martingale. For more on this theory, consult Dellacherie-Meyer [4] or Protter [13]). One of the main purposes of this paper is to find conditions on  $\Phi$ , so that  $\xi$  has càglàd paths.

To this end, let us introduce a functional  $\varphi$  which plays an important role in the sequel. Define  $\varphi : \mathbb{D} \times [0, T] \times \mathbb{R} \mapsto \mathbb{R}$  as follows

$$\varphi(\mathbf{x}, t, y) = E\{\Phi(\mathbf{x}1_{[0,t)} + X^{t,y}1_{[t,T]})\}, \quad (\mathbf{x}, t, y) \in \mathbb{D} \times [0, T] \times \mathbb{R}. \quad (2.16)$$

We note that if  $\Phi(X) = g(X_T)$ , then  $\varphi(\mathbf{x}, t, y) = P_{T-t}g(y)$ , where  $(P_t)$  is the transition semi-group of the strong Markov process  $X$ .

To conclude this section let us recall a “tightness” result of probability measures of Meyer-Zheng [12]. Denote  $\mathcal{M}(\mathbb{D})$  to be the space of all probability measures on

$\mathbb{D}$ . Let  $X$  be any  $\mathbf{F}$ -adapted, càdlàg process defined on  $[0, T]$ , such that  $E|X_t| < \infty$  for all  $t \geq 0$ . For any partition  $\pi : 0 = t_0 < t_1 < \dots < t_n \leq T$ , let us define

$$V_T^\pi(X) \triangleq \sum_{0 \leq i < n} E\{|E\{X_{t_{i+1}} - X_{t_i} | \mathcal{F}_{t_i}\}|\} + E|X_{t_n}|, \quad (2.17)$$

and define the *conditional variation* of  $X$  by  $V_T(X) \triangleq \sup_\pi V_T^\pi(X)$ . If  $V_T(X) < \infty$ , then  $X$  is called a *quasimartingale*. We remark here that the quasimartingale in [12] is defined on  $[0, \infty]$ . However, since  $[0, T]$  and  $[0, \infty]$  are homeomorphic, their results apply as well on  $[0, T]$ . We only point out that the conditional variation  $V_T(X)$  defined here is the finite horizon version of that in [12]. We have the following result.

**Theorem 2.3** (Meyer-Zheng [12]) *Let  $\{P_n\}_{n \geq 1} \subset \mathcal{M}(\mathbb{D})$ , such that under each  $P_n$  the coordinate process  $X_t(\omega) = \omega(t)$ ,  $t \in [0, T]$ ,  $\omega \in \mathbb{D}$ , is a quasimartingale. Assume that  $V_n(X)$ ,  $n \geq 1$ , the conditional variation of  $X$  under  $P_n$ 's, are uniformly bounded in  $n$ . Then there exists a subsequence  $\{P_{n_k}\}$  which converges weakly on  $\mathbb{D}$  to a law  $P^* \in \mathcal{M}(\mathbb{D})$ , and  $X$  is a quasimartingale under  $P^*$ .*

### 3 Discrete Case Revisited

In this section we look at the special case where the functional  $\Phi$  takes a special “discrete” form. To be more precise, we let  $\pi : 0 = t_0 < t_1 < \dots < t_n = T$  be any partition of  $[0, T]$ ; and assume that

$$\Phi(\mathbf{x}) = g(\mathbf{x}_{t_0}, \dots, \mathbf{x}_{t_n}),$$

where  $g \in C_b^1(\mathbb{R}^{n+1})$ . We note that such a case was also studied by Jacod-Méléard-Protter [8], but we shall give a slightly different formula that is more useful for our future discussion. We assume that all the (first order) partial derivatives of  $g$ , denoted by  $\partial_i g$  ( $= \partial_{x_i} g$ ),  $i = 1, 2, \dots$ , are bounded by a common constant  $K > 0$ . Recall the function  $\varphi$  defined by (2.16). Clearly, for  $t \in (t_{i-1}, t_i]$  and  $\mathbf{x} \in \mathbb{D}$ ,

$$\varphi(\mathbf{x}; t, y) = E\left\{g(\mathbf{x}(t_0), \dots, \mathbf{x}(t_{i-1}), X_{t_i}^{t,y}, \dots, X_{t_n}^{t,y})\right\}. \quad (3.1)$$

Since  $X$  is Markovian, it can be shown that the martingale defined by (2.14) can be written as

$$M_t = \varphi(X; t, X_t) = \lim_{\substack{t_{i-1} < u < t \\ u \downarrow t_{i-1}}} \varphi(X 1_{[0,u)}; t, X_t),$$



where the second equality holds true for  $t \in (t_{i-1}, t_i]$ . Note that by (2.15) we have

$$dM_t = \xi_t dX_t + dN_t. \quad (3.2)$$

We shall follow the idea of [8] to identify  $\xi$ . To this end, we define a function  $\phi : [0, T] \times \mathbb{R}^i$  by

$$\phi(t, x_1, \dots, x_{i-1}, y) = E\left\{g(x_1, \dots, x_{i-1}, X_{t_i}^{t,y}, \cdot, X_{t_n}^{t,y})\right\}.$$

Clearly, we have  $\varphi(X; t, X_t) = \phi(t, X_{t_0}, \dots, X_{t_{i-1}}, X_t)$  and (suppressing all variables)  $\partial_t \varphi = \partial_t \phi$ ,  $\partial_y \varphi = \partial_y \phi$ . We shall first assume that  $\phi \in C^{1,2}((t_{i-1}, t_i] \times \mathbb{R}^i)$ . Also, when the context is clear, we shall simply write  $\phi(t, y) = \phi(t, X_{t_0}, \dots, X_{t_{i-1}}, y)$  for notational convenience.

Now, for  $j = 1, \dots, i-1$ , we define  $X_t^j \equiv X_{t_j}$ , for all  $t \geq t_{i-1} \geq t_j$ . Then, applying Itô's formula over  $[t_{i-1}, t_i]$ , and noting that  $dX_s^j \equiv 0$ ,  $j = 1, \dots, i-1$  and  $d\langle X^j, Y \rangle \equiv 0$ , for  $Y = X, X^1, \dots, X^{i-1}$ , and  $t \geq t_{i-1}$ , we get:

$$\begin{aligned} dM_t &= d\phi(t, X_t^1, \dots, X_t^i, X_t) \\ &= \partial_t \phi(t, X_{t-}) dt + \partial_y \phi(t, X_{t-}) dX_t + \frac{1}{2} \partial_{yy} \phi(t, X_{t-}) \sigma^2(t, X_{t-}) dt \\ &\quad + \int_{\mathbb{R}} \left[ \phi(t, X_{t-} + b(t, X_{t-})z) - \phi(t, X_{t-}) - \partial_y \phi(t, X_{t-}) b(t, X_{t-})z \right] \mu(dt, dz). \end{aligned} \quad (3.3)$$

Since  $M$  is a martingale, the finite variation terms on the right side of (3.3) must equal 0; and an argument analogous to that in [8] then shows that

$$\begin{aligned} dM_t &= \partial_y \phi(t, X_{t-}) dX_t + \int_{\mathbb{R}} \left[ \phi(t, X_{t-} + b(t, X_{t-})z) - \phi(t, X_{t-}) \right. \\ &\quad \left. - \partial_y \phi(t, X_{t-}) b(t, X_{t-})z \right] \tilde{\mu}(dtdz) \\ &= (\partial_y \phi \sigma)(t, X_{t-}) dW_t + \int_{\mathbb{R}} \left[ \phi(t, X_{t-} + b(t, X_{t-})z) - \phi(t, X_{t-}) \right] \tilde{\mu}(dtdz). \end{aligned} \quad (3.4)$$

Since  $X$  and  $N$  are orthogonal, by combining (2.10), (3.2) and (3.4) we have

$$\begin{aligned} d\langle M, X \rangle_t &= \xi_t (\sigma^2 + b^2)(t, X_{t-}) dt \\ &= \left[ (\partial_y \phi \sigma^2)(t, X_{t-}) + b(t, X_{t-}) \int_{\mathbb{R}} z \Lambda_t(z) F(dz) \right] dt, \end{aligned}$$

where

$$\Lambda_t(z) \triangleq \phi(X, t, X_{t-} + b(t, X_{t-})z) - \phi(X, t, X_{t-}). \quad (3.5)$$

Consequently, one has, for  $t \in (t_{i-1}, t_i)$ ,

$$\xi_t = \frac{(\partial_y \phi \sigma^2)(t, X_{t-}) + b(t, X_{t-}) \int_{\mathbb{R}} z \Lambda_t(z) F(dz)}{(\sigma^2 + b^2)(t, X_{t-})}. \quad (3.6)$$

Now we are ready to prove the following theorem.

**Theorem 3.1** *Assume (A1), and assume that the function  $g$  is continuously differentiable with bounded derivatives. Then it holds that*

$$\xi_t = \frac{\sigma^2(t, X_{t-}) [\nabla X_{t-}]^{-1} \xi_{t-}^1 + \xi_t^2}{(\sigma^2 + b^2)(t, X_{t-})} \quad (3.7)$$

is càglàd, where

$$\begin{cases} \xi_t^1 \triangleq E \left\{ \sum_{t_j > t} \partial_j g(X_{t_0}, \dots, X_{t_n}) \nabla X_{t_j} \middle| \mathcal{F}_t \right\}; \\ \xi_t^2 \triangleq b(t, X_{t-}) \int_{\mathbb{R}} z \Lambda_t(z) F(dz), \end{cases} \quad (3.8)$$

and  $\Lambda$  is defined by (3.5).

*Proof.* Recall that  $\nu(dt, dx) = dtF(dx)$ . Assume first that  $F$  has compact support, that  $\sigma, b$  are infinitely differentiable with bounded derivatives of all orders, and that  $g$  is bounded, and twice continuously differentiable with bounded derivatives. Following the argument of Lemma 5.1 in [8] we know that in each subinterval  $(t_{i-1}, t_i)$ ,  $\phi \in C^{1,2}((t_{i-1}, t_i) \times \mathbb{R})$ . Then following the same argument as before we see that (3.6) must hold for  $t \in (t_{i-1}, t_i)$ . Now using (3.1) we derive that

$$\partial_y \phi(t, X_t) = [\nabla X_t]^{-1} E \left\{ \sum_{j \geq i} g_j(X_{t_0}, \dots, X_{t_n}) \nabla X_{t_j} \middle| \mathcal{F}_t \right\} = [\nabla X_t]^{-1} \xi_t^1.$$

Noting that  $\partial_y \phi$  is continuous, and taking left limits on both sides above we obtain (3.7) on  $(t_{i-1}, t_i)$ .

To show that  $\xi$  is càglàd we first observe from (3.8) that  $\xi^1$  is obviously càdlàg, hence the mapping  $t \mapsto \sigma^2(t, X_{t-}) [\nabla X_{t-}]^{-1} \xi_{t-}^1$  is càglàd. Furthermore, from (3.5) we see that the process  $\Lambda$  is càglàd. Also, applying Lemma 2.2 one shows that  $|\phi_x| \leq C$  for some constant  $C > 0$ . Thus,

$$|\Lambda_t(z)| \leq C |b(t, X_{t-})| |z| \leq C \left(1 + \sup_{0 \leq s \leq T} |X_s|\right) |z|.$$

Now a simple application of the Dominated Convergence Theorem shows that  $\xi^2$ , defined by (3.8), is also càglàd . Therefore so is  $\xi$ .

It remains to remove the extra assumptions made on  $F$ ,  $g$ ,  $\sigma$  and  $b$ . Here we can follow closely the approximation techniques of [8]. We leave it to the interested readers. The proof is now complete.  $\blacksquare$

For future applications, we now extend Theorem 3.1 slightly:

**Theorem 3.2** *If  $g$  is continuously differentiable, and further it is of linear growth and all its partial derivatives composed with  $X$ ,  $\partial_i g(X_{t_0}, \dots, X_{t_n})$ , are square-integrable, then (3.7) holds and  $\xi$  is càglàd .*

*Proof.* Let  $\{\phi^m\} \subset C_0^1(\mathbb{R}^{n+1})$  be a sequence of *truncation* functions satisfying  $0 \leq \phi^m \leq 1$ ;  $|\partial_i \phi^m| \leq 1$ ; and

$$\phi^m(x_0, \dots, x_n) = \begin{cases} 1, & |(x_0, \dots, x_n)| \leq m; \\ 0, & |(x_0, \dots, x_n)| \geq m + 1. \end{cases}$$

Define  $\Phi_m : \mathbb{D} \mapsto \mathbb{R}$  by

$$\Phi_m(\mathbf{x}) = (g\phi^m)(\mathbf{x}(t_0), \dots, \mathbf{x}(t_n)). \quad (3.9)$$

Then clearly  $\Phi_m$  has compact support with all derivatives bounded. Applying Theorem 3.1 we have

$$\Phi_m(X) = \alpha_m + \int_0^T \xi_t^m dX_t + N_T^m,$$

where  $\xi^m$  is càglàd and satisfies, for  $t \in (t_{i-1}, t_i]$ ,

$$\xi_t^m(\sigma^2 + b^2)(t, X_{t-}) = \sigma^2(t, X_{t-})(\nabla X_{t-})^{-1} \xi_{t-}^{m,1} + \xi_t^{m,2}, \quad \text{a.s.}$$

Here  $\xi^{m,1}$ ,  $\xi^{m,2}$  are defined in the same way as those in (3.8), as well as (3.5), with  $g$  being replaced by  $g\phi^m$ . Since  $g$  is of linear growth, and  $\partial_j g(X_{t_0}, \dots, X_{t_n})$ 's are all square-integrable, letting  $m \rightarrow \infty$  on both sides above and applying Lebesgue's dominated convergence theorem, one concludes that (3.7) holds and  $\xi$  is càglàd .  $\blacksquare$

## 4 $L^\infty$ -Lipschitz Case

In this section we present our first path regularity result, under a rather general condition on the functional  $\Phi$ , which we call the  $L^\infty$ -Lipschitz condition:

$$|\Phi(\mathbf{x}_1) - \Phi(\mathbf{x}_2)| \leq L \|\mathbf{x}_1 - \mathbf{x}_2\|_\infty; \quad \forall \mathbf{x}_1, \mathbf{x}_2 \in \mathbb{D}. \quad (4.1)$$

where  $\|\mathbf{x}\|_\infty \triangleq \sup_{t \in [0, T]} |\mathbf{x}(t)|$ . Two important cases under such an assumption will be studied separately in the next section.

We first give a lemma that shows the implication of (4.1) on the function  $\varphi$  defined by (2.16).

**Lemma 4.1** *Suppose that  $\Phi$  satisfies (4.1), and let  $\varphi$  be defined by (2.16). Then there exists a constant  $C > 0$ , depending only on the time duration  $T$  and the constants  $K$  in (A1) and  $L$  in (4.1), such that for any  $\mathbf{x}, \mathbf{x}_1, \mathbf{x}_2 \in \mathbb{D}$  and  $y, y_1, y_2 \in \mathbb{R}$ ,*

$$\begin{cases} |\varphi(\mathbf{x}_1, t, y) - \varphi(\mathbf{x}_2, t, y)| \leq C \sup_{0 \leq s < t} |\mathbf{x}_1(s) - \mathbf{x}_2(s)|; \\ |\varphi(\mathbf{x}, t, y_1) - \varphi(\mathbf{x}, t, y_2)| \leq C |y_1 - y_2|. \end{cases} \quad (4.2)$$

*Proof.* First note that, by (2.16) and (4.1),

$$\begin{aligned} |\varphi(\mathbf{x}_1, t, y) - \varphi(\mathbf{x}_2, t, y)| &\leq E \left\{ |\Phi(\mathbf{x}_1 1_{[0, t]} + X^{t, y} 1_{[t, T]}) - \Phi(\mathbf{x}_2 1_{[0, t]} + X^{t, y} 1_{[t, T]})| \right\} \\ &\leq L \|\mathbf{x}_1 1_{[0, t]} + X^{t, y} 1_{[t, T]} - \mathbf{x}_2 1_{[0, t]} + X^{t, y} 1_{[t, T]}\|_\infty = L \sup_{0 \leq s < t} |\mathbf{x}_1(s) - \mathbf{x}_2(s)|. \end{aligned}$$

Similarly, for any  $\mathbf{x} \in \mathbb{D}$  and  $y_1, y_2 \in \mathbb{R}$ ,

$$\begin{aligned} |\varphi(\mathbf{x}, t, y_1) - \varphi(\mathbf{x}, t, y_2)| &\leq E \left\{ |\Phi(\mathbf{x} 1_{[0, t]} + X^{t, y_1} 1_{[t, T]}) - \Phi(\mathbf{x} 1_{[0, t]} + X^{t, y_2} 1_{[t, T]})| \right\} \\ &\leq L E \left\{ \sup_{t \leq s \leq T} |X_s^{t, y_1} - X_s^{t, y_2}| \right\} = L |y_1 - y_2| E \left\{ \sup_{t \leq s \leq T} |\Delta_{y_1 - y_2} X_s^{t, y_2}| \right\} \leq C |y_1 - y_2|, \end{aligned}$$

thanks to (2.13) and Lemma 2.2. This completes the proof.  $\blacksquare$

In light of the idea in [11] for the Brownian case, we shall approximate the functional  $\Phi$  by a sequence of discrete functionals, which we now describe. For any partition  $\pi : 0 = t_0 < t_1 < \dots < t_n = T$ , we define a mapping  $\pi : \mathbb{D} \mapsto \mathbb{D}$  by  $\mathbf{x} \mapsto \pi(\mathbf{x}) \triangleq \mathbf{x}_\pi$ , where

$$\mathbf{x}_\pi(t) \triangleq \sum_{i=1}^n \mathbf{x}(t_i) 1_{[t_{i-1}, t_i)}(t) + \mathbf{x}(T) 1_{\{T\}}(t). \quad (4.3)$$

Denote  $|\pi| = \max_i |t_{i+1} - t_i|$  to be the mesh size of the partition  $\pi$ . Then, using the right continuity of  $\mathbf{x}$  it is easily seen that

$$\lim_{|\pi| \rightarrow 0} |\mathbf{x}_\pi(t) - \mathbf{x}(t)| = 0. \quad (4.4)$$

The following lemma is a slight variation of Lemma 4.1 in [11]. We shall state it without proof.

**Lemma 4.2** *Suppose that  $\Phi$  satisfies (4.1). Let  $\Pi = \{\pi\}$  be a family of partitions of  $[0, T]$ . Then there exists a family of discrete functionals  $\{g^\pi : \pi \in \Pi\}$  such that*

(i) *for each  $\pi \in \Pi$ ,  $g^\pi \in C_b^1(\mathbb{R}^{n+1})$ , and satisfies*

$$\sum_{i=0}^n |g_i^\pi(x_0, \dots, x_n) y_i| \leq C \max_{0 \leq i \leq n} |y_i|. \quad (4.5)$$

with constant  $L$  being the same as that in (4.1).

(ii) *for any  $\mathbf{x} \in \mathbb{D}$ , it holds that*

$$|g^\pi(\mathbf{x}(t_0), \dots, \mathbf{x}(t_n)) - \Phi(\mathbf{x}_\pi)| \leq |\pi|. \quad (4.6)$$

Our main theorem of this section is the following.

**Theorem 4.3** *Assume (A1), and that  $\Phi$  satisfies (4.1). Assume further that the function  $\varphi$  defined by (2.16) is càglàd with respect to  $t$ , for each fixed  $(\mathbf{x}, y) \in \mathbb{D} \times \mathbb{R}$ , and that*

$$\lim_{|\pi| \rightarrow 0} |\Phi(\mathbf{x}_\pi) - \Phi(\mathbf{x})| = 0; \quad \forall \mathbf{x} \in \mathbb{D}, \quad (4.7)$$

where  $\mathbf{x}_\pi$  is defined by (4.3). Then  $\xi$  in (2.15) admits a càglàd version.

*Proof.* Let the partition  $\pi : 0 \leq t_0 < t_1 < \dots < t_n = T$  be given. For  $(\mathbf{x}, t, y) \in \mathbb{D} \times [0, T] \times \mathbb{R}$ , define

$$\Phi_\pi(\mathbf{x}) \triangleq g^\pi(\mathbf{x}(t_0), \dots, \mathbf{x}(t_n)); \quad \varphi^\pi(\mathbf{x}, t, y) \triangleq E\{\Phi_\pi(\mathbf{x}1_{[0,t]} + X^{t,y}1_{[t,T]})\}. \quad (4.8)$$

where  $g^\pi$  is the approximation of  $g$  given by Lemma 4.2. By (4.6) and (4.1),

$$|\Phi_\pi(\mathbf{x})| \leq |\Phi(\mathbf{x}_\pi)| + |\pi| \leq C \left(1 + \sup_{0 \leq t \leq T} |\mathbf{x}(t)|\right). \quad (4.9)$$

Applying Lemma 2.1 and the Dominated Convergence Theorem we derive from (4.6) and (4.7) that

$$\lim_{|\pi| \rightarrow 0} E|\Phi_\pi(X) - \Phi(X)|^2 = 0. \quad (4.10)$$

Now Theorem 3.1 tells us that in the following representation

$$\Phi_\pi(X) = \alpha_\pi + \int_0^T \xi_t^\pi dX_t + N_T^\pi. \quad (4.11)$$

the process  $\xi^\pi$  is càglàd ; and it has an explicit form:

$$\xi_t^\pi(\sigma^2 + b^2)(t, X_{t-}) = \sigma^2(t, X_{t-})(\nabla X_{t-})^{-1} \xi_{t-}^{\pi,1} + \xi_t^{\pi,2}, \quad (4.12)$$

where

$$\begin{cases} \xi_t^{\pi,1} \triangleq E\left\{ \sum_{t_j > t}^n g_j^\pi(X_{t_0}, \dots, X_{t_n}) \nabla X_{t_j} \middle| \mathcal{F}_t \right\}; & \forall t \in (t_{i-1}, t_i]; \\ \xi_t^{\pi,2} \triangleq b(t, X_{t-}) \int_{\mathbb{R}} (\varphi^\pi(X, t, X_{t-} + b(t, X_{t-})z) - \varphi^\pi(X, t, X_{t-})) z F(dz). \end{cases} \quad (4.13)$$

Further, by virtue of (4.10), we see that as  $|\pi| \rightarrow 0$  we must have  $\alpha_\pi \rightarrow \alpha$  and

$$E\left\{ \int_0^T |\xi_t^\pi - \xi_t|^2 (\sigma^2 + b^2)(t, X_{t-}) dt \right\} \rightarrow 0. \quad (4.14)$$

To show that  $\xi$  has a càglàd version, it suffices to show that there exist càglàd processes  $\xi^1$  and  $\xi^2$  such that  $\xi$  has the explicit form:

$$\xi_t(\sigma^2 + b^2)(t, X_{t-}) = \sigma^2(t, X_{t-})(\nabla X_{t-})^{-1} \xi_t^1 + \xi_t^2, \quad \text{a.s.} \quad (4.15)$$

To this end, we note that from (4.8) and (4.5) one has

$$\begin{aligned} & |\varphi^\pi(X, t, X_{t-} + b(t, X_{t-})z) - \varphi^\pi(X, t, X_{t-})| \\ & \leq CE \left\{ \sup_{t \leq s \leq T} |X_s^{t, X_{t-} + b(t, X_{t-})z} - X_s^{t, X_{t-}}| \middle| \mathcal{F}_t \right\} \\ & = CE \left\{ |b(t, X_{t-})z| \sup_{t \leq s \leq T} |\Delta_{b(t, X_{t-})z} X_s| \middle| \mathcal{F}_t \right\} \leq C|b(t, X_{t-})||z|. \end{aligned}$$

where the last inequality is due to Lemma 2.1. Thus if we define

$$\xi_t^2 \triangleq b(t, X_{t-}) \int_{\mathbb{R}} (\varphi(X, t, X_{t-} + b(t, X_{t-})z) - \varphi(X, t, X_{t-})) z F(dz), \quad (4.16)$$

and note that  $\int_{\mathbb{R}} F(dz)z^2 = 1$ , then by applying the Dominated Convergence Theorem, we derive from (4.7) that

$$\lim_{|\pi| \rightarrow 0} E\left\{ \int_0^T |\xi_t^{\pi,2} - \xi_t^2| dt \right\} = 0. \quad (4.17)$$

Therefore, possibly along a subsequence,  $\xi^{\pi,2} \rightarrow \xi^2$  as  $|\pi| \rightarrow 0$ , in measure, a.s. Now since  $\varphi$  is càglàd in  $t$ , we see that so is  $\xi^2$ .

It remains to identify  $\xi^1$  and show it has a càglàd version. We shall make use of Meyer-Zheng's tightness criterion again. First, combining (4.12), (4.14), and (4.17) we see that, possibly along a subsequence, one has

$$\xi_{t-}^{\pi,1} \rightarrow \xi_t^1 \triangleq \frac{\xi_t(\sigma^2 + b^2)(t, X_{t-}) - \xi_t^2}{\sigma^2(t, X_{t-})[\nabla X_{t-}]^{-1}}, \quad \text{as } |\pi| \rightarrow 0, \quad (4.18)$$

and the convergence is in measure, a.s.

On the other hand, let  $\pi' : 0 = s_0 < \dots < s_m = T$  be any partition of  $[0, T]$  that is finer than  $\pi$ . We assume that  $t_i = s_{l_i}$ ,  $i = 1, 2, \dots, n$ . Then,

$$\begin{aligned} \sum_{j=1}^m E\{ |E\{\xi_{s_j}^{\pi,1} - \xi_{s_{j-1}}^{\pi,1} | \mathcal{F}_{s_{j-1}}\}| \} &= \sum_{i=1}^n \sum_{j=l_{i-1}+1}^{l_i} E\{ |E\{\xi_{s_j}^{\pi,1} - \xi_{s_{j-1}}^{\pi,1} | \mathcal{F}_{s_{j-1}}\}| \} \\ &= \sum_{i=1}^n E\{ |E\{g_i^\pi \nabla X_{t_i} | \mathcal{F}_{t_{i-1}}\}| \} \leq E\left\{ \sum_{i=1}^n |g_i^\pi \nabla X_{t_i}| \right\} \\ &\leq CE \left\{ \sup_{0 \leq t \leq T} |\nabla X_t| \right\} \leq C. \end{aligned} \quad (4.19)$$

thanks to (4.5). Therefore, the processes  $\xi^{\pi,1}$ 's all have bounded conditional variation, and hence they are all quasimartingales as defined in §2. Furthermore, we note that the uniform bound of these conditional variations are indeed independent of the choice of  $\pi$ . Consequently, applying the Meyer-Zheng theorem (Lemma 2.3), there exists a càdlàg process  $\tilde{\xi}^1$  such that the càdlàg version of  $\xi^{\pi,1}$  converges to  $\tilde{\xi}^1$  weakly under the Meyer-Zheng topology, as  $|\pi| \rightarrow 0$ . Note that the Meyer-Zheng (pseudo-path) topology is equivalent to convergence in measure (see, e.g., [12] or [11]), so if by a slight abuse of notation we denote the càglàd version of  $\tilde{\xi}^1$  by itself, then the uniqueness of the limit shows that  $\tilde{\xi}_t^1 = \xi_t^1$ , a.s.,  $\forall t$ . In other words,  $\xi^1$ , whence  $\xi$ , has a càglàd version. This completes the proof.  $\blacksquare$

## 5 Some Sufficient Conditions

One of the main differences between the Lévy case and the Brownian case is that the  $L^\infty$ -Lipschitz condition alone *does not* guarantee the path regularity of the process  $\xi$ . In fact, in Theorem 4.3 we required an extra assumption on the mapping  $t \mapsto$

$\varphi(\mathbf{x}, t, y)$ , which is not easy to verify in general. In this section we consider two cases where the functional  $\Phi$  satisfies some stronger ‘‘Lipschitz’’ conditions so that the extra assumption can either be removed or be replaced by a more easily verifiable one. These two cases are mainly motivated by connections to finance. The first one corresponds to the ‘‘Asian Option’’, while the second one corresponds to the ‘‘Lookback Option’’. At the end of the section we shall give an example to show that there exists a non-regular  $\xi$  even though  $\Phi$  satisfies the  $L^\infty$ -Lipschitz condition.

**Theorem 5.1** *Suppose that  $\Phi$  satisfies the  $L^1$ -Lipschitz (or Integral Lipschitz) condition:*

$$|\Phi(\mathbf{x}_1) - \Phi(\mathbf{x}_2)| \leq L \int_0^T |\mathbf{x}_1(t) - \mathbf{x}_2(t)| dt. \quad (5.1)$$

then the process  $\xi$  in the representation (2.15) admits a càglàd version.

*Proof.* It is clear that the  $L^1$ -Lipschitz condition (5.1) implies the  $L^\infty$ -Lipschitz condition (4.1). Thus, by Theorem 4.3, it suffices to prove that  $\varphi$  is càglàd and that (4.7) holds.

Note that for any  $\mathbf{x} \in \mathbb{ID}$ , it holds that  $\|\mathbf{x}_\pi\|_\infty \leq \|\mathbf{x}\|_\infty < \infty$ . Then, by (5.1) and the Dominated Convergence Theorem, we have

$$|\Phi(\mathbf{x}_\pi) - \Phi(\mathbf{x})| \leq L \int_0^T |\mathbf{x}_\pi(t) - \mathbf{x}(t)| dt \rightarrow 0; \quad \text{as } |\pi| \rightarrow 0, \quad (5.2)$$

thanks to (4.4). Furthermore, by Lemma 2.1 one has, for  $0 \leq t_1 < t_2 \leq T$ ,

$$\begin{aligned} & |\varphi(\mathbf{x}, t_1, y) - \varphi(\mathbf{x}, t_2, y)| \\ & \leq E \left\{ \left| \Phi(\mathbf{x}1_{[0,t_1]} + X^{t_1,y}1_{[t_1,T]}) - \Phi(\mathbf{x}1_{[0,t_2]} + X^{t_2,y}1_{[t_2,T]}) \right| \right\} \\ & \leq CE \left\{ \int_{t_1}^{t_2} |X_s^{t_1,y} - \mathbf{x}(s)| ds + \int_{t_2}^T |X_s^{t_1,y} - X_s^{t_2,y}| ds \right\} \\ & \leq CE \left\{ \sup_{t_1 \leq s \leq t_2} (|X_s^{t_1,y}| + |\mathbf{x}(s)|)(t_2 - t_1) + |X_{t_2}^{t_1,y} - y| \sup_{t_2 \leq s \leq T} |\Delta_{X_{t_2}^{t_1,y} - y} X_s^{t_2,y}| \right\} \\ & \leq C(1 + |y| + \|\mathbf{x}\|_\infty)(t_2 - t_1)^{\frac{1}{2}}. \end{aligned} \quad (5.3)$$

Therefore,  $\varphi$  is continuous, which, together with (5.2), enables us to apply Theorem 4.3 to conclude that  $\xi$  has a càglàd version.  $\blacksquare$

The second case, motivated by the Lookback option, is a little more involved.



**Theorem 5.2** *If  $\Phi(\mathbf{x}) = g(\sup_{0 \leq t \leq T} h(t, \mathbf{x}(t)))$ , where  $g$  and  $h(t, \cdot)$  are uniformly Lipschitz with a common constant  $K$ , and  $h(\cdot, x)$  is continuous for all  $x$ . Then,  $\xi$  in (2.15) admits a càglàd version.*

*Proof.* That  $\Phi$  satisfies (4.1) is obvious. So again we need to show only that  $\varphi$  is càglàd in  $t$ , and (4.7) holds.

First fix  $\mathbf{x} \in \mathbb{D}$  and let  $\pi : 0 = t_0 < \dots < t_n = T$  be any partition. Since  $g$  is Lipschitz, we have

$$\begin{aligned} |\Phi(\mathbf{x}_\pi) - \Phi(\mathbf{x})| &\leq C \left| \sup_{0 \leq t \leq T} h(t, \mathbf{x}_\pi(t)) - \sup_{0 \leq t \leq T} h(t, \mathbf{x}(t)) \right| \\ &\leq C \left( \sup_{0 \leq t \leq T} h(t, \mathbf{x}(t)) - \max_i h(t_i, \mathbf{x}(t_i)) \right) \\ &\quad + C \max_i \sup_{t_{i-1} \leq t < t_i} |h(t, \mathbf{x}(t_i)) - h(t_i, \mathbf{x}(t_i))|. \end{aligned} \quad (5.4)$$

For any  $\varepsilon > 0$ , choose  $t^\varepsilon \in [0, T]$  such that

$$\sup_{0 \leq t \leq T} h(t, \mathbf{x}(t)) \leq h(t^\varepsilon, \mathbf{x}(t^\varepsilon)) + \varepsilon.$$

Now for each  $i$  we have

$$|h(t^\varepsilon, \mathbf{x}(t^\varepsilon)) - h(t_i, \mathbf{x}(t_i))| \leq |h(t_i, \mathbf{x}(t^\varepsilon)) - h(t^\varepsilon, \mathbf{x}(t^\varepsilon))| + C|\mathbf{x}(t_i) - \mathbf{x}(t^\varepsilon)|,$$

and, thanks to the continuity of  $h$ ,

$$h(t^\varepsilon, \mathbf{x}(t^\varepsilon)) \leq \liminf_{|\pi| \rightarrow 0} \max_i h(t_i, \mathbf{x}(t_i)).$$

Since  $\varepsilon$  is arbitrary, we deduce that

$$\sup_{0 \leq t \leq T} h(t, \mathbf{x}(t)) \leq \liminf_{|\pi| \rightarrow 0} \max_i h(t_i, \mathbf{x}(t_i)). \quad (5.5)$$

Furthermore, since  $\mathbf{x} \in \mathbb{D}$ , it is bounded on  $[0, T]$ . Let  $|\mathbf{x}(t)| \leq K$  for some constant  $K > 0$ . Also, since  $h$  is continuous, it is uniformly continuous on  $[0, T] \times [-K, K]$ .

Thus we have

$$\lim_{|\pi| \rightarrow 0} \max_i \sup_{t_{i-1} \leq t < t_i} |h(t, \mathbf{x}(t_i)) - h(t_i, \mathbf{x}(t_i))| = 0. \quad (5.6)$$

Combining (5.5) and (5.6) we derive from (5.4) that

$$\lim_{|\pi| \rightarrow 0} |\Phi(\mathbf{x}_\pi) - \Phi(\mathbf{x})| = 0. \quad (5.7)$$

It remains to show  $\varphi$  is càglàd in  $t$ . To do this we observe that, for fixed  $\mathbf{x}$ ,  $y$ , and any  $0 \leq t_1 < t_2 \leq T$ ,

$$\begin{aligned}
& |\varphi(\mathbf{x}, t_1, y) - \varphi(\mathbf{x}, t_2, y)| \\
&= \left| E \left\{ \Phi(\mathbf{x}1_{[0,t_1]} + X^{t_1,y}1_{[t_1,T]}) - \Phi(\mathbf{x}1_{[0,t_2]} + X^{t_2,y}1_{[t_2,T]}) \right\} \right| \\
&\leq CE \left\{ \left| \sup_{0 \leq s < t_1} h(s, \mathbf{x}(s)) \vee \sup_{t_1 \leq s \leq T} h(s, X_s^{t_1,y}) - \sup_{0 \leq s < t_2} h(s, \mathbf{x}(s)) \vee \sup_{t_2 \leq s \leq T} h(s, X_s^{t_2,y}) \right| \right\} \\
&\leq CE \left\{ \left| \sup_{0 \leq s < t_1} h(s, \mathbf{x}(s)) - \sup_{0 \leq s < t_2} h(s, \mathbf{x}(s)) \right| + \left| \sup_{t_1 \leq s \leq T} h(s, X_s^{t_1,y}) - \sup_{t_2 \leq s \leq T} h(s, X_s^{t_2,y}) \right| \right\} \\
&\leq CE \left\{ \sup_{t_1 \leq s < t_2} |h(s, \mathbf{x}(s)) - h(t_1, \mathbf{x}(t_1-))| \right. \\
&\quad \left. + \sup_{t_1 \leq s < t_2} |h(s, X_s^{t_1,y}) - h(t_2, y)| + \sup_{t_2 \leq s \leq T} |h(s, X_s^{t_1,y}) - h(s, X_s^{t_2,y})| \right\} \\
&\leq CE \left\{ \sup_{t_1 \leq s < t_2} |h(s, \mathbf{x}(t_1-)) - h(t_1, \mathbf{x}(t_1-))| + \sup_{t_1 \leq s < t_2} |\mathbf{x}(s) - \mathbf{x}(t_1-)| \right. \\
&\quad \left. + \sup_{t_1 \leq s < t_2} |h(s, y) - h(t_2, y)| + \sup_{t_1 \leq s < t_2} |X_s^{t_1,y} - y| + |X_{t_2}^{t_1,y} - y| \sup_{t_2 \leq s \leq T} |\Delta_{X_{t_2}^{t_1,y} - y} X_s^{t_2,y}| \right\} \\
&\leq C \left[ \sup_{t_1 \leq s < t_2} |h(s, \mathbf{x}(t_1-)) - h(t_1, \mathbf{x}(t_1-))| + \sup_{t_1 \leq s < t_2} |\mathbf{x}(s) - \mathbf{x}(t_1-)| \right. \\
&\quad \left. + \sup_{t_1 \leq s < t_2} |h(s, y) - h(t_2, y)| + (t_2 - t_1)^{\frac{1}{2}} \right]. \tag{5.8}
\end{aligned}$$

thanks to Lemma 2.1 and Lemma 2.2. Now fix  $t_0 \in (0, T)$ . For  $\forall \varepsilon > 0$ , since  $\mathbf{x}$  is càdlàg, there exists  $\delta > 0$  such that,

$$\begin{cases} |\mathbf{x}(t) - \mathbf{x}(t_0-)| \leq \varepsilon; & \forall t \in (t_0 - \delta, t_0); \\ |\mathbf{x}(t) - \mathbf{x}(t_0)| \leq \varepsilon; & \forall t \in (t_0, t_0 + \delta). \end{cases} \tag{5.9}$$

Thus

$$\begin{cases} \sup_{t \leq s < t_0} |\mathbf{x}(s) - \mathbf{x}(t-)| \leq 2\varepsilon; & \forall t \in (t_0 - \delta, t_0); \\ \sup_{t_1 \leq s < t_2} |\mathbf{x}(s) - \mathbf{x}(t_1-)| \leq 2\varepsilon; & \forall t_0 < t_1 < t_2 < t_0 + \delta. \end{cases} \tag{5.10}$$

which, combined with (5.8) and the fact that  $h$  is locally uniformly continuous, clearly implies that  $\varphi$  is càglàd.  $\blacksquare$

To conclude this section we give an example which shows that in general the  $L^\infty$ -Lipschitz condition (4.1) alone does not guarantee that  $\xi$  is càglàd.

**Example 5.3** Let  $X_t = N_t - t$ , where  $N$  is the standard Poisson process. (That is,  $\sigma = 0, b = 1$  and  $F(dz) = \delta_{\{1\}}(dz)$ .) Let  $A \in \mathcal{B}([0, T])$  such that  $A$  is dense in  $[0, T]$

and  $0 < |A| < T$ , where  $|A|$  denotes the Lebesgue measure of  $A$ . Define  $\Phi : \mathbb{D} \mapsto \mathbb{R}$  by  $\Phi(\mathbf{x}) = \sup_{t \in A} |\Delta \mathbf{x}(t)|$ . Now consider the equation:

$$M_t = 1 - \int_0^t M_{s-} 1_A(s) dX_s.$$

By the *Doléans-Dade Exponential Formula* (cf. Protter [13]) we have

$$\begin{aligned} M_t &= \exp \left\{ - \int_0^t 1_A(s) dX_s \right\} \prod_{s \leq t} [(1 - 1_A(s) \Delta X_s) e^{1_A(s) \Delta X_s}] & (5.11) \\ &= \exp \left\{ - \sum_{s \leq t} 1_A(s) \Delta N_s + |A \cap [0, t]| \right\} \left[ \prod_{s \leq t} (1 - 1_A(s) \Delta N_s) \right] \exp \left\{ \sum_{s \leq t} 1_A(s) \Delta N_s \right\} \\ &= e^{|A \cap [0, t]|} \prod_{s \leq t} (1 - 1_A(s) \Delta N_s). \end{aligned}$$

Since  $\Phi(N)$  only takes values 0 or 1, and if  $\Phi(N) = 0$ , then for all  $s \in A$ ,  $\Delta N_s = 0$ , thus  $1_A(s) \Delta N_s = 0$  for all  $s \in [0, T]$ , hence  $M_T = e^{|A|}$ , thanks to (5.11). On the other hand, if  $\Phi(N) = 1$ , then there exists  $s \in A$  such that  $\Delta N_s = 1$ , again by (5.11), we have  $M_T = 0$ . So

$$M_T = e^{|A|} 1_{\{\Phi(N)=0\}} = e^{|A|} (1 - \Phi(N)).$$

then

$$\Phi(X) = \Phi(N) = 1 - e^{-|A|} M_T = 1 - e^{-|A|} + \int_0^T e^{-|A|} M_{t-} 1_A(t) dX_t. \quad (5.12)$$

Hence

$$\xi_t = e^{-|A|} M_{t-} 1_A(t)$$

which clearly is not càglàd . ■

## 6 Stochastic Integral Case

In this section we consider the case where  $\Phi(X) = g\left(\int_0^T h(t, X_{t-}) dX_t\right)$ . Note that in this case  $\Phi(X)$  no longer depends on  $X$  in a path by path manner, and none of the “Lipschitz conditions” studied in the previous sections is satisfied. Let us first modify the function  $\varphi$ . Observe that in this case

$$\begin{aligned} &\Phi(X 1_{[0,t)} + X^{t,y} 1_{[t,T)}) \\ &= g\left(\int_0^{t-} h(s, X_{s-}) dX_s + h(t, X_{t-})(y - X_{t-}) + \int_t^T h(s, X_{s-}^{t,y}) dX_s^{t,y}\right), \end{aligned}$$

we shall introduce the following two new functions  $\psi : \Omega \times \mathbb{R}^2 \times [0, T] \times \mathbb{R} \mapsto \mathbb{R}$ , and  $\varphi : \mathbb{R}^2 \times [0, T] \times \mathbb{R} \mapsto \mathbb{R}$ .

$$\begin{cases} \psi(a, x, t, y) \triangleq a + h(t, x)(y - x) + \int_t^T h(s, X_{s-}^{t,y}) dX_s^{t,y}; \\ \varphi(a, x, t, y) \triangleq E\{g(\psi(a, x, t, y))\}. \end{cases} \quad (6.1)$$

The following theorem is an extension of Theorem 5.2.

**Theorem 6.1** *Suppose that  $\Phi(X) = g\left(\int_0^T h(t, X_{t-}) dX_t\right)$ , where  $g$  and  $h$  satisfy*

- (i)  $h$  is bounded;
- (ii) for fixed  $x$ ,  $h(\cdot, x)$  is càglàd ;
- (iii) for fixed  $t$ ,  $g$  and  $h(t, \cdot)$  are uniformly Lipschitz continuous with Lipschitz constant  $K$ ,

*Then,  $\xi$  admits a càglàd version.*

*Proof.* We follow the similar line of the proof of Theorem 5.2, but make necessary adjustments. First let us assume that  $g$  and  $h(t, \cdot)$  are continuously differentiable for fixed  $t$ , and that  $h'_x(\cdot, x)$  is càglàd for fixed  $x$ . We define an approximating discrete functional as follows. For  $\pi : 0 = t_0 < \dots < t_n = T$ , define

$$\Phi_\pi(X) = \tilde{g}(X_{t_0}, \dots, X_{t_n}) \triangleq g\left(\sum_{i=1}^n h(t_{i-1}, X_{t_{i-1}})(X_{t_i} - X_{t_{i-1}})\right); \quad (6.2)$$

Note that  $X$  is a martingale, one has

$$\begin{aligned} & E\{|\Phi_\pi(X) - \Phi(X)|^2\} \\ &= E\left\{\left|g\left(\sum_{i=1}^n h(t_{i-1}, X_{t_{i-1}})(X_{t_i} - X_{t_{i-1}})\right) - g\left(\sum_{i=1}^n \int_{t_{i-1}}^{t_i} h(s, X_{s-}) dX_s\right)\right|^2\right\} \\ &\leq CE\left\{\left|\sum_{i=1}^n \left(h(t_{i-1}, X_{t_{i-1}})(X_{t_i} - X_{t_{i-1}}) - \int_{t_{i-1}}^{t_i} h(s, X_{s-}) dX_s\right)\right|^2\right\} \\ &= CE\left\{\sum_{i=1}^n \left|h(t_{i-1}, X_{t_{i-1}})(X_{t_i} - X_{t_{i-1}}) - \int_{t_{i-1}}^{t_i} h(s, X_{s-}) dX_s\right|^2\right\} \quad (6.3) \\ &\leq CE\left\{\sum_{i=1}^n \int_{t_{i-1}}^{t_i} |h(s, X_{s-}) - h(t_{i-1}, X_{t_{i-1}})|^2 (\sigma^2 + b^2)(s, X_{s-}) ds\right\} \end{aligned}$$

A simple application of the Dominated Convergence Theorem then gives that

$$\lim_{|\pi| \rightarrow 0} E\{|\Phi_\pi(X) - \Phi(X)|^2\} = 0. \quad (6.4)$$

Next, note that by the martingale representation theorem we have

$$\begin{cases} \Phi_\pi(X) = \alpha_\pi + \int_0^T \xi_t^\pi dt + N_T^\pi; \\ \Phi(X) = \alpha + \int_0^T \xi_t dt + N_T, \end{cases}$$

and by (6.4) we have

$$\lim_{|\pi| \rightarrow 0} E \left\{ \int_0^T |\xi_t^\pi - \xi_t|^2 (\sigma^2 + b^2)(t, X_{t-}) dt \right\} = 0. \quad (6.5)$$

Furthermore, by Theorem 3.2 we know that  $\xi^\pi$  admits a càglàd version and has an explicit formula. We want to show that  $\xi$  also has a càglàd version and to identify its explicit form. To this end, let us introduce two functions corresponding to those in (6.1): for  $t \in (t_{i-1}, t_i]$ ,

$$\begin{cases} \psi^\pi(a, x, t, y) \triangleq a + h(t_{i-1}, x)(X_{t_i}^{t,y} - x) + \sum_{j>i} h(t_{j-1}, X_{t_{j-1}}^{t,y})(X_{t_j}^{t,y} - X_{t_{j-1}}^{t,y}); \\ \varphi^\pi(a, x, t, y) \triangleq E\{g(\psi^\pi(a, x, t, y))\}. \end{cases} \quad (6.6)$$

Then, it is easily seen that

$$\Phi_\pi(X1_{[0,t)} + X^{t,y}1_{[t,T)}) = g\left(\psi^\pi\left(\sum_{j<i} h(t_{j-1}, X_{t_{j-1}}), X_{t_{j-1}}, t, y\right)\right), \quad (6.7)$$

and by Theorem 3.2  $\xi^\pi$  can be written explicitly as:

$$\xi_t^\pi = \frac{\sigma^2(t, X_{t-})(\nabla X_{t-})^{-1} \xi_t^{\pi,1} + \xi_t^{\pi,2}}{(\sigma^2 + b^2)(t, X_{t-})}, \quad (6.8)$$

where, for  $t \in (t_{i-1}, t_i]$ ,

$$\begin{cases} \xi_t^{\pi,1} \triangleq \partial_y \varphi^\pi\left(\sum_{j<i} h(t_{j-1}, X_{t_{j-1}})(X_{t_j} - X_{t_{j-1}}), X_{t_{i-1}}, t, X_{t-}\right); \\ \xi_t^{\pi,2} \triangleq b(t, X_{t-}) \int_{\mathbb{R}} z \Lambda_t^\pi(z) F(dz). \end{cases} \quad (6.9)$$

and

$$\begin{aligned} \Lambda_t^\pi(z) &\triangleq \varphi^\pi\left(\sum_{j<i} h(t_{j-1}, X_{t_{j-1}})(X_{t_j} - X_{t_{j-1}}), X_{t_{i-1}}, t, X_{t-} + b(t, X_{t-})z\right) \\ &\quad - \varphi^\pi\left(\sum_{j<i} h(t_{j-1}, X_{t_{j-1}})(X_{t_j} - X_{t_{j-1}}), X_{t_{i-1}}, t, X_{t-}\right). \end{aligned} \quad (6.10)$$

To identify the limit of  $\{\xi^\pi\}$ , we first note that (6.1) and (6.6) yield that

$$\begin{cases} \partial_y \psi(a, x, t, y) = h(t, x) + \int_t^T h'_x(s, X_{s-}^{t,y}) \nabla X_{s-}^{t,y} dX_s^{t,y} \\ \quad + \int_t^T h(s, X_{s-}^{t,y}) d\nabla X_s^{t,y}; \\ \partial_y \varphi(a, x, t, y) = E\{g'(\psi(a, x, t, y)) \partial_y \psi(a, x, t, y)\}. \end{cases} \quad (6.11)$$

and

$$\begin{cases} \partial_y \psi^\pi(a, x, t, y) = h(t_{i-1}, x) \nabla X_{t_i}^{t,y} + \sum_{j>i} h'_x(t_{j-1}, X_{t_{j-1}}^{t,y}) \nabla X_{t_{j-1}}^{t,y} (X_{t_j}^{t,y} - X_{t_{j-1}}^{t,y}) \\ \quad + \sum_{j>i} h(t_{j-1}, X_{t_{j-1}}^{t,y}) (\nabla X_{t_j}^{t,y} - \nabla X_{t_{j-1}}^{t,y}); \\ \partial_y \varphi^\pi(a, x, t, y) = E\{g'(\psi^\pi(a, x, t, y)) \partial_y \psi^\pi(a, x, t, y)\}. \end{cases} \quad (6.12)$$

Since  $h(\cdot, x)$  and  $h'_x(\cdot, x)$  are both bounded and left continuous, applying the Burkholder-Davis-Gundy inequality and the Dominated Convergence Theorem we obtain that

$$\begin{cases} \lim_{|\pi| \rightarrow 0} E\left\{ \left| \sum_{t_j \geq t} h(t_j, X_{t_j}^{t,y}) (X_{t_{j+1}}^{t,y} - X_{t_j}^{t,y}) - \int_t^T h(s, X_{s-}^{t,y}) dX_s^{t,y} \right| \right\} = 0; \\ \lim_{|\pi| \rightarrow 0} E\left\{ \left| \sum_{t_j \geq t} h(t_j, X_{t_j}^{t,y}) (\nabla X_{t_{j+1}}^{t,y} - \nabla X_{t_j}^{t,y}) - \int_t^T h(s, X_{s-}^{t,y}) d\nabla X_s^{t,y} \right| \right\} = 0; \\ \lim_{|\pi| \rightarrow 0} E\left\{ \left| \sum_{t_j \geq t} h'_x(t_j, X_{t_j}^{t,y}) \nabla X_{t_i}^{t,y} (X_{t_{j+1}}^{t,y} - \nabla X_{t_j}^{t,y}) \right. \right. \\ \quad \left. \left. - \int_t^T h'_x(s, X_{s-}^{t,y}) \nabla X_{s-}^{t,y} dX_s^{t,y} \right| \right\} = 0; \end{cases} \quad (6.13)$$

Now if  $a_\pi \rightarrow a$  and  $x_\pi \rightarrow x$  as  $|\pi| \rightarrow 0$ , then (6.13) implies that

$$\begin{cases} \lim_{|\pi| \rightarrow 0} E\{|\psi^\pi(a_\pi, x_\pi, t, y) - \psi(a, x, t, y)|\} = 0; \\ \lim_{|\pi| \rightarrow 0} E\{|\partial_y \psi^\pi(a_\pi, x_\pi, t, y) - \partial_y \psi(a, x, t, y)|\} = 0. \end{cases} \quad (6.14)$$

Further, note that

$$\begin{aligned} & |g'(\psi^\pi(a, x, t, y)) \psi^\pi(a, x, t, y) - g'(\psi(a, x, t, y)) \partial_y \psi(a, x, t, y)| \\ & \leq |g'(\psi^\pi(a, x, t, y)) - g'(\psi(a, x, t, y))| |\partial_y \psi(a, x, t, y)| \\ & \quad + C |\partial_y \psi^\pi(a, x, t, y) - \partial_y \psi(a, x, t, y)|, \end{aligned}$$

applying the Dominated Convergence Theorem we then conclude that

$$\begin{cases} \lim_{|\pi| \rightarrow 0} |\varphi^\pi(a_\pi, x_\pi, t, y) - \varphi(a, x, t, y)| = 0; \\ \lim_{|\pi| \rightarrow 0} |\partial_y \varphi^\pi(a_\pi, x_\pi, t, y) - \partial_y \varphi(a, x, t, y)| = 0. \end{cases} \quad (6.15)$$

Now note that, possibly along a subsequence we must have that, for each  $t \in [0, T]$ ,

$$\sum_{j < i} h(t_{j-1}, X_{t_{j-1}})(X_{t_j} - X_{t_{j-1}}) \rightarrow \int_0^{t-} h(s, X_{s-}) dX_s, \quad \text{a.s.},$$

and  $X_{t_{i-1}} \rightarrow X_{t-}$ , a.s., as  $|\pi| \rightarrow 0$ . Thus, applying the Dominated Convergence Theorem if necessary, we derive from (6.9) that, for all  $t \in [0, T]$

$$\xi_t^{\pi,1} \rightarrow \xi_t^1; \quad \xi_t^{\pi,2} \rightarrow \xi_t^2, \quad \text{as } |\pi| \rightarrow 0,$$

thanks to (6.15), where

$$\begin{cases} \xi_t^1 \triangleq \partial_y \varphi \left( \int_0^{t-} h(s, X_{s-}) dX_s, X_{t-}, t, X_{t-} \right); \\ \xi_t^2 \triangleq b(t, X_{t-}) \int_{\mathbb{R}} \Lambda_t(z) z F(dz), \end{cases} \quad (6.16)$$

and

$$\begin{aligned} \Lambda_t(z) \triangleq & \varphi \left( \int_0^{t-} h(s, X_{s-}) dX_s, X_{t-}, t, X_{t-} + b(t, X_{t-})z \right) \\ & - \varphi \left( \int_0^{t-} h(s, X_{s-}) dX_s, X_{t-}, t, X_{t-} \right). \end{aligned}$$

Furthermore, (6.5) implies that

$$\xi_t = \frac{\sigma^2(t, X_{t-})\xi_t^1 + \xi_t^2}{(\sigma^2 + b^2)(t, X_{t-})}, \quad dP \times dt\text{-a.e.} \quad (6.17)$$

It remains to show that both  $\xi^1$  and  $\xi^2$  have càglàd version. To see this, note that  $X$  is driven by a Lévy process, which has no fixed jump time. Namely, for every  $t \in [0, T]$ , one has  $X_t = X_{t-}$ , a.s. and  $\nabla X_t = \nabla X_{t-}$ , a.s. Recalling (6.1) and (6.11), we see that, for each  $t \in [0, T]$ , it holds almost surely that

$$\begin{aligned} \xi_t^1 &= \partial_y \varphi \left( \int_0^{t-} h(s, X_{s-}) dX_s, X_{t-}, t, X_t \right) \\ &= E \left\{ g' \left( \int_0^T h(s, X_{s-}) dX_s \right) [h(t, X_{t-}) + [\nabla X_{t-}]^{-1} \int_t^T h'_x(s, X_{s-}) \nabla X_{s-} dX_s \right. \\ &\quad \left. + [\nabla X_{t-}]^{-1} \int_t^T h(s, X_{s-}) d\nabla X_s \right] \Big| \mathcal{F}_t \Big\} \\ &= \left[ h(t, X_{t-}) - [\nabla X_{t-}]^{-1} \int_0^{t-} h'_x(s, X_{s-}) \nabla X_{s-} dX_s \right. \\ &\quad \left. - [\nabla X_{t-}]^{-1} \int_0^{t-} h(s, X_{s-}) d\nabla X_s \right] \cdot E \left\{ g' \left( \int_0^T h(s, X_{s-}) dX_s \right) \Big| \mathcal{F}_t \right\} \\ &\quad + [\nabla X_{t-}]^{-1} E \left\{ g' \left( \int_0^T h(s, X_{s-}) dX_s \right) \cdot \left[ \int_0^T h'_x(s, X_{s-}) \nabla X_{s-} dX_s \right. \right. \\ &\quad \left. \left. + \int_0^T h(s, X_{s-}) d\nabla X_s \right] \Big| \mathcal{F}_t \right\}, \end{aligned} \quad (6.18)$$

which clearly has a càglàd version, thanks to the assumption that  $h$  is càglàd with respect to  $t$ .

Furthermore, using (6.1) and (6.11) again we have

$$\begin{cases} |\partial_a \varphi(a, x, t, y)| \leq C; \\ |\partial_x \varphi(a, x, t, y)| \leq C(1 + |y - x|); \\ |\partial_y \varphi(a, x, t, y)| \leq C(1 + |y|), \end{cases} \quad (6.19)$$

and for  $0 \leq t_1 < t_2 \leq T$ , denoting  $X^i = X^{t_i, y}$ ,  $i = 1, 2$ , we have

$$\begin{aligned} & |\varphi(a, x, t_1, y) - \varphi(a, x, t_2, y)| \leq CE\{|\psi(a, x, t_1, y) - \psi(a, x, t_2, y)|\} \\ & \leq CE\left\{|h(t_1, x) - h(t_2, x)||y - x| + \left|\int_{t_1}^{t_2} h(s, X_{s-}^1) dX_s^1\right| \right. \\ & \quad \left. + \left|\int_{t_2}^T [h(s, X_{s-}^1)\sigma(s, X_{s-}^1) - h(s, X_{s-}^2)\sigma(s, X_{s-}^2)] dW_s\right| \right. \\ & \quad \left. + \left|\int_{t_2}^T \int_{\mathbb{R}} [h(s, X_{s-}^1)b(s, X_{s-}^1) - h(s, X_{s-}^2)b(s, X_{s-}^2)] z \tilde{\mu}(ds, dz)\right|\right\} \\ & \leq C\left[|h(t_1, x) - h(t_2, x)||y - x| + (1 + |y|)(t_2 - t_1)^{1/2}\right] \quad (6.20) \\ & \quad + E\left\{\left(\int_{t_2}^T |h(s, X_{s-}^1)\sigma(s, X_{s-}^1) - h(s, X_{s-}^2)\sigma(s, X_{s-}^2)|^2 ds\right)^{\frac{1}{2}}\right\} \\ & \quad + E\left\{\int_{t_2}^T \int_{\mathbb{R}} |h(s, X_{s-}^1)b(s, X_{s-}^1) - h(s, X_{s-}^2)b(s, X_{s-}^2)| |z| F(dz) ds\right\} \\ & \leq C\left[|h(t_1, x) - h(t_2, x)||y - x| + (1 + |y|)(t_2 - t_1)^{\frac{1}{2}}\right. \\ & \quad \left. + E\left(\int_{t_2}^T |X_{s-}^1 - X_{s-}^2|^2 ds\right)^{\frac{1}{2}} + E\left\{\left(\int_{t_2}^T |X_{s-}^1 - X_{s-}^2|^2 ds\right)^{\frac{1}{2}} \left(1 + \sup_{t_2 \leq s \leq T} |X_s^2|\right)\right\}\right. \\ & \quad \left. + E\left\{\int_{t_2}^T |X_{s-}^1 - X_{s-}^2| ds + E\left\{\left(\int_{t_2}^T |X_{s-}^1 - X_{s-}^2| ds\right) \left(1 + \sup_{t_2 \leq s \leq T} |X_s^2|\right)\right\}\right\}\right] \\ & \leq C\left[|h(t_1, x) - h(t_2, x)||y - x| + (1 + |y|^2)(t_2 - t_1)^{\frac{1}{2}}\right], \end{aligned}$$

where  $C > 0$  is again some generic constant that is allowed to vary from line to line. This shows that  $\varphi$  is càglàd with respect to  $t$ . Combining this with (6.19) and (6.9) we see that  $\xi^2$ , whence  $\xi$ , is càglàd.

In the general case that  $g$  and  $h(t, \cdot)$  are only Lipschitz continuous, we can again choose  $g^\varepsilon$  and  $h^\varepsilon$  to be the smooth molifiers with respect to the spatial variable  $x$ , and follow the standard arguments to show that

$$\lim_{\varepsilon \rightarrow 0} E\left\{\int_0^T |\xi_t^\varepsilon - \xi_t|^2 (\sigma^2 + b^2)(t, X_{t-}) dt\right\} = 0. \quad (6.21)$$



Next, since  $(h^\varepsilon)'(\cdot, x)$  is càglàd for each  $x$ , using an argument similar to our previous one, we see that there exists a càglàd process  $\xi^\varepsilon$  of the form:

$$\xi_t^\varepsilon = \frac{\sigma^2(t, X_{t-})\xi_t^{\varepsilon,1} + \xi_t^{\varepsilon,2}}{(\sigma^2 + b^2)(t, X_{t-})}, \quad (6.22)$$

where  $\xi^{\varepsilon,1}$  and  $\xi^{\varepsilon,2}$  are defined in a by now obvious way, and such that

$$\lim_{\varepsilon \rightarrow 0} E \left\{ \int_0^T |\xi_t^{\varepsilon,2} - \xi_t^2|^2 (\sigma^2 + b^2)(t, X_{t-}) dt \right\} = 0; \quad (6.23)$$

where  $\xi^2$  is defined by (6.16). It is not too hard to show, as we did before, that  $\xi^2$  is again càglàd. Define, in light of (6.22),

$$\xi_t^1 \triangleq \frac{\xi_t(\sigma^2 + b^2)(t, X_{t-}) - \xi_t^2}{\sigma^2(t, X_{t-})}.$$

It suffices to show that  $\xi^1$  is càglàd. To this end we denote, by (6.18),

$$\begin{aligned} A_t^\varepsilon &\triangleq \xi_{t+}^{\varepsilon,1} \nabla X_t - h^\varepsilon(t+, X_t) E \left\{ (g^\varepsilon)' \left( \int_0^T h^\varepsilon(s, X_{s-}) dX_s \right) \middle| \mathcal{F}_t \right\} \\ &= E \left\{ (g^\varepsilon)' \left( \int_0^T h^\varepsilon(s, X_{s-}) dX_s \right) \left[ \int_t^T (h^\varepsilon)'_x(s, X_{s-}) \nabla X_{s-} dX_s \right. \right. \\ &\quad \left. \left. + \int_t^T h^\varepsilon(s, X_{s-}) d\nabla X_s \right] \middle| \mathcal{F}_t \right\} \end{aligned} \quad (6.24)$$

For any  $\lambda > 0$ , define  $\Omega_\lambda \triangleq \{\omega : \sup_t \{|\nabla X_t| + \frac{1}{\sigma^2(t, X_t)}\} \leq \lambda\}$ . From (6.21) and (6.23) we know that

$$\begin{aligned} &E \left\{ 1_{\Omega_\lambda} \int_0^T |\xi_t^{\varepsilon,1} \nabla X_{t-} - \xi_t^1 \nabla X_{t-}| dt \right\} \\ &\leq E \left\{ 1_{\Omega_\lambda} \int_0^T \left[ \frac{\sigma^2 + b^2}{\sigma^2} |\xi_t^\varepsilon - \xi_t| |\nabla X_{t-}| + \frac{|\nabla X_{t-}|}{\sigma^2} |\xi_t^{\varepsilon,2} - \xi_t^2| \right] dt \right\} \\ &\leq C(\lambda) \left( E \left\{ 1_{\Omega_\lambda} \int_0^T [|\xi_t^\varepsilon - \xi_t|^2 + |\xi_t^{\varepsilon,2} - \xi_t^2|^2] (\sigma^2 + b^2)(t, X_{t-}) dt \right\} \right)^{\frac{1}{2}} \rightarrow 0, \text{ as } \varepsilon \rightarrow 0. \end{aligned}$$

That is,

$$\lim_{\varepsilon \rightarrow 0} \xi_t^{\varepsilon,1} \nabla X_{t-} = \xi_t^1 \nabla X_{t-}, \quad \text{strongly in } L^1(\Omega_\lambda \times [0, T]). \quad (6.25)$$

Now denote  $G^\varepsilon \triangleq (g^\varepsilon)' \left( \int_0^T h^\varepsilon(s, X_{s-}) dX_s \right)$ . Since it is uniformly bounded, there exists  $G \in L^2(\Omega)$  such that, possibly along a subsequence,  $G^\varepsilon$  converges to  $G$  weakly in  $L^2(\Omega)$ . Noting that  $h^\varepsilon$ 's are uniformly bounded and converge to  $h$  uniformly, we can easily derive

$$\lim_{\varepsilon \rightarrow 0} h^\varepsilon(\cdot+, X_\cdot) E \{ G^\varepsilon | \mathcal{F} \} = h(\cdot+, X_\cdot) E \{ G | \mathcal{F} \}, \quad \text{weakly in } L^2(\Omega \times [0, T]),$$

which, together with (6.25), implies that

$$\lim_{\varepsilon \rightarrow 0} A_{\cdot-}^{\varepsilon} = A^0, \quad \text{weakly in } L^1(\Omega_{\lambda} \times [0, T]), \quad (6.26)$$

where  $A_t^0 \triangleq \xi_t^1 \nabla X_{t-} - h(t, X_{t-}) E\{G | \mathcal{F}_t\}_{\cdot-}$ ,  $t \in [0, T]$ . Now by Mazur's theorem (cf. e.g., [7]), there exists a sequence  $\{B^{n,\lambda}\}$ , where each  $B^{n,\lambda}$  is a convex combinations of  $A^{\varepsilon}$ , such that  $B^{n,\lambda}$  converges to  $A^0$  strongly in  $L^1(\Omega_{\lambda} \times [0, T])$ . Since  $\Omega_{\lambda} \uparrow \Omega$ , a simple diagonalization procedure shows that there exists a sequence  $\{B^n\}$ , where each  $B^n$  is a convex combination of  $B^{n,\lambda}$ 's (hence still a convex combination of  $A^{\varepsilon}$ 's!), such that  $B^n$  converges to  $A^0$ , strongly in  $L^1(\Omega_{\lambda} \times [0, T])$  for all  $\lambda$ , as  $n \rightarrow \infty$ . Consequently, for a.s.  $\omega \in \Omega$ ,  $B^n(\omega)$  converges to  $A^0(\omega)$  in measure. On the other hand, by the definition (6.24) it is not hard to check that  $\{A^{\varepsilon}\}$ , whence  $\{B^n\}$ , is tight under the Meyer-Zheng topology. An argument similar to that in [11] shows that  $A^0$  has a càglàd version. The proof is now complete.  $\blacksquare$

**Remark 6.2** If we let  $h(t, x) = 1_{[0, t_0]}(t)$  for some  $t_0 \in [0, T]$ , then the assumptions (i)–(iii) of the theorem are all satisfied. Therefore our result covers the special case when  $\Phi(X) = g(X_{t_0})$ ; that is, the case considered in [8].

**Remark 6.3** In the theorem we assumed that  $h$  is bounded so that the random variables involved are all square integrable. An alternative assumption (of (i)) could be that (i')  $\int_{\mathbb{R}} z^4 F(dz) < \infty$ . We leave the details to the interested readers.

## 7 General Case

In this section we shall summarize the results from previous sections to study some more general situations. We present them in two theorems.

**Theorem 7.1** *Assume that  $\Phi(X) = g(\int_0^T \mathbf{h}_1(t, X_t) dt, \int_0^T \mathbf{h}_2(t, X_{t-}) dX_t)$ , where*

$$\mathbf{h}_1(t, x) \triangleq (h_1^1, \dots, h_m^1)(t, x), \quad \mathbf{h}_2(t, x) \triangleq (h_1^2, \dots, h_n^2)(t, x).$$

*Assume further that  $\mathbf{h}_2$  is bounded and càglàd with respect to  $t$ . Define  $\varphi : \mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R} \times [0, T] \times \mathbb{R} \mapsto \mathbb{R}$  as follows:*

$$\begin{aligned} & \varphi(A_1, A_2, x, t, y) \\ & \triangleq E \left\{ g(A_1 + \int_t^T \mathbf{h}_1(s, X_s^{t,y}) ds, A_2 + \mathbf{h}_2(t, x)(y - x) + \int_t^T \mathbf{h}_2(s, X_s^{t,y}) dX_s^{t,y}) \right\}. \end{aligned}$$

i) If  $g$ ,  $\mathbf{h}_1$  and  $\mathbf{h}_2$  are continuously differentiable with uniformly bounded derivatives with respect to the spatial variables, then  $\xi$  is càglàd . To be more precise, one has

$$\begin{aligned} & \xi_t(\sigma^2 + b^2)(t, X_{t-}) \\ = & \sigma^2(t, X_{t-})\varphi_y\left(\int_0^{t-} \mathbf{h}_1(s, X_s)ds, \int_0^{t-} \mathbf{h}_2(s, X_{s-})dX_s, X_{t-}, t, X_{t-}\right) \quad (7.1) \\ & + b^2(t, X_{t-}) \int_{\mathbb{R}} \int_0^1 \varphi_y\left(\int_0^{t-} \mathbf{h}_1(s, X_s)ds, \right. \\ & \left. \int_0^{t-} \mathbf{h}_2(s, X_{s-})dX_s, X_{t-}, t, X_{t-} + b(t, X_{t-})zu\right)z^2duF(dz), \end{aligned}$$

where

$$\begin{aligned} \varphi_y(A_1, A_2, x, t, y) = & E\left\{\langle g_1, \int_t^T (\mathbf{h}_1)'_x(s, X_s^{t,y})\nabla X_s^{t,y}ds \rangle \right. \quad (7.2) \\ & \left. + \langle g_2, \mathbf{h}_2(t, x) + \int_t^T (\mathbf{h}_2)'_x(s, X_{s-}^{t,y})\nabla X_{s-}^{t,y}dX_s^{t,y} + \int_t^T \mathbf{h}_2(s, X_{s-}^{t,y})d\nabla X_s^{t,y} \rangle\right\}. \end{aligned}$$

ii) The same holds when  $g$ ,  $\mathbf{h}_1$  and  $\mathbf{h}_2$  are differences of two functions convex with respect to the spatial variables, with right derivatives bounded and all the derivatives appearing in i) are replaced by the corresponding right derivative, provided we have  $X_s^{t,y}$ ,  $\int_0^T h_i^1(t, X_t)dt$  and  $\int_0^T h_j^2(t, X_t)dX_t$  have no atoms.

*Proof.* (i) If  $(\mathbf{h}_2)'_x$  is also càglàd with respect to  $t$ , then similar to (3.8) one can show that (7.1) also holds. In general we can again approximate  $\mathbf{h}_2$  by the molifiers to conclude the same result.

(ii) This is a direct consequence of the arguments of Theorem 2.6-b) of [8].  $\blacksquare$

**Theorem 7.2** Assume that  $\Phi(X) = g(\Phi_1(X), \dots, \Phi_n(X))$ , where  $g$  is uniformly Lipschitz, where  $\Phi_i(X)$ 's are of the form as those in Theorem 5.1, Theorem 5.2 or Theorem 6.1. Then,  $\xi$  is càglàd .

*Proof.* To simplify the presentation, let us assume that

$$\Phi(X) = g(\Phi_1(X), \sup_{0 \leq t \leq T} h(t, X_t), \int_0^T k(t, X_{t-})dX_t),$$

where  $\Phi_1$  satisfies  $L^1$ -Lipschitz condition (5.1), and  $h, k$  satisfy the conditions in Theorem 5.1 and Theorem 6.1, respectively. Define  $\varphi : \mathbb{D} \times \mathbb{R}^2 \times [0, T] \times \mathbb{R} \mapsto \mathbb{R}$  to

be such that

$$\varphi(\mathbf{x}, a, x, t, y) \triangleq E \left\{ g(\Phi_1(\mathbf{x}1_{[0,t]} + X^{t,y}1_{[t,T]}), \sup_{0 \leq s < t} h(s, \mathbf{x}(s)) \vee \sup_{t \leq s \leq T} h(s, X_s^{t,y}), \right. \\ \left. a + k(t, x)(y - x) + \int_t^T k(s, X_{s-}^{t,y}) dX_s^{t,y} \right\}.$$

For any partition  $\pi : 0 = t_0 < \dots < t_n = T$ , define  $\varphi^\pi$  similar to (4.13). Then, combining the arguments in the previous sections we see that  $\xi^\pi$  will converge to  $\xi$  in measure, as  $|\pi| \rightarrow 0$ , and that  $\xi$  is càglàd . ■

## References

- [1] Chan, T. (1999) *Pricing Contingent Claims on Stocks Driven by Lévy Processes*, *Ann. Applied Proba.*, **9**, 504–528.
- [2] Çinlar, E. and Jacod, J. (1981) *Representation of semimartingale Markov processes in terms of Wiener processes and Poisson random measures*, In *Seminar on Stochastic Processes*, Birkhauser, Boston, 159–242,
- [3] Çinlar, E., Jacod, J., Protter, P., and Sharpe, M. (1980) *Semimartingales and Markov processes*, *Z. Wahrsch. Verw. Geb.* **52**, 161–220.
- [4] Dellacherie C. and Meyer, P. A. (1982) *Probabilities and Potential B*, North Holland, Amsterdam.
- [5] Delbaen, F. and Schachermayer, W. (1988) *The Fundamental Theorem for Unbounded Stochastic Processes*, *Math. Ann.*, **312**, 215–250.
- [6] Föllmer, H and Schweizer, M. (1991) *Hedging of Contingent Claims under Incomplete Information*, *Applied Stochastic Analysis*, M.H.A. Davis and R.J. Elliott, eds., Gordon and Breach, London, 389–414.
- [7] Hille, E. and Phillips, R.S. (1957) *Functional Analysis and Semi-Groups*, American Mathematical Society.
- [8] Jacod, J., Méléard S. and Protter P. (2000), *Explicit form and robustness of martingale representations*, to appear in *Annals of Probability*.

- [9] Jacod, J. and Shiryaev (1981) *Limit Theorems for Stochastic Processes*, Springer-Verlag, Heidelberg.
- [10] Ma, J. and Zhang, J. (2000) *Representation theorems for backward stochastic differential equations*, Preprint.
- [11] Ma, J. and Zhang, J. (2000) *Path regularity for solutions of backward SDEs*, Preprint.
- [12] Meyer P. A. and Zheng, W. (1984) *Tightness criteria for laws of semimartingales*, *Ann. Inst. Henri Poincaré*, **20**, 353–372.
- [13] Protter, P. (1990) *Stochastic Integrals and Differential Equations, a New Approach*, Springer-Verlag, Heidelberg.
- [14] Willinger, W. and Taqqu, M.S. (1989) *Pathwise stochastic integration and applications to the theory of continuous trading*, *Stoch. Proc. Appl.*, **32**, 253–280.