CARF Working Paper

CARF-F-445

Asymptotic Expansion for Forward-Backward SDEs with Jumps

Masaaki Fujii
Quantitative Finance Course, Graduate School of Economics,
The University of Tokyo

Akihiko Takahashi
Quantitative Finance Course, Graduate School of Economics,
The University of Tokyo

This version: 7 September, 2018

CARF is presently supported by Dai-ichi Mutual Life Insurance Company, Nomura Holdings, Inc., Sumitomo Mitsui Banking Corporation, MUFG Bank, Finatext Ltd., The University of Tokyo Edge Capital Co., Ltd., Sompo Holdings, Inc., Tokio Marine & Nichido Fire Insurance Co., Ltd. and Nippon Life Insurance Company. This financial support enables us to issue CARF Working Papers.

Asymptotic Expansion for Forward-Backward SDEs with Jumps *

Masaaki Fujii[†] & Akihiko Takahashi[‡]

This version: 7 September, 2018

Abstract

This work provides a semi-analytic approximation method for decoupled forward-backward SDEs (FBSDEs) with jumps. In particular, we construct an asymptotic expansion method for FBSDEs driven by the random Poisson measures with σ-finite compensators as well as the standard Brownian motions around the small-variance limit of the forward SDE. We provide a semi-analytic solution technique as well as its error estimate for which we only need to solve essentially a system of linear ODEs. In the case of a finite jump measure with a bounded intensity, the method can also handle state-dependent and hence non-Poissonian jumps, which are quite relevant for many practical applications.

Keywords : BSDE, jumps, random measure, asymptotic expansion, Lévy process

1 Introduction

Since it was introduced by Bismut (1973) [5] and Pardoux & Peng (1990) [42], the backward stochastic differential equation (BSDE) has attracted many mathematicians because of its deep connections to non-linear partial differential equations. There now exist excellent reviews such as El Karoui & Mazliak (eds.) (1997) [17], Ma & Yong (2000) [38], and Pardoux & Rascanu (2014) [44] for interested readers. BSDEs also have a wide variety of applications to financial as well as operational problems; El Karoui et al. (1997) [18], Lim (2004) [36], Jeanblanc & Hamadène (2007) [28], Cvitanić & Zhang (2013) [11], Touzi (2013) [54] and Crépey, Bielecki & Brigo (2014) [8] to mention only a few. As for BSDEs with jumps, see for example, Barles, Buckdahn & Pardoux (1997) [2], Royer (2006) [49], Crepey & Matoussi (2008) [9], Morlais (2010) [41], Delong (2013) [12] and Quenez & Sulem (2013) [48].

The last financial crisis and a bunch of new regulations that followed have made various problems involving BSDEs such as XVAs, risk measures, optimal executions in illiquid markets and the development for their efficient numerical computation scheme the central issues in the financial industry. Although the backward Monte-Carlo simulation scheme has been proposed and studied by many researchers such as, Bouchard & Touzi (2004) [7], Zhang (2004) [55], Gobet et al. (2004) [27] and Bender & Denk (2007) [3] for continuous

*Forthcoming in *Stochastics*. All the contents expressed in this research are solely those of the authors and do not represent any views or opinions of any institutions. The authors are not responsible or liable in any manner for any losses and/or damages caused by the use of any contents in this research.

[†]Quantitative Finance Course, Graduate School of Economics, The University of Tokyo.

[‡]Quantitative Finance Course, Graduate School of Economics, The University of Tokyo.

BSDEs, and Bouchard & Elie (2008) [6] for BSDEs with jumps, it has not yet become a standard tool for practitioners due to its computational burden. In particular, we can only find simple one-dimensional examples using the Poisson process instead of a random measure in the existing literature. See, for example, Elie (2006) [16] and Lejay et.al. (2014) [34]. See also the discussion in [8] and Crépey & Song (2015) [10] regarding the problems in the existing computation scheme when applied to practical problems [1]. Moreover, in certain applications such as mean-variance hedging and multiple dependent defaults, the solution of one BSDE appears in the driver of another BSDE [2]. In such a case, deriving an analytic approximation for the first BSDE seems the only possibility to obtain a numerical result within reasonable computational time.

As one possible approach to these problems, the current work contributes by providing a straightforward semi-analytic approximation method for BSDEs with jumps, which are especially difficult and time-consuming to evaluate by the standard Monte-Carlo scheme. We develop an asymptotic expansion method for decoupled forward-backward SDEs (FBSDEs) with Lipschitz drivers and the Poisson random measures in addition to the standard Brownian motions. We propose an expansion around a small-variance limit of the forward SDE. It starts from solving a non-linear ODE corresponding to the BSDE where every forward component is replaced by the deterministic mean process. Every higher order approximation yields a linear FBSDE, which can be solved semi-analytically essentially by a system of linear ODEs. More precisely, the approximate solution of the BSDE including the martingale components is explicitly given by a polynomial in the stochastic flows of the forward process whose coefficients can be computed by the linear ODEs.

In order to justify the approximation method and its error estimate, we use the results of Kruse & Popier (2015) [33] for a priori estimates and the existence of unique $\mathbb{L}^p$-solution of BSDEs with jumps, Delong & Imkeller (2010) [13] and Delong [12] for the representation theorem based on the Malliavin's derivative, as well as the idea of Pardoux & Peng (1992) [43] and Ma & Zhang (2002) [39] for controlling the sup-norm of the martingale integrands of the BSDEs. In the case of a finite jump measure with a bounded intensity, the method can also be applied to a system with state-dependent and hence non-Poissonian jumps, which are quite relevant for many practical applications. A closed-form expression of the approximation up to an arbitrary higher order term is available when the forward SDE belongs to (time-inhomogeneous) exponential Lévy type. The current work also serves as a justification of a polynomial expansion method proposed in Fujii (2015) [20] for a certain class of models, which provides a couple of interesting numerical examples.

The organization of the paper is as follows: Section 2 gives some preliminaries, Section 3 the setup of the interested FBSDEs and the representation theorem based on Malliavin's derivative, Section 4 the asymptotic expansion and its error estimate, and finally Section 5 gives the concrete implementation of the scheme. Appendices A and B summarize the relevant a priori estimates, and Appendix C provides the smooth approximation theorem for the FBSDEs, which justifies the assumptions used in the main text.

Remark 1.1. *As for forward SDEs, the asymptotic expansion method around a small-variance limit has been applied to a variety of financial problems. It has been shown, in various numerical examples, that the first few terms of expansion are enough to achieve*

[1] In [10], the authors successfully applied the asymptotic expansion method proposed in [22, 23] to a collateralized debt obligation with 120 underlying names to evaluate credit/funding valuation adjustments.

[2] See Mania & Tevzadze (2003) [40], Pham (2010) [46] and Fujii (2015) [21] for concrete examples.

accurate approximation for option pricing with typical volatilities ranging from 10% to 20% and maturities up to a few years. See a review Takahashi (2015) [50] for the details and a comprehensive list of literature.

Remark 1.2. *The current work can be extended in couple of ways. Firstly, based on the result of Fujii & Takahashi (2017) [25], a similar asymptotic expansion may be justified for a BSDE with a quadratic-exponential growth driver and a bounded terminal condition. This would be done by replacing the estimates of the standard Lipschitz BSDEs with those of local Lipschitz BSDEs with $\mathbb{H}^2$-BMO coefficients. It may also be possible to develop a sub-stepping scheme similar to those in Fujii (2014) [19] and Takahashi & Yamada (2015) [52], which can handle higher volatilities and longer maturities. See an initial attempt in a diffusion setup with quadratic growth driver by Fujii & Takahashi (2016) [24].*

2 Preliminaries

2.1 General Setting

$T > 0$ is some bounded time horizon. The space $(\Omega_W, \mathcal{F}_W, \mathbb{P}_W)$ is the usual canonical space for an l-dimensional Brownian motion equipped with the Wiener measure $\mathbb{P}_W$. $(\Omega_\mu, \mathcal{F}_\mu, \mathbb{P}_\mu)$ denotes a product of canonical spaces $\Omega_\mu := \Omega^1_\mu \times \cdots \times \Omega^k_\mu$, $\mathcal{F}_\mu := \mathcal{F}^1_\mu \times \cdots \times \mathcal{F}^k_\mu$ and $\mathbb{P}^1_\mu \times \cdots \times \mathbb{P}^k_\mu$ with some integer $k \geq 1$, on which each μ^i is a Poisson measure with a compensator $\nu^i(dz)dt$. Here, $\nu^i(dz)$ is a σ-finite measure on $\mathbb{R}_0 := \mathbb{R}\backslash\{0\}$ satisfying $\int_{\mathbb{R}_0} |z|^2\nu^i(dz) < \infty$. Throughout the paper, we work on the filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$, where the space $(\Omega, \mathcal{F}, \mathbb{P})$ is the product of the canonical spaces $(\Omega_W \times \Omega_\mu, \mathcal{F}_W \times \mathcal{F}_\mu, \mathbb{P}_W \times \mathbb{P}_\mu)$, and that the filtration $\mathbb{F} := (\mathcal{F}_t)_{t\in[0,T]}$ is the canonical filtration completed for $\mathbb{P}$ and satisfying the usual conditions. In this construction, $(W, \mu^1, \cdots, \mu^k)$ are independent and it is well-know that *the predictable representation property* holds [3]. We use a vector notation $\mu(\omega, dt, dz) := (\mu^1(\omega, dt, dz^1), \cdots, \mu^k(\omega, dt, dz^k))$. The compensated Poisson measure is denoted by $\widetilde{\mu} := \mu - \nu$. We represent the $\mathbb{F}$-predictable σ-field on $\Omega \times [0, T]$ by $\mathcal{P}$.

2.2 Notation

Let C_p denote a generic constant, which may change line by line, depending on p, T and the Lipschitz constants and the bounds of the relevant functions. For any integer $r \geq 1$, let us introduce a sup-norm for a $\mathbb{R}^r$-valued function $x : [0, T] \to \mathbb{R}^r$ as

$$||x||_{[a,b]} := \sup\{|x_t|, t \in [a, b]\}$$

and write $||x||_t := ||x||_{[0,t]}$.

Let us introduce the following spaces for stochastic processes for $p \geq 2$:
• $\mathbb{S}^p_r[s, t]$ is the set of $\mathbb{R}^r$-valued adapted càdlàg processes X such that

$$||X||_{\mathbb{S}^p_r[s,t]} := \mathbb{E}\left[||X(\omega)||^p_{[s,t]}\right]^{1/p} < \infty \ .$$

[3] See, for example, Chapter XIII in [30]. If one assumes the predictable representation property, this construction is irrelevant.

• $\mathbb{H}^p_r[s,t]$ is the set of progressively measurable $\mathbb{R}^r$-valued processes Z such that

$$||Z||_{\mathbb{H}^p_r[s,t]} := \mathbb{E}\Bigl[\Bigl(\int_s^t |Z_u|^2 du\Bigr)^{p/2}\Bigr]^{1/p} < \infty.$$

• $\mathbb{H}^p_{r,\nu}[s,t]$ is the set of functions $\psi = \{(\psi)_{i,j}, 1 \le i \le r, 1 \le j \le k\}$, $(\psi)_{i,j} : \Omega\times[0,T]\times\mathbb{R}_0 \to \mathbb{R}$ which are $\mathcal{P}\times\mathcal{B}(\mathbb{R}_0)$-measurable and satisfy

$$||\psi||_{\mathbb{H}^p_{r,\nu}[s,t]} := \mathbb{E}\Bigl[\Bigl(\sum_{i=1}^k \int_s^t \int_{\mathbb{R}_0} |\psi_u^{\cdot,i}(z)|^2 \nu^i(dz)du\Bigr)^{p/2}\Bigr]^{1/p} < \infty.$$

For simplicity, we use the notation $(E,\mathcal{E}) := (\mathbb{R}_0^k, \mathcal{B}(\mathbb{R}_0)^k)$ and denote the above maps $\{(\psi)_{i,j}, 1 \le i \le r, 1 \le j \le k\}$ by $\psi : \Omega\times[0,T]\times E \to \mathbb{R}^{r\times k}$ and say ψ is $\mathcal{P}\times\mathcal{E}$-measurable without referring to each component. We also use the notation such that

$$\int_s^t \int_E \psi_u(z)\widetilde{\mu}(du,dz) := \sum_{i=1}^k \int_s^t \int_{\mathbb{R}_0} \psi_u^i(z)\widetilde{\mu}^i(du,dz)$$

for simplicity. The similar abbreviation is used also for the integral with μ and ν. When we use E and $\mathcal{E}$, one should always interpret it in this way so that the integral with the k-dimensional Poisson measure does make sense. On the other hand, when we use the range $\mathbb{R}_0$ with the integrators $(\widetilde{\mu}, \mu, \nu)$, for example,

$$\int_{\mathbb{R}_0} \psi_u(z)\nu(dz) := \Bigl(\int_{\mathbb{R}_0} \psi_u^i(z)\nu^i(dz)\Bigr)_{1\le i\le k}$$

we interpret it as a k-dimensional vector.
• $\mathcal{K}^p[s,t]$ is the set of functions (Y,Z,ψ) in the space $\mathbb{S}^p[s,t]\times\mathbb{H}^p[s,t]\times\mathbb{H}^p_\nu[s,t]$ with the norm defined by

$$||(Y,Z,\psi)||_{\mathcal{K}^p[s,t]} := \Bigl(||Y||^p_{\mathbb{S}^p[s,t]} + ||Z||^p_{\mathbb{H}^p[s,t]} + ||\psi||^p_{\mathbb{H}^p_\nu[s,t]}\Bigr)^{1/p}.$$

• $\mathbb{L}^2(E,\mathcal{E},\nu:\mathbb{R}^r)$ is the set of $\mathbb{R}^{r\times k}$-valued $\mathcal{E}$-measurable functions U satisfying

$$\begin{aligned} ||U||_{\mathbb{L}^2(E)} &:= \Bigl(\int_E |U(z)|^2\nu(dz)\Bigr)^{1/2} \\ &:= \Bigl(\sum_{i=1}^k \int_{\mathbb{R}_0} |U^{\cdot,i}(z)|^2\nu^i(dz)\Bigr)^{1/2} < \infty\ . \end{aligned}$$

We frequently omit the subscripts for its dimension r and the time interval $[s,t]$ when they are obvious in the context.

We use the notation of partial derivatives such that

$$\partial_\epsilon := \frac{\partial}{\partial\epsilon}, \quad \partial_x := (\partial_{x_1},\cdots,\partial_{x_d}) := \Bigl(\frac{\partial}{\partial x_1},\cdots,\frac{\partial}{\partial x_d}\Bigr)$$
$$\partial_x^2 := \partial_{x,x} := \Bigl(\frac{\partial^2}{\partial x_i\partial x_j}\Bigr)_{i,j=\{1,\cdots,d\}}$$

and similarly for every higher order derivative without detailed indexing. We suppress the obvious summation of indexes throughout the paper for notational simplicity.

3 Forward and Backward SDEs

3.1 The setup and some standard results

We work in the filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ defined in the last section. Let us introduce a d-dimensional forward SDE of $(X_s^{t,x,\epsilon}, s \in [t,T])$ with the initial data $(t,x) \in [0,T] \times \mathbb{R}^d$ and a small perturbation parameter $\epsilon \in [0,1]$, and an m-dimensional Markovian BSDE driven by $X^{t,x,\epsilon}$:

$$X_s^{t,x,\epsilon} = x + \int_t^s b(r, X_r^{t,x,\epsilon}, \epsilon)dr + \int_t^s \sigma(r, X_r^{t,x,\epsilon}, \epsilon)dW_r + \int_t^s \int_E \gamma(r, X_{r-}^{t,x,\epsilon}, z, \epsilon)\widetilde{\mu}(dr, dz) \tag{3.1}$$

$$\begin{aligned} Y_s^{t,x,\epsilon} &= \xi(X_T^{t,x,\epsilon}) + \int_s^T f\Big(r, X_r^{t,x,\epsilon}, Y_r^{t,x,\epsilon}, Z_r^{t,x,\epsilon}, \int_{\mathbb{R}_0} \rho(z)\psi_r^{t,x,\epsilon}(z)\nu(dz)\Big)dr \\ &\quad - \int_s^T Z_r^{t,x,\epsilon}dW_r - \int_s^T \int_E \psi_r^{t,x,\epsilon}(z)\widetilde{\mu}(dr, dz), \end{aligned} \tag{3.2}$$

for $s \in [t,T]$. Here, $b : [0,T] \times \mathbb{R}^d \times \mathbb{R} \to \mathbb{R}^d$, $\sigma : [0,T] \times \mathbb{R}^d \times \mathbb{R} \to \mathbb{R}^{d\times l}$ and $\gamma : [0,T] \times \mathbb{R}^d \times E \times \mathbb{R} \to \mathbb{R}^{d\times k}$ for the forward SDE and $\xi : \mathbb{R}^d \to \mathbb{R}^m$, $f : [0,T] \times \mathbb{R}^d \times \mathbb{R}^m \times \mathbb{R}^{m\times l} \times \mathbb{R}^{m\times k} \to \mathbb{R}^m$ and $\rho : E \to \mathbb{R}^k$ for the BSDE are measurable functions.

We shall specify the dependence of (b, σ, γ) in ϵ more explicitly in Section 5.1, where we arrange it so that $X^{t,x,\epsilon}$ becomes deterministic process in the limit of $\epsilon \to 0$. The main goal of the current paper is to obtain the Taylor expansion of the solution $(X^{t,x,\epsilon}, Y^{t,x,\epsilon}, Z^{t,x,\epsilon}, \psi^{t,x,\epsilon})$ around $\epsilon = 0$ and the associated error estimates. Let us fix the order of the highest expansion by $n_{\max}$ $(\in \mathbb{N})$ in the reminder of the paper. For notational simplicity, let us define $n^{\rm ae} := n_{\max} + 2$. [4]

Let us also introduce the function $\eta : \mathbb{R} \to \mathbb{R}$ by $\eta(z) := 1 \wedge |z|$. Now, we make the following assumptions.

Assumption 3.1. *The functions $b(t,x,\epsilon), \sigma(t,x,\epsilon)$ and $\gamma(t,x,z,\epsilon)$ are continuous in all their arguments and $n^{\rm ae}$-time differentiable in (x,ϵ) with continuous derivatives. Furthermore, there exists some positive constant K such that*
(i) for every $0 \le m \le n^{\rm ae}$, $|\partial_\epsilon^m b(t,0,\epsilon)| + |\partial_\epsilon^m \sigma(t,0,\epsilon)| \le K$ uniformly in $(t,\epsilon) \in [0,T] \times [0,1]$,
(ii) for every $1 \le n \le n^{\rm ae}, 0 \le m \le n^{\rm ae}$, $|\partial_x^n \partial_\epsilon^m b(t,x,\epsilon)| + |\partial_x^n \partial_\epsilon^m \sigma(t,x,\epsilon)| \le K$ uniformly in $(t,x,\epsilon) \in [0,T] \times \mathbb{R}^d \times [0,1]$,
(iii) for every $0 \le m \le n^{\rm ae}$ and column $1 \le i \le k$, $|\partial_\epsilon^m \gamma_{\cdot,i}(t,0,z,\epsilon)|/\eta(z) \le K$ uniformly in $(t,z,\epsilon) \in [0,T] \times \mathbb{R}_0 \times [0,1]$,
(iv) for every $1 \le n \le n^{\rm ae}, 0 \le m \le n^{\rm ae}$ and column $1 \le i \le k$, $|\partial_x^n \partial_\epsilon^m \gamma_{\cdot,i}(t,x,z,\epsilon)|/\eta(z) \le K$ uniformly in $(t,x,z,\epsilon) \in [0,T] \times \mathbb{R}^d \times \mathbb{R}_0 \times [0,1]$.

Assumption 3.2. *There exist some positive constants K, q such that*
(i) $\xi(x)$ is $n^{\rm ae}$-time differentiable in x with continuous derivatives. Moreover, it has at

[4] The additional factor +2 (instead of +1) arises basically from the need to bound the approximation error for the control variables (Z, ψ).

most polynomial growth $|\partial_x^n \xi(x)| \leq K(1+|x|^q)$ $x \in \mathbb{R}^d$ *for every* $0 \leq n \leq n^{\rm ae}$,
(ii) $|\rho_i(z)| \leq K\eta(z)$ *for every* $1 \leq i \leq k$ *and* $z \in \mathbb{R}_0$,
(iii) $f(t,x,y,z,u)$ *is continuous in all its arguments and* $n^{\rm ae}$*-time differentiable in* (x,y,z,u) *with continuous derivatives. Moreover, every partial derivative only in* x *has at most polynomial growth* $|\partial_x^n f(t,x,y,z,u)| \leq K(1+|x|^q), 1 \leq n \leq n^{\rm ae}$ *as well as all the other partial derivatives are bounded by* K*, uniformly in* $(t,x,y,z,u) \in [0,T] \times \mathbb{R}^d \times \mathbb{R}^m \times \mathbb{R}^{m\times l} \times \mathbb{R}^m$,
(iv) $|f(t,x,0,0,0)| \leq K(1+|x|^q)$ *for every* $x \in \mathbb{R}^d$ *uniformly in* $t \in [0,T]$.

We define $(\partial_x X_s^{t,x,\epsilon}, s \in [t,T])$ as the solution of the SDE (if exists) given by a formal differentiation:

$$\partial_x X_s^{t,x,\epsilon} = \int_t^s \partial_x b(r, X_r^{t,x,\epsilon}, \epsilon)\partial_x X_r^{t,x,\epsilon} dr + \int_t^s \partial_x \sigma(r, X_r^{t,x,\epsilon}, \epsilon)\partial_x X_r^{t,x,\epsilon} dW_r$$
$$+ \int_t^s \int_E \partial_x \gamma(r, X_r^{t,x,\epsilon}, z, \epsilon)\partial_x X_r^{t,x,\epsilon} \widetilde{\mu}(dr,dz) \ , \tag{3.3}$$

similarly for $(\partial_\epsilon X_s^{t,x,\epsilon}, s \in [t,T])$ and every higher order flow $(\partial_x^n \partial_\epsilon^m X_s^{t,x,\epsilon}, s \in [t,T])_{m,n\geq 0}$.

Proposition 3.1. *Under Assumption 3.1, the SDE (3.1) has a unique solution* $X^{t,x,\epsilon} \in \mathbb{S}_d^p[t,T]$ $\forall p \geq 2$*. Furthermore, for* $0 \leq n, m \leq n^{\rm ae}$*, every* (n,m)*-time classical differentiation of* $X^{t,x,\epsilon}$ *in* (x,ϵ) *is well defined and given by* $\partial_x^n \partial_\epsilon^m X^{t,x,\epsilon} \in \mathbb{S}^p_{d^{n+1}}[t,T]$ $\forall p \geq 2$*, which is a unique solution of the corresponding SDE defined by the formal differentiation of the coefficients as (3.3).*

Proof. The existence of a unique solution $X^{t,x,\epsilon} \in \mathbb{S}_d^p[t,T]$ $\forall p \geq 2$ is standard and can be proved by Lemma A.3. Since every SDE is linear, it is not difficult to recursively show that the same conclusion holds for every $\partial_x^n \partial_\epsilon^m X^{t,x,\epsilon}$. The agreement with the classical differentiation can be proved by following the arguments in Theorem 3.1 of Ma & Zhang (2002) [39]. In particular, one can show

$$\lim_{h\to 0} \mathbb{E}||\nabla X^h - \partial_x X^{t,x,\epsilon}||^2_{[t,T]} = 0$$

where $\nabla X_s^h := \dfrac{X_s^{t,x+h,\epsilon} - X_s^{t,x,\epsilon}}{h}$ ($d = 1$ for simplicity) and similar relations for every higher order derivatives in (x,ϵ). □

Proposition 3.2. *Under Assumptions 3.1 and 3.2, the BSDE (3.2) has a unique solution* $(Y^{t,x,\epsilon}, Z^{t,x,\epsilon}, \psi^{t,x,\epsilon})$ *which belongs to* $\mathbb{S}_m^p[t,T] \times \mathbb{H}^p_{m\times l}[t,T] \times \mathbb{H}^p_{m,\nu}[t,T]$ $\forall p \geq 2$*. Furthermore, it also satisfies*

$$||\hat{\Theta}^{t,x,\epsilon}||_{\mathcal{K}^p[t,T]} \leq C_p(1+|x|^q) \tag{3.4}$$

for every $p \geq 2$.

Proof. The existence of a unique solution follows from Lemma B.2. In addition, one has

$$||\hat{\Theta}^{t,x,\epsilon}||^p_{\mathcal{K}^p[t,T]} \leq C_p \mathbb{E}\left[|\xi(X_T^{t,x,\epsilon})|^p + \Big(\int_t^T |f(s, X_s^{t,x,\epsilon}, 0,0,0)|ds\Big)^p\right]$$

and hence one obtains the desired result by Lemma A.3 and the assumption of polynomial growth of $\xi(x)$ and $f(\cdot, x, 0,0,0)$. □

To lighten the notation, we use the following symbol to represent the collective arguments:

$$\Theta_r^{t,x,\epsilon} := \Big(X_r^{t,x,\epsilon}, Y_r^{t,x,\epsilon}, Z_r^{t,x,\epsilon}, \int_{\mathbb{R}_0} \rho(z)\psi_r^{t,x,\epsilon}(z)\nu(dz)\Big)$$
$$\hat{\Theta}_r^{t,x,\epsilon} := \Big(Y_r^{t,x,\epsilon}, Z_r^{t,x,\epsilon}, \int_{\mathbb{R}_0} \rho(z)\psi_r^{t,x,\epsilon}(z)\nu(dz)\Big).$$

We also use $\partial_\Theta := (\partial_x, \partial_y, \partial_z, \partial_u)$, $\partial_{\hat{\Theta}} := (\partial_y, \partial_z, \partial_u)$ and similarly for their higher order derivatives.

Remark 3.1. *Let us remark on the practical implications of the Assumptions 3.1 and 3.2, since some readers may find that the smoothness assumptions are too restrictive. In Appendix C, we prove a smooth approximation theorem for FBSDEs which justifies Assumptions 3.1 and 3.2 whenever the standard Lipschitz conditions are satisfied.*

Since the financial problems relevant for BSDEs are inevitably non-linear, we are forced to consider in a portfolio level. Thus, ξ and f are likely to be given by complicated piecewise linear functions, which involve a large number of non-smooth points. The first step we can do is to approximate these functions by smooth ones by introducing mollifiers or projecting onto Chebyshev polynomials, for example. In the industry, this is quite common even for linear products such as a digital option to make delta hedging feasible in practice. A small additional fee arising from a mollifier is charged to a client as a hedging cost. It is also used for CVA evaluation by Henry-Labordère (2012) [29].

3.2 Representation theorem for BSDEs

We define the Malliavin derivatives $D_{t,z}$ according to the conventions used in Section 3 of Delong & Imkeller (2010) [13] and Section 2.6 of Delong (2013) [12] (with $\sigma = 1$). See also Di Nunno et al (2009) [14] for details and other applications.

According to their definition, if the random variable $H(\cdot,\omega_\mu)$ is differentiable in the sense of classical Malliavin's calculus for $\mathbb{P}_\mu$-a.e. $\omega_\mu \in \Omega_\mu$, then we have the relation

$$D_{t,0}H(\omega_W,\omega_\mu) = D_t H(\cdot,\omega_\mu)(\omega_W)\ ,$$

where $D_\cdot$ is the Malliavin's derivative with respect to the Wiener direction. For the definition $D_{t,z}H$ with $z \neq 0$, the increment quotient operator is introduced

$$\mathcal{I}_{t,z}H(\omega_W,\omega_\mu) := \frac{H(\omega_W,\omega_\mu^{t,z}) - H(\omega_W,\omega_\mu)}{z}$$

where $\omega_\mu^{t,z}$ transforms a family $\omega_\mu = ((t_1,z_1),(t_2,z_2),\cdots) \in \Omega_\mu$ into a new family $\omega_\mu^{t,z}((t,z),(t_1,z_1),(t_2,z_2),\cdots) \in \Omega_\mu$. This is defined for a one-dimensional Poisson random measure. In the multi-dimensional case, $\mathcal{I}_{t,z}H$ is extended to a k-dimensional vector in the obvious way. It is known that when $\mathbb{E}\Big[\int_0^T\int_E |\mathcal{I}_{t,z}H|^2 z^2\nu(dz)dt\Big] = \mathbb{E}\Big[\sum_{i=1}^k \int_0^T\int_{\mathbb{R}_0} |\mathcal{I}_{t,z_i}H|^2 z_i^2\nu^i(dz_i)dt\Big] < \infty$, one has $D_{t,z}H = \mathcal{I}_{t,z}H$.

Proposition 3.3. *Under Assumption 3.1, the process $X^{t,x,\epsilon}$ is Malliavin differentiable. Moreover, it satisfies*

$$\sup_{(s,z)\in[0,T]\times\mathbb{R}^k} \mathbb{E}\Big[\sup_{r\in[s,T]} |D_{s,z}X_r^{t,x,\epsilon}|^p\Big] < \infty$$

for $\forall p \geq 2$.

Proof. This is a modification of Theorem 4.1.2 in [12] for our setting. The existence of Malliavin derivative follows from Theorem 3 in Petrou (2008) [45].

According to [45], for $z^i \neq 0$, one has

$$\begin{aligned} D_{s,z^i} X_r^{t,x,\epsilon} &= \frac{\gamma^i(s, X_{s-}^{t,x,\epsilon}, z^i, \epsilon)}{z^i} + \int_s^r D_{s,z^i} b(u, X_u^{t,x,\epsilon}, \epsilon) du \\ &+ \int_s^r D_{s,z^i} \sigma(u, X_u^{t,x,\epsilon}, \epsilon) dW_u + \int_s^r \int_E D_{s,z^i} \gamma(u, X_{u-}^{t,x,\epsilon}, z, \epsilon) \widetilde{\mu}(du, dz) \end{aligned} \tag{3.5}$$

for $s \leq r$ and $D_{s,z^i} X_r^{t,x,\epsilon} = 0$ otherwise. Here, γ^i denotes the i-th column vector and

$$D_{s,z^i} b(u, X_u^{t,x,\epsilon}, \epsilon) := \frac{1}{z^i} \left[b(u, X_u^{t,x,\epsilon} + z^i D_{s,z^i} X_u^{t,x,\epsilon}, \epsilon) - b(u, X_u^{t,x,\epsilon}, \epsilon) \right]$$

and similarly for the terms $(D_{s,z^i} \sigma(u, X_u^{t,x,\epsilon}, \epsilon), D_{s,z^i} \gamma(u, X_{u-}^{t,x,\epsilon}, z, \epsilon))$. Due to the uniformly bounded derivative of $\partial_x b, \partial_x \sigma, \partial_x \gamma/\eta$, (3.5) has the unique solution by Lemma A.3. In addition, applying the Burkholder-Davis-Gundy (BDG), Gronwall inequalities and Lemma A.1, one obtains

$$\mathbb{E}||D_{s,z^i} X^{t,x,\epsilon}||_{[s,T]}^p \leq C_p \left(\left| \frac{\gamma^i(s, 0, z^i, \epsilon)}{z^i} \right|^p + \mathbb{E}||X^{t,x,\epsilon}||_T^p \right) .$$

Thus, by Assumption 3.1 (iii), we obtain the desired result. The arguments for the Wiener direction ($z = 0$) are similar. □

Next theorem is an adaptation of Theorem 3.5.1 in [12] and Theorem C.1 in [25]. We suppress the superscripts (t, x, ϵ) denoting the initial data for simplicity.

Theorem 3.1. *Under Assumptions 3.1 and 3.2,*
(a) There exists a unique solution $(Y^{s,0}, Z^{s,0}, \psi^{s,0})$ *belongs to* $\mathcal{K}^p$ $\forall p \geq 2$ *to the BSDE*

$$Y_u^{s,0} = D_{s,0}\xi(X_T) + \int_u^T f^{s,0}(r) dr - \int_u^T Z_r^{s,0} dW_r - \int_u^T \int_E \psi_r^{s,0}(z) \widetilde{\mu}(dr, dz)$$

where

$$\begin{aligned} D_{s,0}\xi(X_T) &:= \partial_x \xi(X_T) D_{s,0} X_T \\ f^{s,0}(r) &:= \partial_x f(r, \Theta_r) D_{s,0} X_r + \partial_y f(r, \Theta_r) Y_r^{s,0} + \partial_z f(r, \Theta_r) Z_r^{s,0} \\ &+ \partial_u f(r, \Theta_r) \int_{\mathbb{R}_0} \rho(z) \psi_r^{s,0}(z) \nu(dz). \end{aligned}$$

(b) For $z^i \neq 0$, *there exists a unique solution* $(Y^{s,z^i}, Z^{s,z^i}, \psi^{s,z^i})$ *belongs to* $\mathcal{K}^p$ $\forall p \geq 2$ *to the BSDE*

$$Y_u^{s,z^i} = D_{s,z^i}\xi(X_T) + \int_u^T f^{s,z^i}(r) dr - \int_u^T Z_r^{s,z^i} dW_r - \int_u^T \int_E \psi_r^{s,z^i}(z) \widetilde{\mu}(dz, dr)$$

where

$$D_{s,z^i}\xi(X_T) := \frac{\xi(X_T + z^i D_{s,z^i}X_T) - \xi(X_T)}{z^i}$$

$$f^{s,z^i}(r) := \Big[f\Big(r, X_r + z^i D_{s,z^i}X_r, Y_r + z^i D_{s,z^i}Y_r, Z_r + z^i D_{s,z^i}Z_r, \int_{\mathbb{R}_0}\rho(e)\big[\psi_r(e) + z^i D_{s,z^i}\psi_r(e)\big]\nu(de)\Big) - f\Big(r, X_r, Y_r, Z_r, \int_{\mathbb{R}_0}\rho(e)\psi_r(e)\nu(de)\Big)\Big]/z^i$$

for every $1 \le i \le k$.
(c)For $u < s \le T$, *set* $(Y_u^{s,z}, Z_u^{s,z}, \psi_u^{s,z}) = 0$ *for* $z \in \mathbb{R}^k$ *(i.e., including Wiener direction* $z = 0$*). Then,* (Y, Z, ψ) *is Malliavin differentiable and* $(Y^{s,z}, Z^{s,z}, \psi^{s,z})$ *is a version of* $(D_{s,z}Y, D_{s,z}Z, D_{s,z}\psi)$.
(d)Set a deterministic function $u(t, x, \epsilon) := Y_t^{t,x,\epsilon}$ *using the solution of the BSDE (3.2). If* u *is continuous in* t *and one-time continuously differentiable with respect to* x*, then*

$$Z_s^{t,x,\epsilon} = \partial_x u(s, X_{s-}^{t,x,\epsilon}, \epsilon)\sigma(s, X_{s-}^{t,x,\epsilon}, \epsilon) \tag{3.6}$$

$$\Big(\psi_s^{t,x,\epsilon}(z)\Big)^i_{1\le i\le k} = \Big(u\big(s, X_{s-}^{t,x,\epsilon} + \gamma^i(s, X_{s-}^{t,x,\epsilon}, z^i, \epsilon), \epsilon\big) - u(s, X_{s-}^{t,x,\epsilon}, \epsilon)\Big)_{1\le i\le k} \tag{3.7}$$

for $t \le s \le T$ *and* $z = (z^i)_{1\le i\le k} \in \mathbb{R}^k$.

Proof. (a) and (b) can be proved by Lemma B.2, the boundedness of derivatives and the fact that $\Theta^{t,x,\epsilon} \in \mathbb{S}^p \times \mathcal{K}^p$ and $D_{s,z}X \in \mathbb{S}^p$ for $\forall p \ge 2$.
(c) can be proved as a simple modification of Theorem 3.5.1 in [12], which is an extension of Proposition 5.3 in El Karoui et.al (1997) [18] to the jump case. The conditions written for ω-dependent driver (assumptions (vii) and (viii) of [12]) can be replaced by our assumption on f, which is Lipschitz with respect to (y, z, u) and has a polynomial growth in x. Note that we already know $X^{t,x,\epsilon}, D_{s,z}X^{t,x,\epsilon} \in \mathbb{S}^p\ \forall p \ge 2$. See also the arguments used in proof of Theorem 6.1 in [25] for a Markovian setup. (d) follows from Theorem 4.1.4 of [12]. □

4 Asymptotic Expansion

As the asymptotic expansion scheme, we want to obtain the Taylor expansion of the solution $(X^{t,x,\epsilon}, Y^{t,x,\epsilon}, Z^{t,x,\epsilon}, \psi^{t,x,\epsilon})$ of the FBSDEs (3.1) and (3.2) around $\epsilon = 0$. It is well-known that this is possible for the forward process $X^{t,x,\epsilon}$. For the backward components $\hat{\Theta}^{t,x,\epsilon}$, we need to prove the existence of classical derivative $\partial_\epsilon^n \hat{\Theta}^{t,x,\epsilon}$ for every $0 \le n \le n_{\max} + 1$ and then to obtain its estimate in an appropriate norm. Since the BSDE corresponding to the classical derivative $\partial_\epsilon^n \hat{\Theta}^{t,x,\epsilon}$ contains the terms proportional to $\prod_{i=1}^j \partial_\epsilon^{k_i} \hat{\Theta}^{t,x,\epsilon}$ with $\sum_{i=1}^j k_i = n$ in its driver, the estimates of $(\partial_\epsilon^{k_i} Z^{t,x,\epsilon}, \partial_\epsilon^{k_i}\psi^{t,x,\epsilon})_{i=1}^j$ with respect to the norm $\mathbb{H}^p \times \mathbb{H}^p_\nu\ \forall p \ge 2$ are not enough to guarantee the well-posedness of the relevant BSDE.

In the following, we shall solve this issue by showing $(\partial_\epsilon^{k_i} Z^{t,x,\epsilon}, \partial_\epsilon^{k_i}\psi^{t,x,\epsilon})$ actually belongs to $\mathbb{S}^p \times \mathbb{S}^p\ \forall p \ge 2$ instead of $\mathbb{H}^p \times \mathbb{H}^p_\nu\ \forall p \ge 2$. This is done by recursively applying the representation theorem and the polynomial growth property of the solutions with respect to x. In order to use the result in Theorem 3.1 (d), we have to start from studying the classical derivatives of the BSDE (3.2) with respect to x.

4.1 Classical derivatives of BSDEs

Lemma 4.1. *Under Assumptions 3.1 and 3.2, $\hat{\Theta}^{t,x,\epsilon}$ is classically differentiable in x, and it is given by $\partial_x\hat{\Theta}^{t,x,\epsilon}$ defined as the unique solution of the BSDE with formal differentiation:*

$$\partial_x Y_s^{t,x,\epsilon} = \partial_x\xi(X_T^{t,x,\epsilon})\partial_x X_T^{t,x,\epsilon} + \int_s^T \partial_\Theta f(r,\Theta_r^{t,x,\epsilon})\partial_x\Theta_r^{t,x,\epsilon}dr \\ -\int_s^T \partial_x Z_r^{t,x,\epsilon}dW_r - \int_s^T\int_E \partial_x\psi_r^{t,x,\epsilon}(z)\widetilde{\mu}(dr,dz) \tag{4.1}$$

and $\partial_x\hat{\Theta}^{t,x,\epsilon}\in\mathcal{K}^p[t,T]$ satisfying

$$||\partial_x\hat{\Theta}^{t,x,\epsilon}||_{\mathcal{K}^p[t,T]} \le C_p(1+|x|^q)$$

for $\forall p\ge 2$.

Proof. The existence and uniqueness can be easily shown by Lemma B.2. Note that the BSDE (4.1) is linear with bounded Lipschitz constants and satisfies

$$||\partial_x\hat{\Theta}^{t,x,\epsilon}||^p_{\mathcal{K}^p[t,T]} \le C_p\mathbb{E}\Big[|\partial_x\xi(X_T^{t,x,\epsilon})|^p|\partial_x X_T^{t,x,\epsilon}|^p + \Big(\int_t^T |\partial_x f(r,\Theta_r^{t,x,\epsilon})||\partial_x X_r^{t,x,\epsilon}|dr\Big)^p\Big] \\ \le C_p||\partial_x X^{t,x,\epsilon}||^p_{\mathbb{S}^{2p}[t,T]}\Big\{\Big(\mathbb{E}|\partial_x\xi(X_T^{t,x,\epsilon})|^{2p}\Big)^{1/2} + \Big(\mathbb{E}\Big(\int_t^T|\partial_x f(r,X_r^{t,x,\epsilon},0)|dr\Big)^{2p}\Big)^{1/2} \\ +||\hat{\Theta}^{t,x,\epsilon}||^p_{\mathcal{K}^{2p}[t,T]}\Big\} \le C_p(1+|x|^{pq})$$

for $\forall p\ge 2$. With a simple modification of Theorem 3.1 of [39], one can also show that

$$\lim_{h\to 0}||\nabla^h\hat{\Theta}^{t,x,\epsilon} - \partial_x\hat{\Theta}^{t,x,\epsilon}||^2_{\mathcal{K}^2[t,T]} = 0$$

where $\nabla^h\hat{\Theta}^{t,x,\epsilon} := \dfrac{\hat{\Theta}^{t,x+h,\epsilon}-\hat{\Theta}^{t,x,\epsilon}}{h}$ with $h\neq 0$ (for each direction). This gives the agreement with the classical differentiation. □

Corollary 4.1. *Under Assumptions 3.1 and 3.2, there exists $\partial_x u(t,x,\epsilon)$ which is continuous in (t,x) and has at most a polynomial growth in x uniformly in $(t,\epsilon)\in[0,T]\times[0,1]$. Furthermore, $Z^{t,x,\epsilon}$ and $\int_{\mathbb{R}_0}\rho(z)\psi^{t,x,\epsilon}(z)\nu(dz)$ belong to $\mathbb{S}^p[t,T]$ $\forall p\ge 2$.*

Proof. Note that $\partial_x u(t,x,\epsilon) = \partial_x Y_t^{t,x,\epsilon}$ and there exists some constant $C>0$ such that

$$|\partial_x u(t,x,\epsilon)| \le ||\partial_x\hat{\Theta}^{t,x,\epsilon}||_{\mathcal{K}^p[t,T]} \le C(1+|x|^q)$$

for every $x\in\mathbb{R}^d$ uniformly in $t\in[0,T]$ by Lemma 4.1. The continuity of $\partial_x u(t,x,\epsilon)$ in (t,x) can be shown in the same way as [39] using the continuity of $X^{t,x,\epsilon}$ in (t,x), which can be seen in Lemma A.3. Then, from the representation given in (3.6), (3.7) and the above result, one sees

$$|Z_s^{t,x,\epsilon}| + \Big|\int_E \rho(z)\psi_s^{t,x,\epsilon}(z)\nu(dz)\Big| \le C(1+|X_{s-}^{t,x,\epsilon}|^{q+1})$$

which gives the desired result $\hat{\Theta}^{t,x,\epsilon}\in\mathbb{S}^p[t,T]^{\otimes 3}$ for any $p\ge 2$. □

Proposition 4.1. *Under Assumptions 3.1 and 3.2, the classical derivative* $\partial_x^n \hat{\Theta}^{t,x,\epsilon}$ *exists for every* $0 \leq n \leq n^{\mathrm{ae}}$ *with* $\partial_x^n \hat{\Theta}^{t,x,\epsilon} \in \mathcal{K}^p[t,T]$ $\forall p \geq 2$ *and is given by the solution of the following BSDE:*

$$\begin{aligned}\partial_x^n Y_s^{t,x,\epsilon} &= \xi_n + \int_s^T \Big\{ H_{n,r} + \partial_\Theta f(r, \Theta_r^{t,x,\epsilon}) \partial_x^n \Theta_r^{t,x,\epsilon} \Big\} dr \\ &\quad - \int_s^T \partial_x^n Z_r^{t,x,\epsilon} dW_r - \int_s^T \int_E \partial_x^n \psi_r^{t,x,\epsilon}(z) \widetilde{\mu}(dr, dz) \end{aligned} \tag{4.2}$$

where

$$\xi_n := n! \sum_{k=1}^{n} \sum_{\beta_1 + \cdots + \beta_k = n, \beta_i \geq 1} \frac{1}{k!} \partial_x^k \xi(X_T^{t,x,\epsilon}) \prod_{j=1}^{k} \frac{1}{\beta_j!} \partial_x^{\beta_j} X_T^{t,x,\epsilon},$$

$$\begin{aligned} H_{n,r} &:= n! \sum_{k=2}^{n} \sum_{\beta_1+\cdots+\beta_k = n, \beta_i \geq 1} \sum_{i_x=0}^{k} \sum_{i_y=0}^{k-i_x} \sum_{i_z=0}^{k-i_x-i_y} \frac{\partial_x^{i_x} \partial_y^{i_y} \partial_z^{i_z} \partial_u^{k-i_x-i_y-i_z} f(r, \Theta_r^{t,x,\epsilon})}{i_x! i_y! i_z! (k - i_x - i_y - i_z)!} \\ &\times \prod_{j_x=1}^{i_x} \frac{1}{\beta_{j_x}!} \partial_x^{\beta_{j_x}} X_r^{t,x,\epsilon} \prod_{j_y=i_x+1}^{i_x+i_y} \frac{1}{\beta_{j_y}!} \partial_x^{\beta_{j_y}} Y_r^{t,x,\epsilon} \prod_{j_z=i_x+i_y+1}^{i_x+i_y+i_z} \frac{1}{\beta_{j_z}!} \partial_x^{\beta_{j_z}} Z_r^{t,x,\epsilon} \\ &\times \prod_{j_u=i_x+i_y+i_z+1}^{k} \frac{1}{\beta_{j_u}!} \int_{\mathbb{R}_0} \rho(z) \partial_x^{\beta_{j_u}} \psi_r^{t,x,\epsilon}(z) \nu(dz) \ . \end{aligned}$$

Moreover, for every $0 \leq n \leq n_{\max} + 1$, $\partial_x^n \hat{\Theta}^{t,x,\epsilon} \in \mathbb{S}^p[t,T]^{\otimes 3}$ $\forall p \geq 2$.

Proof. We can prove recursively by the arguments used in Proposition 3.2, Lemma 4.1 and Corollary 4.1. We already know that $\hat{\Theta}^{t,x,\epsilon} \in \mathbb{S}^p[t,T]^{\otimes 3}$ and $\partial_x \hat{\Theta}^{t,x,\epsilon} \in \mathcal{K}^p[t,T]$ for any $p \geq 2$. The BSDE for $\partial_x^2 \hat{\Theta}^{t,x,\epsilon}$ has bounded Lipschitz constants and $H_{2,r}$ is at most quadratic in $(\partial_x \hat{\Theta}_r^{t,x,\epsilon})$. From the fact that $\xi(x), f(\cdot, x, 0)$ have at most a polynomial growth in x and that $(\partial_x^m X^{t,x,\epsilon})_{0 \leq m \leq n^{\mathrm{ae}}}, \hat{\Theta}^{t,x,\epsilon} \in \mathbb{S}^p[t,T]$ $\forall p \geq 2$, one can prove the existence of the unique solution $\partial_x^2 \hat{\Theta}^{t,x,\epsilon} \in \mathcal{K}^p[t,T]$ $\forall p \geq 2$ by Lemma B.2. Furthermore, one can show as in Lemma 4.1 that $||\partial_x^2 \hat{\Theta}^{t,x,\epsilon}||_{\mathcal{K}^p[t,T]}$ has at most polynomial growth in x. By following the arguments of Theorem 3.1 of [39], one sees this agrees with the classical differentiation in the sense of Lemma 4.1. This in turn shows the existence $\partial_x^2 u(t,x,\epsilon) = \partial_x^2 Y_t^{t,x,\epsilon}$ and also the fact that $\partial_x^2 u(t,x,\epsilon)$ has at most a polynomial growth in x. This implies that, together with Assumption 3.1 and the representation theorem (3.6) (3.7), $\partial_x Z^{t,x,\epsilon}$ and $\int_{\mathbb{R}_0} \rho(z) \partial_x \psi^{t,x,\epsilon}(z) \nu(dz)$ are in $\mathbb{S}^p[t,T]$ $\forall p \geq 2$. Thus, we get $\partial_x \hat{\Theta}^{t,x,\epsilon} \in \mathbb{S}^p[t,T]^{\otimes 3}$ $\forall p \geq 2$.

In the same manner, if we assume that $\left(\partial_x^i \hat{\Theta}^{t,x,\epsilon}\right)_{i \leq n} \in \mathbb{S}^p[t,T]^{\otimes 3}$ and that $\partial_x^{n+1} \hat{\Theta}^{t,x,\epsilon} \in \mathcal{K}^p[t,T]$ for $\forall p \geq 2$ with the $\mathcal{K}^p$-norm at most a polynomial growth in x , then one can show that the existence of the unique solution $\partial_x^{n+2} \hat{\Theta}^{t,x,\epsilon} \in \mathcal{K}^p[t,T]$ with the norm at most a polynomial growth in x by Lemma B.2. It then implies from the representation theorem that $\partial_x^{n+1} \hat{\Theta}^{t,x,\epsilon} \in \mathbb{S}^p[t,T]^{\otimes 3}$ $\forall p \geq 2$. By repeating the procedures, one obtains the desired result. □

4.2 Asymptotic expansion

We are now going to prove $\partial_\epsilon^n \hat{\Theta}^{t,x,\epsilon} \in \mathbb{S}^p[t,T]^{\otimes 3}$ $\forall p \geq 2$ for every $0 \leq n \leq n_{\max} + 1$. Although the strategy is similar to the previous section, we actually have to study the properties of $\left(\partial_x^m \partial_\epsilon^n \hat{\Theta}^{t,x,\epsilon}\right)$ since ϵ affects $u(s, X_{s-}^{t,x,\epsilon}, \epsilon)$ not only through its explicit dependence but also through $X^{t,x,\epsilon}$ indirectly.

Lemma 4.2. *Under Assumptions 3.1 and 3.2, $\hat{\Theta}^{t,x,\epsilon}$ is classically differentiable in ϵ, and it is given by $\partial_\epsilon \hat{\Theta}^{t,x,\epsilon}$ defined as the unique solution of the BSDE with formal differentiation:*

$$\partial_\epsilon Y_s^{t,x,\epsilon} = \partial_x \xi(X_T^{t,x,\epsilon}) \partial_\epsilon X_T^{t,x,\epsilon} + \int_s^T \partial_\Theta f(r, \Theta_r^{t,x,\epsilon}) \partial_\epsilon \Theta_r^{t,x,\epsilon} dr$$
$$- \int_s^T \partial_\epsilon Z_r^{t,x,\epsilon} dW_r - \int_s^T \int_E \partial_\epsilon \psi_r^{t,x,\epsilon} \widetilde{\mu}(dr, dz) \ .$$

One has $\partial_\epsilon \hat{\Theta}^{t,x,\epsilon} \in \mathcal{K}^p[t,T]$ satisfying

$$||\partial_\epsilon \hat{\Theta}^{t,x,\epsilon}||_{\mathcal{K}^p[t,T]} \leq C_p(1 + |x|^q)$$

for any $\forall p \geq 2$.

Proof. The proof can be done similarly as in Lemma 4.1. □

We now get the following result.

Proposition 4.2. *Under Assumptions 3.1 and 3.2, the classical derivative $\partial_\epsilon^n \hat{\Theta}^{t,x,\epsilon}$ exists for every $0 \leq n \leq n^{\mathrm{ae}}$ with $\partial_\epsilon^n \hat{\Theta}^{t,x,\epsilon} \in \mathcal{K}^p[t,T]$ $\forall p \geq 2$ and is given by the unique solution of the following BSDE:*

$$\partial_\epsilon^n Y_s^{t,x,\epsilon} = \widetilde{\xi}_n + \int_s^T \left\{ \widetilde{H}_{n,r} + \partial_\Theta f(r, \Theta_r^{t,x,\epsilon}) \partial_\epsilon^n \Theta_r^{t,x,\epsilon} \right\} dr$$
$$- \int_s^T \partial_\epsilon^n Z_r^{t,x,\epsilon} dW_r - \int_s^T \int_E \partial_\epsilon^n \psi_r^{t,x,\epsilon} \widetilde{\mu}(dr, dz) \ .$$

Here, $\widetilde{\xi}_n$ and $\widetilde{H}_{n,r}$ are given by the expressions of ξ_n and $H_{n,r}$ in Proposition 4.1 with $\partial_x^{\beta_{j\Theta}}$ replaced by $\partial_\epsilon^{\beta_{j\Theta}}$. Moreover, for every $0 \leq n \leq n_{\max} + 1$, $\partial_\epsilon^n \hat{\Theta}^{t,x,\epsilon} \in \mathbb{S}^p[t,T]^{\otimes 3}$ $\forall p \geq 2$.

Proof. We start from the result of Lemma 4.2, which implies $\partial_\epsilon u(t,x,\epsilon)$ has at most polynomial growth in x. Using the fact that $\partial_\epsilon \Theta^{t,x,\epsilon} \in \mathbb{S}^p[t,T] \times \mathcal{K}^p[t,T]$ and $\partial_x \Theta^{t,x,\epsilon} \in \mathbb{S}^p[t,T]^{\otimes 4}$, one can show that $\partial_x \partial_\epsilon \hat{\Theta}^{t,x,\epsilon}$ exists and satisfies $\partial_x \partial_\epsilon \hat{\Theta}^{t,x,\epsilon} \in \mathcal{K}^p[t,T]$ for $\forall p \geq 2$ as in Lemma 4.1. The corresponding norm has at most polynomial growth in x and so is $\partial_x \partial_\epsilon u(t,x,\epsilon)$. This implies, together with the representations (3.6) and (3.7), that $\partial_\epsilon \hat{\Theta}^{t,x,\epsilon} \in \mathbb{S}^p[t,T]$ for $\forall p \geq 2$.

As in Proposition 4.1, one can recursively prove that the classical derivative $\partial_x^n \partial_\epsilon \hat{\Theta}^{t,x,\epsilon}$ exists and belongs to $\mathcal{K}^p[t,T]$ $\forall p \geq 2$ for every $0 \leq n \leq n^{\mathrm{ae}}$ and moreover that it belongs to $\mathbb{S}^p[t,T]^{\otimes 3}$ $\forall p \geq 2$ for every $0 \leq n \leq n_{\max} + 1$ by induction. Then, by Lemma B.2, it is straightforward to check $\partial_x^n \partial_\epsilon^2 \hat{\Theta}^{t,x,\epsilon}$ exists and belongs to $\mathcal{K}^p[t,T]$ $\forall p \geq 2$ for $0 \leq n \leq n^{\mathrm{ae}}$. By the representation theorem, it then implies $\partial_x^n \partial_\epsilon^2 \hat{\Theta}^{t,x,\epsilon}$ in fact belongs to $\in \mathbb{S}^p[t,T]$ $\forall p \geq 2$ for $0 \leq n \leq n_{\max} + 1$. By repeating the same procedures, one can show that, for every $0 \leq n, m \leq n_{\max} + 1$, $\partial_x^n \partial_\epsilon^m \hat{\Theta}^{t,x,\epsilon}$ exists and belongs to $\mathbb{S}^p[t,T]^{\otimes 3}$ $\forall p \geq 2$. Thus the claims of the proposition are proved. □

We have shown that $\Theta^{t,x,\epsilon}$ is n^{ae}-time classically differentiable with respect to (x, ϵ) and, in particular for $n \leq n_{\max} + 1$, $\partial_\epsilon^n \Theta^{t,x,\epsilon} \in \mathbb{S}^p[t,T]^{\otimes 4}$ $\forall p \geq 2$. Let us define for $s \in [t,T]$

and $0 \le n \le n_{\max}$ that

$$\Theta_s^{[n]} := \frac{1}{n!}\partial_\epsilon^n \Theta_s^{t,x,\epsilon}\Big|_{\epsilon=0}.$$

Using the differentiability and the Taylor formula, one has for any $1 \le N \le n_{\max}$

$$\Theta_s^{t,x,\epsilon} = \Theta_s^{[0]} + \sum_{n=1}^{N} \epsilon^n \Theta_s^{[n]} + \frac{\epsilon^{N+1}}{N!}\int_0^1 (1-u)^N \left(\partial_\alpha^{N+1}\Theta_s^{t,x,\alpha}\right)\Big|_{\alpha=u\epsilon} du \ . \tag{4.3}$$

As we shall see later, each $\Theta^{[m]}, m \in \{1, 2, \cdots, n_{\max}\}$ can be evaluated by solving the system of linear ODEs. Although $\Theta^{[0]}$ requires to solve a non-linear ODE as an exception, the existence of the bounded solution is guaranteed under the Assumptions 3.1 and 3.2.

The next theorem is the main result of the paper which gives the error estimate for the approximation of $\Theta^{t,x,\epsilon}$ by the series of $\Theta^{[m]}, m \in \{0, 1, \cdots, n_{\max}\}$.

Theorem 4.1. *Under Assumptions 3.1 and 3.2, the asymptotic expansion of the forward-backward SDEs (3.1) and (3.2) is given by (4.3) for every* $1 \le N \le n_{\max}$ *and satisfies, with some positive constant* C_p*, that*

$$\left|\left|\Theta^{t,x,\epsilon} - \left(\Theta^{[0]} + \sum_{n=1}^{N}\epsilon^n \Theta^{[n]}\right)\right|\right|_{\mathbb{S}^p[t,T]} \le \epsilon^{N+1} C_p \ . \tag{4.4}$$

Proof. This immediately follows from the fact that $\partial_\epsilon^{N+1}\Theta^{t,x,\epsilon}$ is in $\mathbb{S}^p[t,T]$ $\forall p \ge 2$ and continuous with respect to ϵ by Propositions 3.1 and 4.2. □

4.3 State-dependent jump intensity

When ν is a finite measure $\nu(E) < \infty$, all the previous results hold true with slightly weaker assumptions with $\eta, \rho \equiv 1$ in Assumptions 3.1 and 3.2. In practical applications, however, there are many cases where we want to make the jump intensity state dependent. In this section, we solve this problem when the intensity is bounded.

In particular, we consider the forward-backward SDEs (3.1) and (3.2) but with the compensated random measure $\widetilde{\mu}(dr, dz)$ given by, for $1 \le i \le k$,

$$\widetilde{\mu}^i(dr, dz) = \mu^i(dr, dz) - \lambda^i(r, X_r^{t,x,\epsilon})\nu^i(dz)dr$$

where ν^i is normalized as $\nu^i(\mathbb{R}_0) = 1$ and $\lambda^i : [0,T]\times\mathbb{R}^d \to \mathbb{R}$. One can see that the random measure is not Poissonian any more and depends implicitly on ϵ through its intensity.

Assumption 4.1. *For every* $1 \le i \le k$, $\nu^i(\mathbb{R}_0) = 1$ *and there exist some positive constants* K, c_1, c_2 *such that*
(i) $\lambda^i(t,x)$ *is continuous in* (t,x), $n^{\rm ae}$*-time differentiable in* x *with continuous derivatives satisfying* $|\partial_x^n \lambda^i(t,x)| \le K$ *for every* $1 \le n \le n^{\rm ae}$ *uniformly in* $(t,x) \in [0,T]\times\mathbb{R}^d$,
(ii) $0 < c_1 \le \lambda^i(t,x) \le c_2$ *uniformly in* $(t,x) \in [0,T]\times\mathbb{R}^d$,
(iii) $|\partial_\epsilon^m \gamma_{\cdot,i}(t,x,z,\epsilon)| \le K$ *for every* $1 \le m \le n^{\rm ae}$ *uniformly in* $(t,x,z,\epsilon) \in [0,T]\times\mathbb{R}^d\times\mathbb{R}_0\times[0,1]$.

Lemma 4.3. *Under Assumption 4.1, one can define an equivalent probability measure* $\mathbb{Q}$ *by, for* $s \in [t,T]$,

$$\frac{d\mathbb{Q}}{d\mathbb{P}}\Big|_{\mathcal{F}_s} = M_s$$

where M is a strictly positive $\mathbb{P}$-martingale given by

$$M_s = 1 + \sum_{i=1}^{k} \int_t^s M_{r-}\Big(\frac{c_2}{\lambda^i(r, X_{r-}^{t,x,\epsilon})} - 1\Big)\widetilde{\mu}^i(dr, \mathbb{R}_0)\ .$$

Under the new measure $\mathbb{Q}$, the compensated random measure becomes

$$\widetilde{\mu}^{\mathbb{Q}}(dr, dz) = \mu(dr, dz) - c_2\nu(dz)dt$$

and hence μ is Poissonian. Moreover, for $\forall s \in [t, T]$,

$$M_s \geq \exp\big(-(c_2 - c_1)k(T - t)\big)\ .$$

Proof. By Kazamaki (1979) [32], it is known that if X is a BMO martingale satisfying $\Delta X_t \geq -1 + \delta$ a.s. for all $t \in [0, T]$ with some strictly positive constant $\delta > 0$, then Doléans-Dade exponential $\mathcal{E}(X)$ is uniformly integrable. One can easily confirm that this condition is satisfied for a martingale

$$\int^{\cdot} \Big(c_2/\lambda(s, X_s^{t,x,\epsilon}) - 1\Big)\widetilde{\mu}(ds, \mathbb{R}_0)\ .$$

Thus the given measure change is well-defined and the first claim follows from Theorem 41 in Chapter 3 of [47]. The second claim directly follows from the explicit expression

$$\begin{aligned} M_s &= \prod_{i=1}^{k} \left\{ \prod_{0<r\leq s} \Big(\frac{c_2}{\lambda^i(r, X_{r-}^{t,x,\epsilon})}\Big)^{\Delta\mu^i(r,\mathbb{R}_0)} \exp\Big(-\int_t^s (c_2 - \lambda^i(r, X_{r-}^{t,x,\epsilon}))dr\Big) \right\} \\ &\geq \exp\Big(-\int_t^s k(c_2 - c_1)dr\Big)\ . \end{aligned}$$

□

Under the measure $\mathbb{Q}$, we have

$$\begin{aligned} X_s^{t,x,\epsilon} &= x + \int_t^s \widetilde{b}(r, X_r^{t,x,\epsilon}, \epsilon)dr + \int_t^s \sigma(r, X_r^{t,x,\epsilon}, \epsilon)dW_r \\ &\quad + \int_t^s \int_E \gamma(r, X_{r-}^{t,x,\epsilon}, z, \epsilon)\widetilde{\mu}^{\mathbb{Q}}(dr, dz) \end{aligned} \tag{4.5}$$

$$\begin{aligned} Y_s^{t,x,\epsilon} &= \xi(X_T^{t,x,\epsilon}) + \int_s^T \widetilde{f}\Big(r, X_r^{t,x,\epsilon}, Y_r^{t,x,\epsilon}, Z_r^{t,x,\epsilon}, \int_{\mathbb{R}_0} \psi_r^{t,x,\epsilon}(z)\nu(dz)\Big)dr \\ &\quad - \int_s^T Z_r^{t,x,\epsilon}dW_r - \int_s^T \int_E \psi_r^{t,x,\epsilon}(z)\widetilde{\mu}^{\mathbb{Q}}(dr, dz) \end{aligned} \tag{4.6}$$

where

$$\widetilde{b}(s, x, \epsilon) = b(s, x, \epsilon) + \sum_{i=1}^{k}(c_2 - \lambda^i(s, x)) \int_{\mathbb{R}_0} \gamma^i(s, x, z^i, \epsilon)\nu(dz^i)$$

$$\widetilde{f}(s, x, y, z, u) = f(s, x, y, z, u) - \sum_{i=1}^{k}(c_2 - \lambda^i(s, x))u^i.$$

Theorem 4.2. *Under Assumptions 3.1, 3.2 with ρ and η replaced by 1, and Assumption 4.1, the solution $\Theta^{t,x,\epsilon}$ of the forward-backward SDEs (3.1) and (3.2) allows the asymptotic expansion with respect to ϵ and satisfies the same error estimate (4.4) in the original measure $\mathbb{P}$.*

Proof. Assumption 4.1 makes $(\widetilde{b}, \widetilde{f})$ once again satisfy Assumptions 3.1 and 3.2 with ρ, η replaced by 1. Therefore, all the results in the previous sections hold true under the measure $\mathbb{Q}$ to the equivalent FBSDEs (4.5) and (4.6). In particular this implies from Lemma 4.3 that, with some positive constant C_p,

$$\begin{aligned}
\epsilon^{p(N+1)}C_p &\geq \mathbb{E}^{\mathbb{Q}}\left[\sup_{s\in[t,T]}\left|\Theta_s^{t,x,\epsilon} - \Big(\Theta_s^{[0]} + \sum_{n=1}^{N}\epsilon^n\Theta_s^{[n]}\Big)\right|^p\right] \\
&= \mathbb{E}\left[M_T \sup_{s\in[t,T]}\left|\Theta_s^{t,x,\epsilon} - \Big(\Theta_s^{[0]} + \sum_{n=1}^{N}\epsilon^n\Theta_s^{[n]}\Big)\right|^p\right] \\
&\geq \exp\big(-k(c_2-c_1)(T-t)\big)\mathbb{E}\left[\sup_{s\in[t,T]}\left|\Theta_s^{t,x,\epsilon} - \Big(\Theta_s^{[0]} + \sum_{n=1}^{N}\epsilon^n\Theta_s^{[n]}\Big)\right|^p\right] .
\end{aligned}$$

This proves the claim. □

5 Implementation of the asymptotic expansion

5.1 Evaluation scheme

In this section, we explain how to calculate $\Theta^{[n]}$, $n \in \{0, 1, \cdots, n_{\max}\}$ (semi)-analytically. As we shall see, if we introduce ϵ in a specific way to the forward SDE (3.1), then the grading structure introduced by the asymptotic expansion yields a very simple scheme requiring only a system of linear ODEs to be solved with only one exception at the zero-th order. It is also remarkable that one can directly approximate not only $(Y^{t,x}, Z^{t,x})$ but also the $\mathbb{L}^2(E;\nu)$-valued process $\psi^{t,x}(\cdot)$. This looks almost infeasible for the standard regression-based simulation scheme.

Let us put the initial time as $t = 0$, and take $(m = d = l = 1)$ for *notational simplicity.* The extension to higher dimensional setups is straightforward for which one only needs proper indexing of each variable. Let us adopt a following parametrization of X with ϵ which obviously leads to *small-variance* expansion;

$$X_s^\epsilon = x + \int_0^s b(r, X_r^\epsilon, \epsilon)dr + \int_0^s \epsilon\sigma(r, X_r^\epsilon)dW_r + \int_0^s \int_{\mathbb{R}_0} \epsilon\gamma(r, X_{r-}^\epsilon, z)\widetilde{\mu}(dr, dz) ,$$

where we omit the superscript denoting the initial data $(0, x)$. One can see that the process X^ϵ becomes deterministic when $\epsilon \to 0$. Similar to the standard applications [50], this parameterization is crucial to obtain semi-analytic approximations. We make Assumptions 3.1 and 3.2 (or those replaced by $\rho = \eta = 1$ and Assumption 4.1) the standing assumptions for this section.

Lemma 5.1. *The zero-th order solution* $\left(\Theta_s^{[0]}, s \in [0,T]\right)$ *is given by*

$$
\begin{aligned}
X_s^{[0]} &= x + \int_0^s b(r, X_r^{[0]}, 0)dr \\
Y_s^{[0]} &= \xi(X_T^{[0]}) + \int_s^T f(r, X_r^{[0]}, Y_r^{[0]}, 0, 0)dr \qquad (5.1)\\
Z^{[0]} &= \psi^{[0]}(\cdot) \equiv 0 \ .
\end{aligned}
$$

which is continuous, deterministic and bounded.

Proof. Thanks to the Lipschitz continuity of b, f with respect to x, y respectively, the claim can be proved by the standard results for the ODEs. □

Let us introduce some notations. We denote, for $s \in [0,T]$,

$$
\begin{aligned}
&b^{[0]}(s) := b(s, X_s^{[0]}, 0), \quad \sigma^{[0]}(s) := \sigma(s, X_s^{[0]}), \quad \gamma^{[0]}(s,z) := \gamma(s, X_s^{[0]}, z) \\
&\xi^{[0]} := \xi(X_T^{[0]}), \quad f^{[0]}(s) := f(s, X_s^{[0]}, Y_s^{[0]}, 0, 0), \\
&\Gamma^{[0]}(s) := \int_{\mathbb{R}_0} \rho(z)\gamma^{[0]}(s,z)\nu(dz) \ .
\end{aligned}
$$

As for derivatives, we denote for example

$$
\begin{aligned}
&\partial_x b^{[0]}(s) := \partial_x b(s, x, 0)\Big|_{x = X_s^{[0]}}, \quad \partial_\epsilon b^{[0]}(s) = \partial_\epsilon b(s, X_s^{[0]}, \epsilon)\Big|_{\epsilon=0} \\
&\partial_x \Gamma^{[0]}(s) := \int_{\mathbb{R}_0} \rho(z)\partial_x \gamma(s, x, z)\Big|_{x = X_s^{[0]}} \nu(dz)
\end{aligned}
$$

and the other terms in the obvious way.

For the first order of the expansion, we have to solve

$$
X_s^{[1]} = \int_0^s \left[\partial_\epsilon b^{[0]}(r) + \partial_x b^{[0]}(r) X_r^{[1]}\right] dr + \int_0^s \sigma^{[0]}(r) dW_r + \int_0^s \int_{\mathbb{R}_0} \gamma^{[0]}(r,z)\widetilde{\mu}(dr,dz), \tag{5.2}
$$

$$
Y_s^{[1]} = \partial_x \xi^{[0]} X_T^{[1]} + \int_s^T \partial_\Theta f^{[0]}(r)\Theta_r^{[1]} dr - \int_s^T Z_r^{[1]} dW_r - \int_s^T \int_{\mathbb{R}_0} \psi_r^{[1]}(z)\widetilde{\mu}(dr,dz) \ . \tag{5.3}
$$

Lemma 5.2. *There exists a unique solution* $\Theta^{[1]}$ *to (5.2) and (5.3) which belongs to* $\mathbb{S}^p[0,T]^{\otimes 4}$ $\forall p \geq 2$. $\hat{\Theta}^{[1]}$ *is given by, for* $s \in [0,T]$ *and* $z \in \mathbb{R}_0$,

$$
\begin{aligned}
Y_s^{[1]} &= y_1^{[1]}(s) X_s^{[1]} + y_0^{[1]}(s) \\
Z_s^{[1]} &= y_1^{[1]}(s)\sigma^{[0]}(s) \qquad (5.4)\\
\psi_s^{[1]}(z) &= y_1^{[1]}(s)\gamma^{[0]}(s,z) \ .
\end{aligned}
$$

Here, $\left(y_1^{[1]}(s), y_0^{[1]}(s), s \in [0,T]\right)$ *are the solutions to the following linear ODEs:*

$$-\frac{dy_1^{[1]}(s)}{ds} = \big(\partial_x b^{[0]}(s) + \partial_y f^{[0]}(s)\big) y_1^{[1]}(s) + \partial_x f^{[0]}(s),$$
$$-\frac{dy_0^{[1]}(s)}{ds} = \partial_y f^{[0]}(s) y_0^{[1]}(s) + \Big(\partial_\epsilon b^{[0]}(s) + \partial_z f^{[0]}(s)\sigma^{[0]}(s) + \partial_u f^{[0]}(s)\Gamma^{[0]}(s)\Big) y_1^{[1]}(s) \tag{5.5}$$

with the terminal conditions $y_1^{[1]}(T) = \partial_x \xi^{[0]}$ *and* $y_0^{[1]}(T) = 0$.

Proof. The existence of the unique solution for $\Theta^{[1]}$ is obvious from Lemmas A.3 and B.2. Since the ODEs are linear with bounded coefficients as well the terminal conditions, they obviously have bounded solutions $(y_0^{[1]}, y_1^{[1]})$. The form of $Y^{[1]}$ is naturally expected from the linear structure of the BSDE and the order of ϵ. It automatically fixes the form of $Z^{[1]}$ and $\psi^{[1]}$. By applying Itô-formula to the hypothesized $Y^{[1]}$ in (5.4) and using (5.5), one can directly confirm (5.4) gives the solution to the BSDE (5.3). This also proves $\Theta^{[1]} \in \mathbb{S}^p[0,T]^{\otimes 4}$ $\forall p \geq 2$. Since the solution of the BSDE is unique, we are done. □

In the second order of ϵ, we need to solve

$$X_s^{[2]} = \int_0^s \Big(\partial_x b^{[0]}(r) X_r^{[2]} + \frac{1}{2}\partial_x^2 b^{[0]}(r)(X_r^{[1]})^2 + \partial_x\partial_\epsilon b^{[0]}(r) X_r^{[1]} + \frac{1}{2}\partial_\epsilon^2 b^{[0]}(r)\Big) dr$$
$$+ \int_0^s \partial_x \sigma^{[0]}(r) X_r^{[1]} dW_r + \int_{\mathbb{R}_0} \partial_x \gamma^{[0]}(r,z) X_r^{[1]} \widetilde{\mu}(dr,dz) \tag{5.6}$$

and

$$Y_s^{[2]} = \partial_x \xi^{[0]} X_T^{[2]} + \frac{1}{2}\partial_x^2 \xi^{[0]} (X_T^{[1]})^2 + \int_s^T \Big(\partial_\Theta f^{[0]}(r)\Theta_r^{[2]} + \frac{1}{2}\partial_\Theta^2 f^{[0]}(r)\Theta_r^{[1]}\Theta_r^{[1]}\Big) dr$$
$$- \int_s^T Z_r^{[2]} dW_r - \int_s^t \psi_r^{[2]}(z)\widetilde{\mu}(dr,dz)\ . \tag{5.7}$$

You can see that the dynamics of $X^{[2]}$ is linear in $X^{[2]}$ and contains $\{(X^{[1]})^j, j \leq 2\}$. The BSDE for $\hat{\Theta}^{[2]}$ is linear in itself and contains $\{(\Theta^{[1]})^j, j \leq 2\}$. Since we have seen $\hat{\Theta}^{[1]}$ is linear in $X^{[1]}$, the driver contains $\{(X^{[1]})^j, j \leq 2\}$. Suppose that $\hat{\Theta}^{[2]}$ is linear in $X^{[2]}$ and quadratic in $X^{[1]}$. Then, one can check that this is also the case for the driver of $Y^{[2]}$ and hence consistent with the initial assumption. In fact, although it becomes a bit more tedious, one can prove the next lemma exactly in the same way as Lemma 5.2 by directly comparing the result of Itô-formula with the driver of the BSDE.

Lemma 5.3. *There exists a unique solution* $\Theta^{[2]}$ *to (5.6) and (5.7) which belongs to* $\mathbb{S}^p[0,T]^{\otimes 4}$ $\forall p \geq 2$. $\hat{\Theta}^{[2]}$ *is given by, for* $s \in [0,T]$ *and* $z \in \mathbb{R}_0$,

$$Y_s^{[2]} = y_2^{[2]}(s) X_s^{[2]} + y_{1,1}^{[2]}(s)(X_s^{[1]})^2 + y_1^{[2]}(s) X_s^{[1]} + y_0^{[2]}(s)$$
$$Z_s^{[2]} = X_{s-}^{[1]}\Big(y_2^{[2]}(s)\partial_x\sigma^{[0]}(s) + 2y_{1,1}^{[2]}\sigma^{[0]}(s)\Big) + y_1^{[2]}(s)\sigma^{[0]}(s)$$
$$\psi_s^{[2]}(z) = X_{s-}^{[1]}\Big(y_2^{[2]}(s)\partial_x\gamma^{[0]}(s,z) + 2y_{1,1}^{[2]}(s)\gamma^{[0]}(s,z)\Big) + y_{1,1}^{[2]}(s)(\gamma^{[0]}(s,z))^2 + y_1^{[2]}(s)\gamma^{[0]}(s,z)\ .$$

Here, $\left(y_2^{[2]}(s), y_{1,1}^{[2]}(s), y_1^{[2]}(s), y_0^{[2]}(s), s \in [0,T]\right)$ *are the solutions to the following linear*

ODEs:

$$
\begin{aligned}
-\frac{dy_2^{[2]}(s)}{ds} &= \Big(\partial_x b^{[0]}(s)+\partial_y f^{[0]}(s)\Big)y_2^{[2]}(s)+\partial_x f^{[0]}(s)\\
-\frac{dy_{1,1}^{[2]}(s)}{ds} &= \Big(2\partial_x b^{[0]}(s)+\partial_y f^{[0]}(s)\Big)y_{1,1}^{[2]}(s)+\frac{1}{2}\partial_x^2 f^{[0]}(s)\\
&\quad +\frac{1}{2}\partial_x^2 b^{[0]}(s)y_2^{[2]}(s)+\partial_x\partial_y f^{[0]}(s)y_1^{[1]}(s)+\frac{1}{2}\partial_y^2 f^{[0]}(s)(y_1^{[1]}(s))^2
\end{aligned}
$$

$$
\begin{aligned}
-\frac{dy_1^{[2]}(s)}{ds} &= \Big(\partial_x b^{[0]}(s)+\partial_y f^{[0]}(s)\Big)y_1^{[2]}(s)+\partial_x\partial_\epsilon b^{[0]}(s)y_2^{[2]}(s)+2\partial_\epsilon b^{[0]}(s)y_{1,1}^{[2]}(s)\\
&\quad +\partial_z f^{[0]}(s)\Big(y_2^{[2]}(s)\partial_x\sigma^{[0]}(s)+2y_{1,1}^{[2]}(s)\sigma^{[0]}(s)\Big)\\
&\quad +\partial_u f^{[0]}(s)\Big(y_2^{[2]}(s)\partial_x\Gamma^{[0]}(s)+2y_{1,1}^{[2]}(s)\Gamma^{[0]}(s)\Big)\\
&\quad +\partial_y^2 f^{[0]}(s)y_1^{[1]}(s)y_0^{[1]}(s)+\partial_x\partial_y f^{[0]}(s)y_0^{[1]}(s)\\
&\quad +y_1^{[1]}(s)\Big(\partial_x\partial_z f^{[0]}(s)\sigma^{[0]}(s)+\partial_x\partial_u f^{[0]}(s)\Gamma^{[0]}(s)\Big)\\
&\quad +(y_1^{[1]}(s))^2\Big(\partial_y\partial_z f^{[0]}(s)\sigma^{[0]}(s)+\partial_y\partial_u f^{[0]}(s)\Gamma^{[0]}(s)\Big)\\
-\frac{dy_0^{[2]}(s)}{ds} &= \partial_y f^{[0]}(s)y_0^{[2]}(s)+y_{1,1}^{[2]}(s)\Big((\sigma^{[0]}(s))^2+\int_{\mathbb{R}_0}(\gamma^{[0]}(s,z))^2\nu(dz)\Big)\\
&\quad +\frac{1}{2}\partial_\epsilon^2 b^{[0]}(s)y_2^{[2]}(s)+\partial_\epsilon b^{[0]}(s)y_1^{[2]}(s)+y_1^{[2]}(s)\Big(\partial_z f^{[0]}(s)\sigma^{[0]}(s)+\partial_u f^{[0]}(s)\Gamma^{[0]}(s)\Big)\\
&\quad +y_{1,1}^{[2]}(s)\partial_u f^{[0]}(s)\int_{\mathbb{R}_0}\rho(z)(\gamma^{[0]}(s,z))^2\nu(dz)+\frac{1}{2}\partial_y^2 f^{[0]}(s)(y_0^{[1]}(s))^2\\
&\quad +(y_1^{[1]}(s))^2\Big(\frac{1}{2}\partial_z^2 f^{[0]}(s)(\sigma^{[0]}(s))^2+\frac{1}{2}\partial_u^2 f^{[0]}(s)(\Gamma^{[0]}(s))^2+\partial_z\partial_u f^{[0]}(s)\sigma^{[0]}(s)\Gamma^{[0]}(s)\Big)\\
&\quad +(y_1^{[1]}(s)y_0^{[1]}(s))\Big(\partial_y\partial_z f^{[0]}(s)\sigma^{[0]}(s)+\partial_y\partial_u f^{[0]}(s)\Gamma^{[0]}(s)\Big)
\end{aligned}
$$

with terminal conditions $y_2^{[2]}(T)=\partial_x\xi^{[0]}$*,* $y_{1,1}^{[2]}(T)=\frac{1}{2}\partial_x^2\xi^{[0]}$*,* $y_1^{[2]}(T)=y_0^{[2]}(T)=0$.

One can repeat the procedures to an any order $n\le n_{\max}$. This can be checked in the following way. By a simple modification of (4.2) gives

$$
Y_s^{[n]}=G_n+\int_s^T\Big\{F_{n,r}+\partial_\Theta f^{[0]}(r)\Theta_r^{[n]}\Big\}dr-\int_s^T Z_r^{[n]}dW_r-\int_s^T\int_{\mathbb{R}_0}\psi_r^{[n]}(z)\widetilde{\mu}(dr,dz)
$$

where

$$
\begin{aligned}
G_n &:= \sum_{k=1}^n\sum_{\beta_1+\cdots+\beta_k=n,\beta_i\ge 1}\frac{1}{k!}\partial_x^k\xi(X_T^{[0]})\prod_{j=1}^k X_T^{[\beta_j]},\\
F_{n,r} &:= \sum_{k=2}^n\sum_{\beta_1+\cdots+\beta_k=n,\beta_i\ge 1}\sum_{i_x=0}^k\sum_{i_y=0}^{k-i_x}\sum_{i_z=0}^{k-i_x-i_y}\frac{\partial_x^{i_x}\partial_y^{i_y}\partial_z^{i_z}\partial_u^{k-i_x-i_y-i_z}f^{[0]}(r)}{i_x!i_y!i_z!(k-i_x-i_y-i_z)!}\\
&\times\prod_{j_x=1}^{i_x}X_r^{[\beta_{j_x}]}\prod_{j_y=i_x+1}^{i_x+i_y}Y_r^{[\beta_{j_y}]}\prod_{j_z=i_x+i_y+1}^{i_x+i_y+i_z}Z_r^{[\beta_{j_z}]}\prod_{j_u=i_x+i_y+i_z+1}^{k}\int_{\mathbb{R}_0}\rho(z)\psi_r^{[\beta_{j_u}]}(z)\nu(dz).
\end{aligned}
$$

From the shapes of $G_n, F_{n,r}$, one can confirm that $\hat{\Theta}_r^{[n]}$ is given by the polynomials

$$\left\{\prod_{j=1}^{k} X_r^{[\beta_j]}; \beta_1+\cdots+\beta_k = m\ (\beta_i\geq 1),\ k\leq m,\ m\leq n\right\}$$

by induction. Since $\Theta^{[n]}$ appears only linearly both in the forward and backward SDEs the relevant ODEs become always linear.

5.2 A polynomial scheme

We have just seen that the grading structure both for $\{X^{[n]}\}_{n\geq 0}$ and $\{\hat{\Theta}^{[n]}\}_{n\geq 0}$ played an important role. In particular, even if $\{\hat{\Theta}^{[n]}\}_{n\geq 0}$ has a grading structure, one cannot obtain the system of linear ODEs unless $\{X^{[n]}\}_{n\geq 0}$ shares the same features. Suppose that the dynamics of $X^{t,x}$ is linear in itself. Then, one need not expand the forward SDE and thus one may obtain the expansion of $\hat{\Theta}^{t,x,\epsilon}$ in terms of polynomials of $X^{t,x}$. If this is the case, the ODEs for the associated coefficients required in each order will be greatly simplified.

Let us consider the following forward-backward SDEs for $s\in[t,T]$:

$$\begin{aligned}
X_s^{t,x} &= x+\int_t^s\Big(b^0(r)+b^1(r)X_r^{t,x}\Big)dr+\int_t^s\Big(\sigma^0(r)+\sigma^1(r)X_r^{t,x}\Big)dW_r\\
&\quad+\int_s^t\int_E\Big(\gamma^0(r,z)+\gamma^1(r,z)X_{r-}^{t,x}\Big)\widetilde{\mu}(dr,dz) \qquad (5.8)\\
Y_s^{t,x,\epsilon} &= \xi(\epsilon X_T^{t,x})+\int_s^T f\Big(r,\epsilon X_r^{t,x},Y_r^{t,x,\epsilon},Z_r^{t,x,\epsilon},\int_{\mathbb{R}_0}\rho(z)\psi_r^{t,x,\epsilon}(z)\nu(dz)\Big)dr\\
&\quad-\int_s^T Z_r^{t,x,\epsilon}dW_r-\int_s^T\int_E\psi_r^{t,x,\epsilon}(z)\widetilde{\mu}(dr,dz)\ . \qquad (5.9)
\end{aligned}$$

where $b^0:[0,T]\to\mathbb{R}^d$, $b^1:[0,T]\to\mathbb{R}^{d\times d}$, $\sigma^0:[0,T]\to\mathbb{R}^{d\times l}$, $\sigma^1:[0,T]\to\mathbb{R}^{d\times d\times l}$, $\gamma^0:[0,T]\times E\to\mathbb{R}^{d\times k}$, $\gamma^1:[0,T]\times E\to\mathbb{R}^{d\times d\times k}$ are measurable functions and ξ, f are defined as before.

Assumption 5.1. *The functions $\{b^i(t),\sigma^i(t),\gamma^i(t,z)\}, i\in\{0,1\}$ are continuous. Furthermore, there exists some positive constant K such that $\Big(|b^i(t)|+|\sigma^i(t)|+|\gamma^i(t,z)|/\eta(z)\leq K\Big)$ for $i\in\{0,1\}$ uniformly in $(t,z)\in[0,T]\times E$.*

With slight abuse of notation, let us use $\Theta_r^{t,x,\epsilon} := \Big(\epsilon X_r^{t,x}, Y_r^{t,x,\epsilon}, Z_r^{t,x,\epsilon}, \int_{\mathbb{R}_0}\rho(z)\psi_r^{t,x,\epsilon}(z)\nu(dz)\Big)$ in this subsection.

Theorem 5.1. *Under Assumptions 3.2 and 5.1, there exists a unique solution $\hat{\Theta}^{t,x,\epsilon}$ to the BSDE (5.9), its classical derivative $\partial_\epsilon^n\hat{\Theta}^{t,x,\epsilon}\in\mathcal{K}^p[0,T]\ \forall p\geq 2$ exists for every $0\leq n\leq n^{\mathrm{ae}}$ and is given by the solution of the following BSDE:*

$$\begin{aligned}
\partial_\epsilon^n Y_s^{t,x,\epsilon} &= g_n(X_T^{t,x})^n+\int_s^T\Big\{h_{n,r}+\partial_x^n f(r,\Theta_r^{t,x,\epsilon})(X_r^{t,x})^n+\partial_{\hat{\Theta}}f(r,\Theta_r^{t,x,\epsilon})\partial_\epsilon^n\hat{\Theta}_r^{t,x,\epsilon}\Big\}dr\\
&\quad-\int_s^T\partial_\epsilon^n Z_r^{t,x,\epsilon}dW_r-\int_s^T\int_E\partial_\epsilon^n\psi_r^{t,x,\epsilon}(z)\widetilde{\mu}(dr,dz)
\end{aligned}$$

where $g_n := \partial_x^n \xi(\epsilon X_T^{t,x})$ *and*

$$h_{n,r} := n! \sum_{k=2}^{n} \sum_{i_x=0}^{k-1} \sum_{i_y=0}^{k-i_x} \sum_{i_z=0}^{k-i_x-i_y} \sum_{\beta_{i_x+1}+\cdots+\beta_k=n-i_x,\ \beta_i\geq 1} \frac{\partial_x^{i_x}\partial_y^{i_y}\partial_z^{i_z}\partial_u^{k-i_x-i_y-i_z} f(r,\Theta_r^{t,x,\epsilon})}{i_x! i_y! i_z! (k-i_x-i_y-i_z)!} (X_r^{t,x})^{i_x}$$
$$\times \prod_{j_y=i_x+1}^{i_x+i_y} \frac{1}{\beta_{j_y}!} \partial_\epsilon^{\beta_{j_y}} Y_r^{t,x,\epsilon} \prod_{j_z=i_x+i_y+1}^{i_x+i_y+i_z} \frac{1}{\beta_{j_z}!} \partial_\epsilon^{\beta_{j_z}} Z_r^{t,x,\epsilon} \prod_{j_u=i_x+i_y+i_z+1}^{k} \frac{1}{\beta_{j_u}!} \int_{\mathbb{R}_0} \rho(z) \partial_\epsilon^{\beta_{j_u}} \psi_r^{t,x,\epsilon}(z)\nu(dz)\ .$$

Moreover, for every $0 \leq n \leq n_{\max}+1$, $\partial_\epsilon^n \hat{\Theta}^{t,x,\epsilon} \in \mathbb{S}^p[t,T]^{\otimes 3}$ $\forall p \geq 2$. *The asymptotic expansion of* $\hat{\Theta}^{t,x,\epsilon}$ *with respect to* ϵ *satisfies, with some positive constant* C_p, *that*

$$\left|\left| \hat{\Theta}^{t,x,\epsilon} - \Big(\hat{\Theta}^{[0]} + \sum_{n=1}^{N} \epsilon^n \hat{\Theta}^{[n]} \Big) \right|\right|_{\mathbb{S}^p[t,T]} \leq \epsilon^{N+1} C_p.$$

for every $1 \leq N \leq n_{\max}$.

Proof. One can follow the same arguments in Proposition 4.2 and Theorem 4.1 by replacing $(X^{t,x,\epsilon})$ by $(\epsilon X^{t,x})$. Since there is no ϵ dependence through $X^{t,x}$ in the expressions $Y_s^{t,x,\epsilon} = u(s, X_s^{t,x}, \epsilon)$ and $Z_s^{t,x,\epsilon} = \partial_x u(x, X_{s-}^{t,x}, \epsilon)\sigma(s, X_{s-}^{t,x}, \epsilon)$, one-time differentiability with respect to x and its polynomial growth property are enough to show recursively that $\partial_\epsilon^n \hat{\Theta}^{t,x,\epsilon} \in \mathbb{S}^p[t,T]$ for $\forall p \geq 2$. □

Remark 5.1. *The above result also justifies the method proposed in Fujii (2015) [20] for the underlying* X *with linear dynamics. As for a general Affine-like process* X *(such as* $\sigma(x) = \sqrt{x}$*), it is difficult to prove within the current technique due to its non-Lipschitz nature.*

It is not difficult to see that $(\hat{\Theta}_s^{[n]}, s \in [t,T])$ is given by the unique solution to the following BSDE:

$$Y_s^{[n]} = \frac{1}{n!}\partial_x^n \xi(0)(X_T^{t,x})^n + \int_s^T \Big\{ \widetilde{h}_{n,r} + \frac{1}{n!}\partial_x^n f^{[0]}(r)(X_r^{t,x})^n + \partial_{\hat{\Theta}} f^{[0]}(r)\hat{\Theta}_r^{[n]} \Big\} dr$$
$$- \int_s^T Z_r^{[n]} dW_r - \int_s^T \int_E \psi_r^{[n]}(z)\widetilde{\mu}(dr,dz) \tag{5.10}$$

where

$$\widetilde{h}_{n,r} := \sum_{k=2}^{n} \sum_{i_x=0}^{k-1} \sum_{i_y=0}^{k-i_x} \sum_{i_z=0}^{k-i_x-i_y} \sum_{\beta_{i_x+1}+\cdots+\beta_k=n-i_x,\beta_i\geq 1} \frac{\partial_x^{i_x}\partial_y^{i_y}\partial_z^{i_z}\partial_u^{k-i_x-i_y-i_z} f^{[0]}(r)}{i_x! i_y! i_z! (k-i_x-i_y-i_z)!} (X_r^{t,x})^{i_x}$$
$$\times \prod_{j_y=i_x+1}^{i_x+i_y} Y_r^{[\beta_{j_y}]} \prod_{j_z=i_x+i_y+1}^{i_x+i_y+i_z} Z_r^{[\beta_{j_z}]} \prod_{j_u=i_x+i_y+i_z+1}^{k} \int_{\mathbb{R}_0} \rho(z)\psi_r^{[\beta_{j_u}]}(z)\nu(dz)$$

and $f^{[0]}(r) := f(r, 0, Y_r^{[0]}, 0, 0)$. Since $(i_x + \sum_{j_y}\beta_{j_y} + \sum_{j_z}\beta_{j_z} + \sum_{j_u}\beta_{j_u}) = n$, one can recursively show that $\hat{\Theta}_r^{[n]}$ is given by the polynomials $\Big\{ (X_r^{t,x})^j, 0 \leq j \leq n \Big\}$ and every coefficient is determined by the system of linear ODEs as in Section 5.1, which we leave as a simple exercise.

An exponential Lévy case

In the reminder of this section, let us deal with a special example of an exponential (time-inhomogeneous) Lévy dynamics for X. Let us put $m = d = l = k = 1$ and $t = 0$ for simplicity and consider $b^0 = \sigma^0 = \gamma^0 = 0$

$$X_s = x + \int_t^s X_r\Big(b(r)dr + \sigma(r)dW_r\Big) + \int_s^t \int_{\mathbb{R}_0} X_{r-}\gamma(r,z)\widetilde{\mu}(dr,dz) \tag{5.11}$$

with $b := b^1, \sigma := \sigma^1, \gamma := \gamma^1$ in (5.8). We omit the superscript denoting the initial data $(0, x)$. Let us introduce the notations: $q(s,j) := \int_{\mathbb{R}_0}(\gamma(s,z))^j\nu(dz)$ for $j \geq 2$, $\Gamma(s,j) := \int_{\mathbb{R}_0}\rho(z)\big[(1+\gamma(s,z))^j - 1\big]\nu(dz)$ for $j \geq 1$ and $C_{n,j} := n!/(j!(n-j)!)$ for $j \leq n, n \geq 2$.

Theorem 5.2. *Under Assumptions 3.1, 5.1, $m = d = l = k = 1$ and $t = 0$, the asymptotic expansion of the forward-backward SDEs (5.11) and (5.9) is given by, for $s \in [0,T]$,*

$$\begin{aligned} Y_s^{[0]} &= \xi(0) + \int_s^T f(r,0,Y_r^{[0]},0,0)dr \\ Z^{[0]} &= \psi^{[0]} = 0 \end{aligned} \tag{5.12}$$

and, for $1 \leq n \leq n_{\max}$,

$$\begin{aligned} Y_s^{[n]} &= (X_s)^n y^{[n]}(s) \\ Z_s^{[n]} &= (X_{s-})^n y^{[n]}(s) n\sigma(s) \\ \psi_s^{[n]}(z) &= (X_{s-})^n y^{[n]}(s)\big[(1+\gamma(s,z))^n - 1\big] \end{aligned}$$

where the functions $\{y^{[j]}(s), s \in [0,T]\}_{1\leq j\leq n}$ are determined recursively by the following system of linear ODEs:

$$\begin{aligned} -\frac{dy^{[n]}(s)}{ds} \quad = \quad & \Big(nb(s) + \frac{1}{2}n(n-1)\sigma^2(s) + \sum_{j=2}^n C_{n,j}q(s;j) + \partial_y f^{[0]}(s) \\ & + \partial_z f^{[0]}(s)n\sigma(s) + \partial_u f^{[0]}(s)\Gamma(s;n)\Big)y^{[n]}(s) + \frac{1}{n!}\partial_x^n f^{[0]}(s) \\ & + \sum_{k=2}^n \sum_{i_x=0}^{k-1} \sum_{i_y=0}^{k-i_x-i_y} \sum_{\beta_{i_x+1}+\cdots+\beta_k = n-i_x, \beta_i \geq 1} \left\{ \frac{\partial_x^{i_x}\partial_y^{i_y}\partial_z^{i_z}\partial_u^{k-i_x-i_y-i_z} f^{[0]}(s)}{i_x! i_y! i_z!(k-i_x-i_y-i_z)!} \right. \\ & \times \prod_{j_y=i_x+1}^{i_x+i_y} \Big(y^{[\beta_{j_y}]}(s)\Big) \prod_{j_z=i_x+i_y+1}^{i_x+i_y+i_z} \Big(\beta_{j_z}\sigma(s)y^{[\beta_{j_z}]}(s)\Big) \\ & \left. \times \prod_{j_u=i_x+i_y+i_z+1}^{k} \Big(\Gamma(s;\beta_{j_u})y^{[\beta_{j_u}]}(s)\Big) \right\} \end{aligned}$$

with a terminal condition $y^{[n]}(T) = \partial_x^n \xi(0)/n!$ for every n. Here, $f^{[0]}(r)$ is defined by $f(r,0,Y_r^{[0]},0,0)$ using $Y^{[0]}$ determined by (5.12).

Proof. If one supposes the form of the solution as $Y_s^{[n]} = (X_s)^n y^{[n]}(s)$, then $Z^{[n]}$ and $\psi^{[n]}$ must have the form as given. Comparing the result of Itô formula applied to $X^n y^{[n]}$ and the form of the BSDE (5.10) substituted by the hypothesized form of $\{\hat{\Theta}^{[\beta]}\}_{\beta\leq n}$, one obtains the system of ODEs given above. Since every ODE is linear, there exists a solution for every $y^{[n]}$, $1 \le n \le n_{\max}$. Since the solution of the BSDE is unique, this must be the desired solution. $\square$

Remark

It is interesting to observe the difference from the linearization scheme proposed in [22] for a Brownian setup. There, the BSDE is expanded around a linear driver in the first step. Then in the second step the resultant set of linear BSDEs are evaluated by the small-variance asymptotic expansion of the forward SDE, or by the interacting particle simulation method proposed in Fujii & Takahashi (2015) [23]. Hence, in order for the scheme of [22] works well, it requires the smallness of the non-linear terms in the driver f, although it naturally arises in many applications. Furthermore, due to the presence of large number of conditional expectations, calculating them analytically without invoking the particle simulation technique [23] is unrealistic in most of the practical situations.

On the other hand, in the current scheme, the expansion of the driver is not directly performed and the significant part of non-linearity is taken into account at the zero-th order around the mean dynamics of the forward SDE as observed in (5.1). The effects of the stochasticity from the forward SDE are then taken into account perturbatively around this "mean" solution. Therefore, the current scheme is expected to be more advantageous when there exists significant non-linearity in the driver. Furthermore, the special grading structure of approximating FBSDEs makes them explicitly solvable by ODEs without using any Monte-Carlo simulation. Since the approximate solution of $(Y, Z, \psi(\cdot))$ is explicitly given as a polynomial in the stochastic flows of X, one can obtain not only the current value $(Y_0, Z_0, \psi_0(\cdot))$ but also its evolution by simply simulating the flows of X (or X itself for the polynomial case). Some numerical examples and empirical error estimates are available in Fujii (2015) [20] based on this property for a certain class of models.

A Useful a priori estimates: forward SDEs

Let us summarize the useful a priori estimates for FSDEs with jumps. The following result taken from Lemma 5-1 of Bichteler, Gravereaux and Jacod (1987) [4] is essential for analysis of a σ-finite random measure.

Lemma A.1. *Let $\eta : \mathbb{R} \to \mathbb{R}$ be defined by $\eta(z) = 1 \wedge |z|$. Then, for $\forall p \ge 2$, there exists a constant δ_p depending on p, T, m, k such that*

$$\mathbb{E}\left[\sup_{t\in[0,T]}\left|\int_0^t \int_E U(s,z)\widetilde{\mu}(ds,dz)\right|^p\right] \le \delta_p \int_0^T \mathbb{E}|L_s|^p ds \tag{A.1}$$

if U is an $\mathbb{R}^{m\times k}$-valued $\mathcal{P}\otimes\mathcal{E}$-measurable function on $\Omega\times[0,T]\times E$ and L is a predictable process satisfying $|U_{\cdot,i}(\omega,s,z)| \le L_s(\omega)\eta(z)$ for each column $1 \le i \le k$.

Since $\int_E \eta(z)^p \nu(dz) < \infty$ for $\forall p \ge 2$, the above lemma tells that one can use a BDG-like inequality with a compensator ν whenever the integrand of the random measure divided by η is dominated by some integrable random variable. The following result from Chapter 1 Section 9 Lemma 6 of Liptser & Shiryayev (1989) [37] or Lemma 2.1 of Dzhaparidze & Valkeila (1990) [15] is also important.

Lemma A.2. *Let ψ belong to $\mathbb{H}^2_\nu[0,T]$. Then, for $p \geq 2$, there exists some constant $C_p > 0$ depending only on p such that*

$$\mathbb{E}\Big(\int_0^T \int_E |\psi_s(z)|^2 \nu(dz)ds\Big)^{p/2} \leq C_p \mathbb{E}\Big(\int_0^T \int_E |\psi_s(z)|^2 \mu(ds,dz)\Big)^{p/2} .$$

For $t_1 \leq t_2 \leq T$ and $\mathbb{R}^d$-valued $\mathcal{F}_{t_i}$-measurable random variable x^i, let us consider $\{X^i_t, t \in [t_i, T]\}_{1 \leq i \leq 2}$ as a solution of the following SDE:

$$X^i_t = x^i + \int_{t_i}^t \widetilde{b}^i(s, X^i_s)ds + \int_{t_i}^t \widetilde{\sigma}^i(s, X^i_s)dW_s + \int_{t_i}^t \int_E \widetilde{\gamma}^i(s, X^i_{s-}, z)\widetilde{\mu}(ds, dz) \qquad \text{(A.2)}$$

where $\widetilde{b}^i : \Omega \times [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$, $\widetilde{\sigma}^i : \Omega \times [0,T] \times \mathbb{R}^d \to \mathbb{R}^{d \times l}$, and $\widetilde{\gamma}^i : \Omega \times [0,T] \times \mathbb{R}^d \times E \to \mathbb{R}^{d \times k}$.

Assumption A.1. *For $i \in \{1,2\}$, the map $(\omega,t) \mapsto \widetilde{b}^i(\omega,t,\cdot)$ is $\mathbb{F}$-progressively measurable, $(\omega,t) \mapsto \widetilde{\sigma}^i(\omega,t,\cdot), \widetilde{\gamma}^i(\omega,t,\cdot)$ are $\mathbb{F}$-predictable, and there exists some constant $K > 0$ such that, for every $x, x' \in \mathbb{R}^d$ and $z \in E$,*

$$\begin{aligned}
&|\widetilde{b}^i(\omega,t,x) - \widetilde{b}^i(\omega,t,x')| + |\widetilde{\sigma}^i(\omega,t,x) - \widetilde{\sigma}^i(\omega,t,x')| \leq K|x - x'| \\
&|\widetilde{\gamma}^i_{\cdot,j}(\omega,t,x,z) - \widetilde{\gamma}^i_{\cdot,j}(\omega,t,x',z)| \leq K\eta(z)|x - x'|, \quad 1 \leq j \leq k
\end{aligned}$$

$d\mathbb{P} \otimes dt$-a.e. in $\Omega \times [0,T]$. Furthermore, for some $p \geq 2$,

$$\mathbb{E}\left[|x^i|^p + \Big(\int_{t_i}^T |\widetilde{b}^i(s,0)|ds\Big)^p + \Big(\int_{t_i}^T |\widetilde{\sigma}^i(s,0)|^2 ds\Big)^{p/2} + \int_{t_i}^T |L^i_s|^p ds\right] < \infty$$

where L^i is some $\mathbb{F}$-predictable process satisfying $|\widetilde{\gamma}^i_{\cdot,j}(\omega,t,0,z)| \leq L^i_t(\omega)\eta(z)$ for every column vector $\{\widetilde{\gamma}^i_{\cdot,j},\ 1 \leq j \leq k\}$.

The following lemma is an extension of Lemma A.1 given in [6] to a σ-finite measure by using (A.1).

Lemma A.3. *Under Assumption A.1, the SDE (A.2) has a unique solution and there exists some constant $C_p > 0$ such that,*

$$\begin{aligned}
||X^i||^p_{\mathbb{S}^p_d[t_i,T]} \leq C_p \mathbb{E}\Bigg[&|x^i|^p + \Big(\int_{t_i}^T |\widetilde{b}^i(s,0)|ds\Big)^p \\
&+ \Big(\int_{t_i}^T |\widetilde{\sigma}^i(s,0)|^2 ds\Big)^{p/2} + \int_{t_i}^T |L^i_s|^p ds\Bigg]
\end{aligned} \qquad \text{(A.3)}$$

and, for all $t_i \leq s \leq t \leq T$,

$$\mathbb{E}\left[\sup_{s \leq u \leq t} |X^i_u - X^i_s|^p\right] \leq C_p A^i_p |t - s| \qquad \text{(A.4)}$$

where

$$A^i_p := \mathbb{E}\left[|x^i|^p + ||\widetilde{b}^i(\cdot,0)||^p_{[t_i,T]} + ||\widetilde{\sigma}^i(\cdot,0)||^p_{[t_i,T]} + ||L^i||^p_{[t_i,T]}\right] .$$

Moreover, for $t_2 \le t \le T$,

$$||\delta X||^p_{\mathbb{S}^p_d[t_2,T]} \le C_p\Big(\mathbb{E}|x^1-x^2|^p + A^1_p|t_2-t_1|\Big)$$
$$+C_p\mathbb{E}\left[\Big(\int_{t_2}^T |\delta\widetilde{b}_t|dt\Big)^p + \Big(\int_{t_2}^T |\delta\widetilde{\sigma}_t|^2dt\Big)^{p/2} + \int_{t_2}^T |\delta L_t|^p dt\right] \qquad \text{(A.5)}$$

where $\delta X := X^1 - X^2$, $\delta\widetilde{b}_\cdot := (\widetilde{b}^1 - \widetilde{b}^2)(\cdot, X^1_\cdot)$, $\delta\widetilde{\sigma}_\cdot := (\widetilde{\sigma}^1 - \widetilde{\sigma}^2)(\cdot, X^1_\cdot)$ and δL is a predictable process satisfying $|\delta\widetilde{\gamma}|(\omega,t,z) \le \delta L_t(\omega)\ \eta(z)$, $d\mathbb{P}\otimes dt$-a.e. in $\Omega\times[0,T]$, where $\delta\widetilde{\gamma}(\omega,t,z) := (\widetilde{\gamma}^1-\widetilde{\gamma}^2)(\omega,t,X^1_{t-}(\omega),z)$.

Proof. The existence of a unique solution is given in pp.237 of Gikhman & Skorohod (1972) [26] or Section 6.2 of Applebaum (2009) [1], for example. For the sake of completeness, let us give a sketch of proof for the other estimates.

Set a sequence of stopping times $\Big(\tau_n := \inf\{t \ge t_i; |X^i_s| \ge n\}\wedge T,\ n\in\mathbb{N}\Big)$. Then, using the fact that $|\widetilde{\gamma}^i(s,X^i_{s-},z)| \le (L^i_s + K|X^i_{s-}|)\eta(z)$, Lemma A.1 and the Burkholder-Davis-Gundy (BDG) inequality, one obtains

$$\begin{aligned}\mathbb{E}|X^i_{\tau_n}|^p &\le C_p\int_{t_i}^{\tau_n}\mathbb{E}|X^i_s|^p ds\\ &+ C_p\mathbb{E}\left[|x^i|^p + \Big(\int_{t_i}^{\tau_n}|\widetilde{b}^i(s,0)|ds\Big)^p + \Big(\int_{t_i}^{\tau_n}|\widetilde{\sigma}^i(s,0)|^2ds\Big)^{p/2} + \int_{t_i}^{\tau_n}|L^i_s|^p ds\right]\ .\end{aligned}$$

Using the Gronwall inequality and passing to the limit $n\to\infty$, one obtains the estimate for $\Big(\sup_{t\in[t_i,T]}\mathbb{E}|X^i_t|^p\Big)$. Using the BDG inequality and Lemma A.1 once again, one obtains the first estimate (A.3). A similar analysis yields

$$\mathbb{E}\sup_{u\in[s,t]}|X^i_u - X^i_s|^p \le C_p\mathbb{E}\left[\Big(\int_s^t|\widetilde{b}^i(r,0)|dr\Big)^p + \Big(\int_s^t|\widetilde{\sigma}^i(r,0)|^2dr\Big)^{p/2} + \int_s^t|L^i_r|^p dr\right]$$
$$+C_p(t-s)\mathbb{E}||X^i||^p_{[t_i,T]},$$

which gives second estimate (A.4).

As for the last estimate (A.5), notice first that

$$|\widetilde{\gamma}^1-\widetilde{\gamma}^2|(s,X^1_{s-},z) \le (L^1_s + L^2_s + 2K|X^1_{s-}|)\eta(z)\ .$$

Since $X^1\in\mathbb{S}^p$, there exists a predictable process δL satisfying $|\widetilde{\gamma}^1-\widetilde{\gamma}^2|(s,X^1_{s-},z) \le \delta L_s\eta(z)$, $d\mathbb{P}\otimes ds$-a.e. and $\int_{t_2}^T\mathbb{E}|\delta L_r|^p dr < \infty$ as desired. Separating the integration range, applying the BDG inequality and Lemma A.1, one obtains

$$\begin{aligned}\mathbb{E}||\delta X||^p_{[t_2,t]} \le\ & C_p\mathbb{E}\left[|x^1-x^2|^p + \Big(\int_{t_1}^{t_2}|\widetilde{b}^1(s,0)|ds\Big)^p + \Big(\int_{t_1}^{t_2}|\widetilde{\sigma}^1(s,0)|^2ds\Big)^{p/2}\right.\\ &\left.+\int_{t_1}^{t_2}|L^1_s|^p ds + (t_2-t_1)||X^1||^p_{[t_1,t_2]}\right] + C_p\mathbb{E}\left[\int_{t_2}^t|\delta X_s|^p ds\right.\\ &\left.+\Big(\int_{t_2}^t|\delta\widetilde{b}_s|ds\Big)^p + \Big(\int_{t_2}^t|\delta\widetilde{\sigma}_s|^2ds\Big)^{p/2} + \int_{t_2}^t|\delta L_s|^p ds\right]\ .\end{aligned}$$

Using the first two results and the Gronwall inequality, one obtains (A.5). □

Remark

Note that when $p=2$, one can replace $\int^\cdot |L^i_s|^2 ds$ (resp. $\int^\cdot |\delta L_s|^2 ds$) by $\int^\cdot \int_E |\widetilde{\gamma}^i(s,0,z)|^2\nu(dz)ds$ (resp. $\int^\cdot \int_E |\delta\widetilde{\gamma}(s,z)|^2\nu(dz)ds$) by simply applying the BDG inequality. Furthermore, when the compensator is finite $\nu(E)<\infty$, the above replacement is possible for any $\forall p\geq 2$ thanks to Lemma B.3 (see below).

B Useful a priori estimates: BSDEs

Consider the following BSDE:

$$Y_t = \widetilde{\xi} + \int_t^T \widetilde{f}(s,Y_s,Z_s,\psi_s)ds - \int_t^T Z_s dW_s - \int_t^T \int_E \psi_s(z)\widetilde{\mu}(ds,dz) \ , \tag{B.1}$$

where $\xi:\Omega\to\mathbb{R}^m$, $\widetilde{f}:\Omega\times[0,T]\times\mathbb{R}^m\times\mathbb{R}^{m\times l}\times\mathbb{L}^2(E,\mathcal{E},\nu;\mathbb{R}^m)\to\mathbb{R}^m$. In this section, we use $\langle\cdot,\cdot\rangle$ to denote an inner product of m-dimensional vectors for clarity.

Assumption B.1. *(i) $\widetilde{\xi}$ is $\mathcal{F}_T$-measurable and the map $(\omega,t)\mapsto\widetilde{f}(\omega,t,\cdot)$ is $\mathbb{F}$-progressively measurable. There exists a solution (Y,Z,ψ) to the BSDE (B.1).*
(ii) For $\forall\lambda\in(0,1)$, there exist an $\mathbb{F}$-progressively measurable continuous process with bounded variation $(V^\lambda_s, s\in[0,T])$ with $V^\lambda_0=0$ and an $\mathbb{F}$-progressively measurable increasing process $(N^\lambda_s, s\in[0,T])$ with $N_0=0$ such that, as a signed measure on $\mathbb{R}_+$,

$$\langle Y_s, \widetilde{f}(s,Y_s,Z_s,\psi_s)\rangle ds \leq |Y_s|^2 dV^\lambda_s + |Y_s|dN^\lambda_s + \lambda(|Z_s|^2 + ||\psi_s||^2_{\mathbb{L}^2(E)})ds \ .$$

(iv) There exists some $p\geq 2$ such that $\mathbb{E}\Big[||e^{V^\lambda}Y||^p_T + \Big(\int_0^T e^{V^\lambda_s}dN^\lambda_s\Big)^p\Big]<\infty$ is satisfied for every $\forall\lambda\in(0,1)$.

Lemma B.1. *Suppose Assumption B.1 hold true. Then, there exists some $\exists\lambda\in(0,1)$ such that the following inequality is satisfied;*

$$\begin{aligned}\mathbb{E}||e^{V^\lambda}Y||^p_T + \mathbb{E}\Big(\int_0^T e^{2V^\lambda_s}|Z_s|^2 ds\Big)^{\frac{p}{2}} + \mathbb{E}\Big(\int_0^T\int_E e^{2V^\lambda_s}|\psi_s(z)|^2\mu(ds,dz)\Big)^{\frac{p}{2}} \\ + \mathbb{E}\Big(\int_0^T\int_E e^{2V^\lambda_s}|\psi_s(z)|^2\nu(dz)ds\Big)^{\frac{p}{2}} \leq C_{p,\lambda}\mathbb{E}\Big[e^{pV^\lambda_T}|\widetilde{\xi}|^p + \Big(\int_0^T e^{V^\lambda_s}dN^\lambda_s\Big)^p\Big] \ ,\end{aligned}$$

where $C_{p,\lambda}$ is a positive constant depending only on p,λ.

Proof. The following proof is an improvement of Proposition 2 of Kruse & Popier (2015) [33] by following the idea of Proposition 6.80 of Pardoux & Rascanu (2014) [44], which yields a slightly sharper a priori estimate for $p\geq 2$.

First step: Introduce a sequence of stopping times with $n\in\mathbb{N}$,

$$\begin{aligned}\tau_n \ := \ & \inf\Big\{t\geq 0; \int_0^t e^{2V^\lambda_s}|Z_s|^2 ds + \int_0^t\int_E e^{2V^\lambda_s}|\psi_s(z)|^2\big(\mu(ds,dz)+\nu(dz)ds\big) \\ & +||e^{V^\lambda}Y||_t + \int_0^t e^{V^\lambda_s}dN^\lambda_s \geq n\Big\}\wedge T.\end{aligned}$$

One obtains by applying Itô formula

$$\begin{aligned}
&|Y_0|^2+\int_0^{\tau_n} e^{2V_s^\lambda}|Z_s|^2ds+\int_0^{\tau_n}\int_E e^{2V_s^\lambda}|\psi_s(z)|^2\mu(ds,dz)\\
&=e^{2V_{\tau_n}^\lambda}|Y_{\tau_n}|^2+\int_0^{\tau_n} e^{2V_s^\lambda}2\Big(\langle Y_s,\widetilde{f}(s,Y_s,Z_s,\psi_s)\rangle ds-|Y_s|^2dV_s^\lambda\Big)\\
&\quad-\int_0^{\tau_n} e^{2V_s^\lambda}2\langle Y_s,Z_sdW_s\rangle-\int_0^{\tau_n} e^{2V_s^\lambda}2\langle Y_{s-},\psi_s(z)\rangle\widetilde{\mu}(ds,dz)\\
&\le e^{2V_{\tau_n}^\lambda}|Y_{\tau_n}|^2+\int_0^{\tau_n} e^{2V_s^\lambda}2\Big(|Y_s|dN_s^\lambda+\lambda(|Z_s|^2+||\psi_s||^2_{\mathbb{L}^2(E)})ds\Big)\\
&\quad-\int_0^{\tau_n} e^{2V_s^\lambda}2\langle Y_s,Z_sdW_s\rangle-\int_0^{\tau_n}\int_E e^{2V_s^\lambda}2\langle Y_{s-},\psi_s(z)\rangle\widetilde{\mu}(ds,dz)\ .
\end{aligned}$$

The BDG (or Davis when $p=2$) inequality yields, with some positive constant C_p depending only on p,

$$\begin{aligned}
&\mathbb{E}\left[\Big(\int_0^{\tau_n} e^{2V_s^\lambda}|Z_s|^2ds\Big)^{\frac{p}{2}}+\Big(\int_0^{\tau_n}\int_E e^{2V_s^\lambda}|\psi_s(z)|^2\mu(ds,dz)\Big)^{\frac{p}{2}}\right]\\
&\le C_p\mathbb{E}\left[||e^{V^\lambda}Y||^p_{\tau_n}+\Big(\int_0^{\tau_n} e^{V_s^\lambda}dN_s^\lambda\Big)^p\right]\\
&\quad+\lambda^{\frac{p}{2}}C_p\mathbb{E}\left[\Big(\int_0^{\tau_n} e^{2V_s^\lambda}|Z_s|^2ds\Big)^{\frac{p}{2}}+\Big(\int_0^{\tau_n} e^{2V_s^\lambda}||\psi_s||_{\mathbb{L}^2(E)}ds\Big)^{\frac{p}{2}}\right]\\
&\quad+C_p\mathbb{E}\left[\Big(\int_0^{\tau_n} e^{4V_s^\lambda}|Y_s|^2|Z_s|^2ds\Big)^{\frac{p}{4}}+\Big(\int_0^{\tau_n}\int_E e^{4V_s^\lambda}|Y_s|^2|\psi_s(z)|^2\mu(ds,dz)\Big)^{\frac{p}{4}}\right]\ .
\end{aligned}$$

With an arbitrary constant $\epsilon>0$, one has

$$\begin{aligned}
C_p\mathbb{E}\left[\Big(\int_0^{\tau_n} e^{4V_s^\lambda}|Y_s|^2|Z_s|^2ds\Big)^{\frac{p}{4}}\right]&\le C_p\mathbb{E}\left[||e^{V^\lambda}Y||^{\frac{p}{2}}_{\tau_n}\Big(\int_0^{\tau_n} e^{2V_s^\lambda}|Z_s|^2ds\Big)^{\frac{p}{4}}\right]\\
&\le\frac{C_p^2}{4\epsilon}\mathbb{E}\Big[||e^{V^\lambda}Y||^p_{\tau_n}\Big]+\epsilon\mathbb{E}\left[\Big(\int_0^{\tau_n} e^{2V_s^\lambda}|Z_s|^2ds\Big)^{\frac{p}{2}}\right]
\end{aligned}$$

and similarly

$$\begin{aligned}
&C_p\mathbb{E}\left[\Big(\int_0^{\tau_n}\int_E e^{4V_s^\lambda}|Y_s|^2|\psi_s(z)|^2\mu(ds,dz)\Big)^{\frac{p}{4}}\right]\\
&\qquad\le\frac{C_p^2}{4\epsilon}\mathbb{E}\Big[||e^{V^\lambda}Y||^p_{\tau_n}\Big]+\epsilon\mathbb{E}\left[\Big(\int_0^{\tau_n}\int_E e^{2V_s^\lambda}|\psi_s(z)|^2\mu(ds,dz)\Big)^{\frac{p}{2}}\right].
\end{aligned}$$

Thus, one obtains

$$\begin{aligned}
&(1-\epsilon-\lambda^{\frac{p}{2}}C_p)\mathbb{E}\Big(\int_0^{\tau_n} e^{2V_s^\lambda}|Z_s|^2ds\Big)^{\frac{p}{2}}+(1-\epsilon)\mathbb{E}\Big(\int_0^{\tau_n}\int_E e^{2V_s^\lambda}|\psi_s(z)|^2\mu(ds,dz)\Big)^{\frac{p}{2}}\\
&-\lambda^{\frac{p}{2}}C_p\mathbb{E}\Big(\int_0^{\tau_n}\int_E e^{2V_s^\lambda}|\psi_s(z)|^2\nu(dz)ds\Big)^{\frac{p}{2}}\le C_p'\mathbb{E}\left[||e^{V^\lambda}Y||^p_{\tau_n}+\Big(\int_0^{\tau_n} e^{V_s^\lambda}dN_s^\lambda\Big)^p\right]\ .
\end{aligned}$$

Firstly, choose some $\epsilon\in(0,1)$. Then, by Lemma A.2, there exists a $\lambda\in(0,1)$ depending only on p so that the 3rd term is absolutely smaller than the 2nd term. Redefining the coefficients and passing to the limit $\tau_n \to T$ yields

$$\mathbb{E}\left[\Big(\int_0^T e^{2V_s^\lambda}|Z_s|^2 ds\Big)^{\frac{p}{2}} + \Big(\int_0^T \int_E e^{2V_s^\lambda}|\psi_s(z)|^2 \mu(ds,dz)\Big)^{\frac{p}{2}}\right]$$
$$+\mathbb{E}\left[\Big(\int_0^T \int_E e^{2V_s^\lambda}|\psi_s(z)|^2 \nu(dz)ds\Big)^{\frac{p}{2}}\right] \leq C_{p,\lambda}\mathbb{E}\left[||e^{V^\lambda}Y||_T^p + \Big(\int_0^T e^{V_s^\lambda} dN_s^\lambda\Big)^p\right]. \quad \text{(B.2)}$$

Second step: Put $\theta(y) := |y|^p$. Then, Itô formula yields

$$d(e^{pV_s^\lambda}|Y_s|^p) = e^{pV_s^\lambda}\Big(p|Y_s|^p dV_s^\lambda + p|Y_{s-}|^{p-2}\langle Y_{s-}, dY_s\rangle + \frac{1}{2}\mathrm{Tr}(\partial_y^2\theta(Y_s)Z_sZ_s^\top)ds\Big)$$
$$+\int_E e^{pV_s^\lambda}\Big(|Y_{s-}+\psi_s(z)|^p - |Y_{s-}|^p - p|Y_{s-}|^{p-2}\langle Y_{s-},\psi_s(z)\rangle\Big)\mu(ds,dz).$$

Using the same sequence of stopping times $(\tau_n)_{n\in\mathbb{N}}$,

$$e^{pV_t^\lambda}|Y_t|^p = e^{pV_{\tau_n}^\lambda}|Y_{\tau_n}|^p + \int_t^{\tau_n} e^{pV_s^\lambda}p|Y_s|^{p-2}\Big(\langle Y_s, \widetilde{f}(s,Y_s,Z_s,\psi_s)\rangle ds - |Y_s|^2 dV_s^\lambda\Big)$$
$$-\int_t^{\tau_n} e^{pV_s^\lambda}\frac{1}{2}\mathrm{Tr}(\partial_y^2\theta(Y_s)Z_sZ_s^\top)ds$$
$$-\int_t^{\tau_n}\int_E e^{pV_s^\lambda}\Big(|Y_{s-}+\psi_s(z)|^p - |Y_{s-}|^p - p|Y_{s-}|^{p-2}\langle Y_{s-},\psi_s(z)\rangle\Big)\mu(ds,dz)$$
$$-\int_t^{\tau_n} e^{pV_s^\lambda}p|Y_s|^{p-2}\langle Y_s, Z_s dW_s\rangle - \int_t^{\tau_n}\int_E e^{pV_s^\lambda}p|Y_{s-}|^{p-2}\langle Y_{s-},\psi_s(z)\rangle\widetilde{\mu}(ds,dz)\ .$$

Let us mention the fact that

$$\mathrm{Tr}(\partial_y^2\theta(Y_s)Z_sZ_s^\top) \geq p|Y_s|^{p-2}|Z_s|^2,$$
$$|Y_{s-}+\psi_s^i(z)|^p - |Y_{s-}|^p - p|Y_{s-}|^{p-2}\langle Y_{s-},\psi_s^i(z)\rangle \geq p(p-1)3^{1-p}|Y_{s-}|^{p-2}|\psi_s^i(z)|^2,$$

for every $i \in \{1,\cdots,k\}$. The latter is obtained by evaluating the residual of Taylor formula [33]. Setting $\kappa_p := \min\Big(\frac{p}{2}, p(p-1)3^{1-p}\Big)$, one obtains

$$e^{pV_t^\lambda}|Y_t|^p + \kappa_p\int_t^{\tau_n} e^{pV_s^\lambda}|Y_s|^{p-2}|Z_s|^2 ds + \kappa_p\int_t^{\tau_n}\int_E e^{pV_s^\lambda}|Y_{s-}|^{p-2}|\psi_s(z)|^2\mu(ds,dz)$$
$$\leq e^{pV_{\tau_n}^\lambda}|Y_{\tau_n}|^p + \int_t^{\tau_n} e^{pV_s^\lambda}p|Y_s|^{p-2}\Big(|Y_s|dN_s^\lambda + \lambda(|Z_s|^2 + ||\psi_s||^2_{\mathbb{L}^2(E)})ds\Big)$$
$$-\int_t^{\tau_n} e^{pV_s^\lambda}p|Y_s|^{p-2}\langle Y_s, Z_s dW_s\rangle - \int_t^{\tau_n}\int_E e^{pV_s^\lambda}p|Y_{s-}|^{p-2}\langle Y_{s-},\psi_s(z)\rangle\widetilde{\mu}(ds,dz). \quad \text{(B.3)}$$

Putting $t = 0$ and taking expectation give

$$\mathbb{E}\left[\kappa_p\int_0^{\tau_n} e^{pV_s^\lambda}|Y_s|^{p-2}|Z_s|^2 ds + \kappa_p\int_0^{\tau_n}\int_E e^{pV_s^\lambda}|Y_{s-}|^{p-2}|\psi_s(z)|^2\mu(ds,dz)\right]$$
$$\leq \mathbb{E}\left[e^{pV_{\tau_n}^\lambda}|Y_{\tau_n}|^p + \int_0^{\tau_n} e^{pV_s^\lambda}p|Y_s|^{p-1}dN_s^\lambda\right] + \lambda\mathbb{E}\left[\int_0^{\tau_n} e^{pV_s^\lambda}p|Y_s|^{p-2}\big(|Z_s|^2 + ||\psi_s||^2_{\mathbb{L}^2(E)}\big)ds\right]\ .$$

By Lemma A.2, one obtains

$$\mathbb{E}\left[\int_0^{\tau_n} e^{pV_s^\lambda}|Y_s|^{p-2}|Z_s|^2ds+\int_0^{\tau_n}\int_E e^{pV_s^\lambda}|Y_{s-}|^{p-2}|\psi_s(z)|^2\mu(ds,dz)\right]$$
$$\leq C_{p,\lambda}\mathbb{E}\left[e^{pV_{\tau_n}^\lambda}|Y_{\tau_n}|^p+\int_0^{\tau_n} e^{pV_s^\lambda}|Y_s|^{p-1}dN_s^\lambda\right] \tag{B.4}$$

by choosing a small $\lambda \in (0,1)$.

Now, applying the Davis inequality (See Chap.I, Sec. 9, Theorem 6 in [37]) to (B.3),

$$\mathbb{E}\Big[||e^{V^\lambda}Y||_{\tau_n}^p\Big]+\mathbb{E}\left[\kappa_p\int_0^{\tau_n} e^{pV_s^\lambda}|Y_s|^{p-2}|Z_s|^2ds+\kappa_p\int_0^{\tau_n}\int_E e^{pV_s^\lambda}|Y_{s-}|^{p-2}|\psi_s(z)|^2\mu(ds,dz)\right]$$
$$\leq \mathbb{E}\left[e^{pV_{\tau_n}^\lambda}|Y_{\tau_n}|^p+\int_0^{\tau_n} e^{pV_s^\lambda}p|Y_s|^{p-1}dN_s^\lambda\right]+\lambda\mathbb{E}\left[\int_0^{\tau_n} e^{pV_s^\lambda}p|Y_s|^{p-2}\big(|Z_s|^2+||\psi_s||^2_{\mathbb{L}^2(E)}\big)ds\right]$$
$$+C\mathbb{E}\Big(\int_0^{\tau_n} e^{2pV_s^\lambda}|Y_s|^{2p-2}|Z_s|^2ds\Big)^{\frac{1}{2}}+C\mathbb{E}\Big(\int_0^{\tau_n}\int_E e^{2pV_s^\lambda}|Y_{s-}|^{2p-2}|\psi_s(z)|^2\mu(ds,dz)\Big)^{\frac{1}{2}},$$

where C is some positive constant. By Lemma A.2, one can choose $\lambda \in (0,1)$ small enough (depending only on p) so that

$$\mathbb{E}\Big[||e^{V^\lambda}Y||_{\tau_n}^p\Big]\leq \mathbb{E}\left[e^{pV_{\tau_n}^\lambda}|Y_{\tau_n}|^p+\int_0^{\tau_n} e^{pV_s^\lambda}p|Y_s|^{p-1}dN_s^\lambda\right]$$
$$+C\mathbb{E}\Big(\int_0^{\tau_n} e^{2pV_s^\lambda}|Y_s|^{2p-2}|Z_s|^2ds\Big)^{\frac{1}{2}}+C\mathbb{E}\Big(\int_0^{\tau_n}\int_E e^{2pV_s^\lambda}|Y_{s-}|^{2p-2}|\psi_s(z)|^2\mu(ds,dz)\Big)^{\frac{1}{2}}.$$

By retaking a smaller λ in the first step if necessary, one can use a common $\lambda \in (0,1)$ both in the first and second steps.

Note that

$$C\mathbb{E}\Big(\int_0^{\tau_n} e^{2pV_s^\lambda}|Y_s|^{2p-2}|Z_s|^2ds\Big)^{\frac{1}{2}}\leq C\mathbb{E}\left[||e^{V^\lambda}Y||_{\tau_n}^{\frac{p}{2}}\Big(\int_0^{\tau_n} e^{pV_s^\lambda}|Y_s|^{p-2}|Z_s|^2ds\Big)^{\frac{1}{2}}\right]$$
$$\leq \epsilon\mathbb{E}\Big[||e^{V^\lambda}Y||_{\tau_n}^p\Big]+\frac{C^2}{4\epsilon}\mathbb{E}\left[\int_0^{\tau_n} e^{pV_s^\lambda}|Y_s|^{p-2}|Z_s|^2ds\right],$$

and similarly

$$C\mathbb{E}\Big(\int_0^{\tau_n}\int_E e^{2pV_s^\lambda}|Y_{s-}|^{2p-2}|\psi_s(z)|^2\mu(ds,dz)\Big)^{\frac{1}{2}}$$
$$\leq \epsilon\mathbb{E}\Big[||e^{V^\lambda}Y||_{\tau_n}^p\Big]+\frac{C^2}{4\epsilon}\mathbb{E}\left[\int_0^{\tau_n}\int_E e^{pV_s^\lambda}|Y_{s-}|^{p-2}|\psi_s(z)|^2\mu(ds,dz)\right].$$

Thus, taking $\epsilon = 1/4$, one obtains

$$\mathbb{E}\Big[||e^{V^\lambda}Y||_{\tau_n}^p\Big]\leq C_p\mathbb{E}\left[e^{pV_{\tau_n}^\lambda}|Y_{\tau_n}|^p+\int_0^{\tau_n} e^{pV_s^\lambda}|Y_s|^{p-1}dN_s^\lambda\right]$$
$$+C_p\mathbb{E}\left[\int_0^{\tau_n} e^{pV_s^\lambda}|Y_s|^{p-2}|Z_s|^2ds+\int_0^{\tau_n}\int_E e^{pV_s^\lambda}|Y_{s-}|^{p-2}|\psi_s(z)|^2\mu(ds,dz)\right].$$

Then the inequality (B.4) implies

$$\mathbb{E}\Big[||e^{V^\lambda}Y||^p_{\tau_n}\Big] \le C_{p,\lambda}\mathbb{E}\left[e^{pV^\lambda_{\tau_n}}|Y_{\tau_n}|^p + \int_0^{\tau_n} e^{pV^\lambda_s}|Y_s|^{p-1}dN^\lambda_s\right] .$$

Passing to the limit $\tau_n \to T$, the monotone convergence in the left and the dominated convergence in the right-hand side give

$$\mathbb{E}\Big[||e^{V^\lambda}Y||^p_T\Big] \le C_{p,\lambda}\mathbb{E}\left[e^{pV^\lambda_T}|\widetilde{\xi}|^p + \int_0^T e^{pV^\lambda_s}|Y_s|^{p-1}dN^\lambda_s\right] .$$

By Young's inequality, for an arbitrary $\epsilon > 0$, one has that

$$\begin{aligned}
&\mathbb{E}\left[\int_0^T e^{pV^\lambda_s}|Y_s|^{p-1}dN^\lambda_s\right] \le \mathbb{E}\left[||e^{V^\lambda}Y||^{p-1}_T \int_0^T e^{V^\lambda_s}dN^\lambda_s\right] \\
&\le \frac{p-1}{p}\epsilon^{\frac{p}{p-1}}\mathbb{E}\Big[||e^{V^\lambda}Y||^p_T\Big] + \frac{1}{p\epsilon^p}\mathbb{E}\left[\Big(\int_0^T e^{V^\lambda_s}dN^\lambda_s\Big)^p\right] .
\end{aligned}$$

Hence, by taking ϵ small, one obtains

$$\mathbb{E}\Big[||e^{V^\lambda}Y||^p_T\Big] \le C_{p,\lambda}\mathbb{E}\left[e^{pV^\lambda_T}|\widetilde{\xi}|^p + \Big(\int_0^T e^{V^\lambda_s}dN^\lambda_s\Big)^p\right] ,$$

Combining with the result (B.2) in *First step*, one obtains the desired result. □

Now, let us introduce the maps $\widetilde{\xi}^i : \Omega \to \mathbb{R}^m$ and $\widetilde{f}^i : \Omega \times [0,T] \times \mathbb{R}^m \times \mathbb{R}^{m\times l} \times \mathbb{L}^2(E,\mathcal{E},\nu;\mathbb{R}^m) \to \mathbb{R}^m$ with $i \in \{1,2\}$.

Assumption B.2. *(i) For $i \in \{1,2\}$, $\widetilde{\xi}^i$ is $\mathcal{F}_T$-measurable and the map $(\omega,t) \mapsto \widetilde{f}^i(\omega,t,\cdot)$ is $\mathbb{F}$-progressively measurable.*
(ii) For every (y,z,ψ), $(y',z',\psi') \in \mathbb{R}^m \times \mathbb{R}^{m\times l} \times \mathbb{L}^2(E,\mathcal{E},\nu;\mathbb{R}^m)$, there exists a positive constant $K > 0$ such that

$$|\widetilde{f}^i(\omega,t,y,z,\psi) - \widetilde{f}^i(\omega,t,y',z',\psi')| \le K\Big(|y-y'| + |z-z'| + ||\psi-\psi'||_{\mathbb{L}^2(E)}\Big)$$

$d\mathbb{P}\otimes dt$-a.e. in $\Omega \times [0,T]$.
(iii) For both $i \in \{1,2\}$, there exists some $p \ge 2$ such that

$$\mathbb{E}\left[|\widetilde{\xi}|^p + \Big(\int_0^T |\widetilde{f}(s,0,0,0)|ds\Big)^p\right] < \infty .$$

Lemma B.2. *(a)Under Assumption B.2, the BSDE*

$$Y^i_t = \widetilde{\xi}^i + \int_t^T \widetilde{f}^i(s,Y^i_s,Z^i_s,\psi^i_s)ds - \int_t^T Z^i_s dW_s - \int_t^T\int_E \psi^i_s(z)\widetilde{\mu}(ds,dz) \tag{B.5}$$

has a unique solution (Y^i,Z^i,ψ^i) which belongs to $\mathbb{S}^p_m[0,T] \times \mathbb{H}^p_{m\times l}[0,T] \times \mathbb{H}^p_{m,\nu}[0,T]$ satisfying the inequality

$$||(Y^i,Z^i,\psi^i)||^p_{\mathcal{K}^p[0,T]} \le C_p\mathbb{E}\left[|\widetilde{\xi}|^p + \Big(\int_0^T |\widetilde{f}^i(s,0,0,0)|ds\Big)^p\right] \tag{B.6}$$

where C_p is some positive constant depending only on (p,K,T). Moreover, if $A^i_2 :=$

$\mathbb{E}\Big[|\widetilde{\xi}^i|^2 + ||\widetilde{f}^i(\cdot,0)||_T^2\Big] < \infty$, *then*

$$\mathbb{E}\Big[\sup_{s\le u\le t}|Y_u^i - Y_s^i|^2\Big] \le C_2\left[A_2^i|t-s|^2 + \Big(\int_s^t |Z_u^i|^2du\Big) + \int_s^t\int_E |\psi_u^i(z)|^2\nu(dz)du\right]. \quad \text{(B.7)}$$

(b) Fix $\widetilde{\xi}^1,\widetilde{\xi}^2 \in \mathbb{L}^p(\Omega,\mathcal{F}_T,\mathbb{P};\mathbb{R}^m)$ *and let* (Y^i,Z^i,ψ^i) *be the solution of (B.5) for* $i\in\{1,2\}$. *Then, for all* $t\in[0,T]$,

$$\mathbb{E}\left[||\delta Y||_{[t,T]}^p + \Big(\int_t^T |\delta Z_s|^2ds\Big)^{p/2} + \Big(\int_t^T\int_E |\delta\psi_s(z)|^2\mu(ds,dz)\Big)^{p/2}\right]$$
$$+\mathbb{E}\left[\Big(\int_t^T\int_E |\delta\psi_s(z)|^2\nu(dz)ds\Big)^{p/2}\right] \le C_p\mathbb{E}\left[|\delta\xi|^p + \Big(\int_t^T |\delta\widetilde{f}_s|ds\Big)^p\right] \quad \text{(B.8)}$$

where $\delta\xi := \widetilde{\xi}^1 - \widetilde{\xi}^2$, $\delta Y := Y^1 - Y^2$, $\delta Z := Z^1 - Z^2$, $\delta\psi := \psi^1 - \psi^2$ *and* $\delta\widetilde{f}_\cdot := (\widetilde{f}^1 - \widetilde{f}^2)(\cdot, Y_\cdot^1, Z_\cdot^1, \psi_\cdot^1)$.

Remark

Note that in [33], the estimates (B.6) and (B.8) are slightly weaker, where the right hand side is given by $\Big(\int_0^T |\widetilde{f}(s,0,0,0)|^pds\Big)$ *instead of* $\Big(\int_0^T |\widetilde{f}(s,0,0,0)|ds\Big)^p$. *This stems from Lemma B.1 and can be crucial if one needs to apply a fixed-point theorem for a short maturity* T.

Proof. Firstly, assume the existence of a solution to (B.5) such that $(Y^i,Z^i,\psi^i)\in\mathcal{K}^p[0,T]$ for both $i\in\{1,2\}$. One has

$$\langle Y_s^i, \widetilde{f}^i(s,Y_s^i,Z_s^i,\psi_s^i)\rangle ds \le |Y_s^i|\Big(|\widetilde{f}^i(s,0)| + K\big(|Y_s^i| + |Z_s^i| + ||\psi_s^i||_{\mathbb{L}^2(E)}\big)\Big)ds$$
$$\le |Y_s^i|^2\Big(K + \frac{K^2}{2\lambda}\Big)ds + |Y_s^i||\widetilde{f}^i(s,0)|ds + \lambda(|Z_s^i|^2 + ||\psi_s^i||_{\mathbb{L}^2(E)}^2)ds$$

for $\forall\lambda>0$. One can easily check that Assumption B.1 is satisfied by choosing

$$V_t^\lambda := \Big(K + \frac{K^2}{2\lambda}\Big)t, \quad N_t^\lambda := \int_0^t |\widetilde{f}^i(s,0)|ds,$$

for $t\in[0,T]$. Thus Lemma B.1 proves the inequality (B.6).

The BDG inequality yields

$$\mathbb{E}\left[\sup_{u\in[s,t]}|Y_u^i - Y_s^i|^2\right] \le C_2\mathbb{E}\left[\Big(\int_s^t |\widetilde{f}^i(r,Y_r^i,Z_r^i,\psi_r^i)|dr\Big)^2 + \int_s^t\Big(|Z_r^i|^2 + ||\psi_r^i||_{\mathbb{L}^2(E)}^2\Big)dr\right]$$

which, together with the estimate (B.6), proves (B.7). For (b), it is easy to check

$$|f^1(s,Y_s^1,Z_s^1,\psi_s^1) - f^2(s,Y_s^2,Z_s^2,\psi_s^2)| \le |\delta f_s| + K\Big(|\delta Y_s| + |\delta Z_s| + ||\delta\psi_s||_{\mathbb{L}^2(E)}\Big).$$

Thus, Assumption B.1 is satisfied once again for $(\delta Y,\delta Z,\delta\psi)$ by choosing

$$V_t^\lambda := \Big(K + \frac{K^2}{2\lambda}\Big)t, \quad N_t^\lambda := \int_0^t |\delta f(s)|ds\ .$$

Therefore, the estimate (B.8) immediately follows from Lemma B.1.

Now, let us prove the existence in (a). The uniqueness is already proved by (b). The following is a simple modification of Theorem 5.17 [44] given for a diffusion setup. Consider a sequence of BSDEs (the superscript $i \in \{1,2\}$ is omitted), for $n \in \mathbb{N}$,

$$Y_t^{n+1} = \widetilde{\xi} + \int_t^T \widetilde{f}(s, Y_s^n, Z_s^n, \psi_s^n)ds - \int_t^T Z_s^{n+1} dW_s - \int_t^T \int_E \psi_s^{n+1}(z)\widetilde{\mu}(ds, dz) \ .$$

Suppose that $(Y^n, Z^n, \psi^n) \in \mathcal{K}^p[0,T]$. Then, from the linear growth property, it is obvious that

$$\widetilde{\xi} + \int_t^T \widetilde{f}(s, Y_s^n, Z_s^n, \psi_s^n)ds \in \mathbb{L}^p(\Omega, \mathcal{F}_T, \mathbb{P}; \mathbb{R}^m) \ .$$

Thus the martingale representation theorem (see, for example, Theorem 5.3.6 in [1]) implies that there exists a unique solution $(Y^{n+1}, Z^{n+1}, \psi^{n+1}) \in \mathcal{K}^p[0,T]$. Let us define this map as $(Y^{n+1}, Z^{n+1}, \psi^{n+1}) = \Phi(Y^n, Z^n, \psi^n)$. Denote $(\delta Y^n, \delta Z^n, \delta\psi^n) := (Y^n - Y^{n-1}, Z^n - Z^{n-1}, \psi^n - \psi^{n-1})$. Then (B.8) (with a zero Lipschitz constant) implies

$$\begin{aligned}
&||(\delta Y^{n+1}, \delta Z^{n+1}, \delta\psi^{n+1})||^p_{\mathcal{K}^p[0,T]} \\
&\leq C_p \mathbb{E}\Big[\Big(\int_0^T |\widetilde{f}(s, Y_s^n, Z_s^n, \psi^n) - \widetilde{f}(s, Y_s^{n-1}, Z_s^{n-1}, \psi_s^{n-1})|ds\Big)^p\Big] \\
&\leq C_p' \mathbb{E}\Big[\Big(\int_0^T \big[|\delta Y_s^n| + |\delta Z_s^n| + ||\delta\psi_s^n||_{\mathbb{L}^2(E)}\big]ds\Big)^p\Big] \\
&\leq C_p' \max(T^p, T^{\frac{p}{2}})||(\delta Y^n, \delta Z^n, \delta\psi^n)||^p_{\mathcal{K}^p[0,T]} \ . \qquad \text{(B.9)}
\end{aligned}$$

Note in particular that C_p' is independent of the terminal condition. Thus, if the terminal time T is small enough so that $\alpha := C_p' \max(T^p, T^{\frac{p}{2}}) < 1$, then the map Φ is strictly contracting. In this case, by the fixed point theorem in the Banach space, there exists a solution $(Y, Z, \psi) \in \mathcal{K}^p[0,T]$ to the BSDE (B.5). For general T, one can consider a time partition $0 = T_0 < T_1 < \cdots < T_N = T$. By taking $[T_{N-1}, T]$ small enough, the above arguments guarantee that there exists a solution $(Y, Z, \psi) \in \mathcal{K}^p[T_{N-1}, T]$. By the uniqueness of the solution, one can repeat the same procedures for the interval $[T_{N-2}, T_{N-1}]$ with the new terminal value $Y_{T_{N-1}}$. Repeating N times, one proves the desired result. □

The following lemma is useful when one deals with the jumps of finite measure.

Lemma B.3. *Suppose $\nu^i(\mathbb{R}_0) < \infty$ for every $1 \leq i \leq k$. Given $\psi \in \mathbb{H}^2_\nu[0,T]$, let M be defined by $M_t := \int_0^t \int_E \psi_s(z)\widetilde{\mu}(ds, dz)$ on $[0,T]$. Then, for $\forall p \geq 2$, $k_p||\psi||^p_{\mathbb{H}^p_\nu[0,T]} \leq ||M||^p_{\mathbb{S}^p[0,T]} \leq K_p||\psi||^p_{\mathbb{H}^p_\nu[0,T]}$, where k_p, K_p are positive constant depend only on $p, \nu(E)$ and T.*

Proof. See pp.125 of [16], for example. □

C Smooth approximation theorem

In the reminder of the paper, we provide a justification to use smooth coefficients in the forward-backward SDEs for any numerical approximation purpose. Since ϵ is a perturbation parameter, we can always introduce it so that all the functions depend smoothly on ϵ. This is actually the case for the examples used in Sections 5.1 and 5.2. Thus we concentrate on the other parameters and omit ϵ dependence from the functions in the following. Let us first consider the forward component:

$$\widetilde{X}_s = x + \int_t^s \widetilde{b}(r, \widetilde{X}_r)dr + \int_t^s \widetilde{\sigma}(r, \widetilde{X}_r)dW_r + \int_t^s \int_E \widetilde{\gamma}(r, \widetilde{X}_r, z)\widetilde{\mu}(dr, dz) \ , \tag{C.1}$$

where $x \in \mathbb{R}^d$ and $\widetilde{b} : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$, $\widetilde{\sigma} : [0,T] \times \mathbb{R}^d \to \mathbb{R}^{d\times l}$, $\widetilde{\gamma} : [0,T] \times \mathbb{R}^d \times E \to \mathbb{R}^{d\times k}$ are measurable functions. We omit the superscripts denoting the initial data (t, x).

Assumption C.1. *$\widetilde{b}, \widetilde{\sigma}, \widetilde{\gamma}$ are continuous in (t, x, z). There exists some positive constant K such that, for every $x, x' \in \mathbb{R}^d$,*
(i) $|\widetilde{b}(t, x) - \widetilde{b}(t, x')| + |\widetilde{\sigma}(t, x) - \widetilde{\sigma}(t, x')| \leq K|x - x'|$ uniformly in $t \in [0,T]$,
(ii) $|\widetilde{\gamma}_j(t, x, z) - \widetilde{\gamma}_j(t, x', z)| \leq K\eta(z)|x - x'|$ for $1 \leq j \leq k$ uniformly in $(t, z) \in [0,T] \times \mathbb{R}_0$,
(iii) $||\widetilde{b}(\cdot, 0)||_T + ||\widetilde{\sigma}(\cdot, 0)||_T + ||\widetilde{\gamma}(\cdot, 0, z)||_T/\eta(z) \leq K$ uniformly in $z \in E$.

The regularization technique by the convolution with appropriate mollifiers gives us the following approximating functions.

Lemma C.1. *Under Assumption C.1, one can choose a sequence of functions $b_n : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$, $\sigma_n : [0,T] \times \mathbb{R}^d \to \mathbb{R}^{d\times l}$, $\gamma_n : [0,T] \times \mathbb{R}^d \times E \to \mathbb{R}^{d\times k}$ with $n \in \mathbb{N}$, which are continuous in all their arguments, infinitely differentiable in x with continuous derivatives, and also satisfy, for each $n \geq 1$;*
(i) for every $m \geq 1$, $|\partial_x^m b_n(t, x)| + |\partial_x^m \sigma_n(t, x)| + |\partial_x^m \gamma_n(t, x, z)|/\eta(z)$ is uniformly bounded in $(t, x, z) \in [0,T] \times \mathbb{R}^d \times E$,
(ii) for every $(t, x, z) \in [0,T] \times \mathbb{R}^d \times E$, $b_n(t, x), \sigma_n(t, x)$ and $\gamma_n(t, x, z)$ converge pointwise to $\widetilde{b}(t, x)$, $\widetilde{\sigma}(t, x)$ and $\widetilde{\gamma}(t, x, z)$, respectively,
(iii) $(b_n, \sigma_n, \gamma_n)$ satisfy the properties in Assumption C.1 with some positive constant K' independent of n.

Proof. We consider a sequence of (symmetric) mollifiers $\varrho_n \in \mathcal{C}_0^\infty : \mathbb{R}^d \to \mathbb{R}_+$ with compact support satisfying $\int_{\mathbb{R}^d} \varrho_n(x)dx = 1$ and $\varrho_n(x) \to \delta(x)$ as $n \to \infty$ in the space of Schwartz distributions, where $\delta(\cdot)$ is a Dirac delta function. Let us define intermediate mollified functions as

$$\bar{b}_n(t, x) := \varrho_n * \widetilde{b}(t, x), \quad \bar{\sigma}_n(t, x) := \varrho_n * \widetilde{\sigma}(t, x), \quad \bar{\gamma}_n(t, x, z) := \varrho_n * \widetilde{\gamma}(t, x, z)$$

where $*$ denotes a convolution with respect to x, such as

$$\bar{b}_n(t, x) = \int_{\mathbb{R}^d} \varrho_n(x - y)\widetilde{b}(t, y)dy = \int_{\mathbb{R}^d} \widetilde{b}(t, x - y)\varrho_n(y)dy \ .$$

Since $\widetilde{b}, \widetilde{\sigma}, \widetilde{\gamma}$ are continuous, every point $x \in \mathbb{R}^d$ is a Lebesgue point. Thus, the approximated functions $\bar{b}_n, \bar{\sigma}_n, \bar{\gamma}_n$ are known to converge pointwise to $\widetilde{b}, \widetilde{\sigma}, \widetilde{\gamma}$ from the Lebesgue differentiation theorem (see, for example, Theorem 8.7 in Igari (1996) [31] or Theorem C.19 in Leoni (2009) [35]). The Lipschitz property can be shown as, for every $x, x' \in \mathbb{R}^d$,

$$\begin{aligned} |\bar{\gamma}_{n,j}(t, x, z) - \bar{\gamma}_{n,j}(t, x', z)| &\leq \int_{\mathbb{R}^d} |\widetilde{\gamma}_j(t, x - y, z) - \widetilde{\gamma}_j(t, x' - y, z)|\varrho_n(y)dy \\ &\leq K\eta(z)|x - x'| \int_{\mathbb{R}^d} \varrho_n(y)dy = K|x - x'|\eta(z) \end{aligned}$$

and similarly for the others. It is easy to see that there exists some positive constant C' satisfying

$$||\bar{b}_n(\cdot,0)||_T + ||\bar{\sigma}_n(\cdot,0)||_T + ||\bar{\gamma}_n(\cdot,0,z)||_T/\eta(z) \leq C'$$

uniformly in $z \in E$ as well as $n \in \mathbb{N}$ since ϱ_n has a compact support shrinking to the origin as $n \to \infty$. We prepare another (symmetric) mollifiers $\varsigma_n \in \mathcal{C}_0^\infty : \mathbb{R}^d \times E \to \mathbb{R}_+$ in the following way:

$$\varsigma_n(x,z) = \begin{cases} 1 & \text{for } |x|+|z| \leq n \\ 0 & \text{for } |x|+|z| \geq 2n \end{cases} . \tag{C.2}$$

We then define the mollified functions as

$$b_n(t,x) := \varsigma_n(x,0)\bar{b}_n(t,x), \quad \sigma_n(t,x) := \varsigma_n(x,0)\bar{\sigma}_n(t,x), \quad \gamma_n(t,x,z) := \varsigma_n(x,z)\bar{\gamma}_n(t,x,z) .$$

Since they are smooth in x and have compact supports, they have bounded derivatives of all orders with respect to x uniformly in (t,x,z) for each n. The pointwise convergence is clearly preserved. Lastly, one has to check that there exists a Lipschitz constant K' independent of n. By the construction in (C.2), one can arrange the mollifier in the following way: there exists a positive constant C such that

$$\sup_{(x,z)\in\mathbb{R}^d\times E} \Big|\partial_x \varsigma_n(x,z)\Big| \leq C/n$$

for every $n \in \mathbb{N}$. Then, for $\forall n \in \mathbb{N}$, one sees

$$\begin{aligned} |\partial_x \gamma_n(t,x,z)| &\leq \varsigma_n(x,z)|\partial_x \bar{\gamma}_n(t,x,z)| + |\partial_x \varsigma_n(x,z)||\bar{\gamma}_n(t,x,z)| \\ &\leq K\eta(z) + \eta(z)C/n(C' + K(2n)) \leq K'\eta(z) \end{aligned}$$

uniformly in (t,x,z). Here, we have used the fact that $\partial_x \varsigma_n(x,z)$ vanishes when $|x| \geq 2n$ and the linear growth property of $\bar{\gamma}_n$. One can similarly check $|\partial_x b_n(t,x)|, |\partial_x \sigma_n(t,x)| \leq K'$ for $\forall n \in \mathbb{N}$. The property (iii) of Assumption C.1 is obviously preserved in the second mollification. □

This yields the following result.

Theorem C.1. *Under Assumption C.1, consider the process $\widetilde{X}$ of (C.1) and the sequence of processes $(X_s^n, s \in [t,T])_{n\geq 1}$ defined by*

$$X_s^n = x + \int_t^s b_n(r,X_r^n)dr + \int_t^s \sigma_n(r,X_r^n)dW_r + \int_t^s \int_E \gamma_n(r,X_r^n,z)\widetilde{\mu}(dr,dz) \tag{C.3}$$

with b_n, σ_n and γ_n given in Lemma C.1. Then, there exist unique solutions $\widetilde{X}, X^n$ in $\mathbb{S}^p[t,T]$ $\forall p \geq 2$. Moreover, the following relation holds

$$\lim_{n\to\infty} \mathbb{E}\Big[||\widetilde{X} - X^n||_{[t,T]}^p\Big] = 0$$

for $\forall p \geq 2$.

Proof. The existence of the unique solution for (C.1) as well as (C.3) in $\mathbb{S}^p$ for $\forall p \geq 2$ is clear from Lemma A.3. We also have, for $\forall p \geq 2$,

$$||\widetilde{X} - X^n||_{\mathbb{S}^p}^p \leq C_p \mathbb{E}\Big[\Big(\int_t^T |\delta\widetilde{b}_n(r, \widetilde{X}_r)|dr\Big)^p + \Big(\int_t^T |\delta\widetilde{\sigma}_n(r, \widetilde{X}_r)|^2 dr\Big)^{p/2} + \int_t^T |\delta L_r^n|^p dr\Big]$$

where $\delta\widetilde{b}_n := \widetilde{b} - b_n$, $\delta\widetilde{\sigma}_n := \widetilde{\sigma} - \sigma_n$. Furthermore δL^n is a predictable process satisfying $|\delta\widetilde{\gamma}_n|(t, \widetilde{X}_{t-}, z) \leq \delta L_t^n \eta(z)$, $d\mathbb{P} \otimes dt$-a.e. in $\Omega \times [0, T]$, where $\delta\widetilde{\gamma}_n := \widetilde{\gamma} - \gamma_n$. We can take δL^n such that $\int_t^T \mathbb{E}|\delta L_r^n|^p dr < \infty$, since we have $|\delta\widetilde{\gamma}_n|(s, \widetilde{X}_{s-}, z) \leq 2K(1 + |\widetilde{X}_{s-}|)\eta(z)$ in the current setup. See also the related discussion in Lemma A.3.

Note that C_p is independent of n thanks to Lemma C.1 (iii). Due to the linear growth property, the inside of the expectation is dominated by $C(1 + ||\widetilde{X}||_{[t,T]}^p)$ with some positive constant C independent of n. From Lemma C.1 (ii), $(\delta\widetilde{b}_n, \delta\widetilde{\sigma}_n, \delta\widetilde{\gamma}_n)$ converge pointwise to zero. Thus, one can also take a sequence of $(\delta L^n, n \in \mathbb{N})$ converging pointwise to zero. Since $\widetilde{X} \in \mathbb{S}^p$ for $\forall p \geq 2$, the dominated convergence theorem give the desired result in the limit $n \to \infty$. [5] □

The above result implies that by choosing a large enough n one can work on X^n that is an arbitrary accurate approximation in the $\mathbb{S}^p$ sense of the original process $\widetilde{X}$, and involves only smooth coefficients $(b_n, \sigma_n, \gamma_n)$. This conclusion can be extended to the forward-backward system. Consider the BSDE driven by $\widetilde{X}$;

$$\begin{aligned}\widetilde{Y}_s = \widetilde{\xi}(\widetilde{X}_T) + \int_s^T \widetilde{f}\Big(r, \widetilde{X}_r, \widetilde{Y}_r, \widetilde{Z}_r, \int_{\mathbb{R}_0} \rho(z)\widetilde{\psi}_r(z)\nu(dz)\Big)dr \\ - \int_s^T \widetilde{Z}_r dW_r - \int_s^T \int_E \widetilde{\psi}_r(z)\widetilde{\mu}(dr, dz) \qquad \text{(C.4)}\end{aligned}$$

for $s \in [t, T]$ where $\widetilde{\xi} : \mathbb{R}^d \to \mathbb{R}^m$, $\widetilde{f} : [0, T] \times \mathbb{R}^d \times \mathbb{R}^m \times \mathbb{R}^{m\times l} \times \mathbb{R}^{m\times k} \to \mathbb{R}^m$ are measurable functions and ρ is defined as before.

Assumption C.2. *The functions $\widetilde{\xi}$ and $\widetilde{f}$ are continuous in all their arguments. There exist some positive constants $K, q \geq 0$ such that*
(i)$|\widetilde{\xi}(x)| + |\widetilde{f}(t, x, 0, 0, 0)| \leq K(1 + |x|^q)$ for every $x \in \mathbb{R}^d$ uniformly in $t \in [0, T]$.
(ii) $|\widetilde{f}(t, x, y, z, u) - \widetilde{f}(t, x, y', z', u')| \leq K(|y - y'| + |z - z'| + |u - u'|)$ for every $(y, z, u), (y', z', u') \in \mathbb{R}^m \times \mathbb{R}^{m\times l} \times \mathbb{R}^{m\times k}$ uniformly in $(t, x) \in [0, T] \times \mathbb{R}^d$.

Lemma C.2. *Under Assumption C.2, one can choose a sequence of functions $\xi_n : \mathbb{R}^d \to \mathbb{R}^m$, $f_n : [0, T] \times \mathbb{R}^d \times \mathbb{R}^m \times \mathbb{R}^{m\times l} \times \mathbb{R}^{m\times k} \to \mathbb{R}^m$ with $n \in \mathbb{N}$, which are continuous in all their arguments, infinitely differentiable in (x, y, z, u) with continuous derivatives, and also satisfy, for each $n \geq 1$;*
(i) for every $i \geq 1$, all the ith order partial derivatives of (ξ_n, f_n) are uniformly bounded in $(t, x, y, z, u) \in [0, T] \times \mathbb{R}^d \times \mathbb{R}^m \times \mathbb{R}^{m\times l} \times \mathbb{R}^{m\times k}$,
(ii) for every $(t, x, y, z, u) \in [0, T] \times \mathbb{R}^d \times \mathbb{R}^m \times \mathbb{R}^{m\times l} \times \mathbb{R}^{m\times k}$, ξ_n and f_n converge pointwise to $\widetilde{\xi}$ and $\widetilde{f}$, respectively,
(iii) (ξ_n, f_n) satisfy Assumption C.2 with some positive constant K'' independent of n.

Proof. The first step of the mollification can be done exactly the same way as in Lemma C.1, which gives us $\bar{\xi}_n(x)$ and $\bar{f}_n(t, x, \hat{\Theta}) := \bar{f}_n(t, x, y, z, u)$. In order to achieve the property

[5] In $p = 2$, one can see more directly $||\widetilde{X} - X^n||_{\mathbb{S}^2}^2 \to 0$ since the integral of δL^n can be replaced by that of $\delta\widetilde{\gamma}_n$ (See a remark below Lemma A.3.). Taking an appropriate subsequence if necessary, one can also show that $(X_s^n, s \in [t, T])_{n\geq 1}$ is almost surely uniformly convergent to $(\widetilde{X}_s, s \in [t, T])$ by the Borel-Cantelli lemma.

(iii), one has to take care of the polynomial growth property of the driver with respect to x. One can take the second sequence of mollifiers as

$$\varsigma_n(x,\hat{\Theta}) = \begin{cases} 1 & \text{for } |x|^q + |\hat{\Theta}| \le n \\ 0 & \text{for } |x|^q + |\hat{\Theta}| \ge 2n \end{cases}$$

and then control their first derivatives, with some positive constant C, by

$$\sup_{(x,\hat{\Theta})\in\mathbb{R}^d\times\mathbb{R}^{m(1+l+k)}} |\partial_{\hat{\Theta}}\varsigma_n(x,\hat{\Theta})| \le C/n$$

for $\forall n \in \mathbb{N}$. Then, one can check that

$$\xi_n(x) := \varsigma_n(x,0)\bar{\xi}_n(x), \quad f_n(t,x,\hat{\Theta}) := \varsigma_n(x,\hat{\Theta})\bar{f}_n(t,x,\hat{\Theta})$$

satisfy the desired property similarly as in Lemma C.1. □

Finally, we obtain the main approximation theorem.

Theorem C.2. *Under Assumptions C.1 and C.2, consider the process* $(\widetilde{Y},\widetilde{Z},\widetilde{\psi})$ *of (C.4) and the sequence of processes* $(Y_s^{n,m},Z_s^{n,m},\psi_s^{n,m})_{s\in[t,T]}, (n,m) \in \{1,2,\cdots\}$ *defined as the solution to following BSDE*

$$\begin{aligned} Y_s^{n,m} &= \xi_m(X_T^n) + \int_s^T f_m\Big(r, X_r^n, Y_r^{n,m}, Z_r^{n,m}, \int_{\mathbb{R}_0} \rho(z)\psi_r^{n,m}(z)\nu(dz)\Big)dr \\ &\quad - \int_s^T Z_r^{n,m}dW_r - \int_s^T\int_E \psi_r^{n,m}(z)\widetilde{\mu}(dr,dz) \end{aligned} \tag{C.5}$$

where X^n *is the solution of (C.3),* (ξ_m, f_m) *are the mollified functions given in Lemma C.2. Then, there exist unique solutions* $(\widetilde{Y},\widetilde{Z},\widetilde{\psi})$, $(Y^{n,m},Z^{n,m},\psi^{n,m})_{n,m\ge1} \in \mathcal{K}^p[t,T]$ $\forall p \ge 2$. *Moreover, the following relation holds*

$$\lim_{m\to\infty}\lim_{n\to\infty}\Big|\Big|(\delta Y^{n,m},\delta Z^{n,m},\delta\psi^{n,m})\Big|\Big|_{\mathcal{K}^p[t,T]} = 0 \quad \forall p \ge 2$$

where $\delta Y^{n,m} := \widetilde{Y} - Y^{n,m}$, $\delta Z^{n,m} := \widetilde{Z} - Z^{n,m}$ *and* $\delta\psi^{n,m} := \widetilde{\psi} - \psi^{n,m}$.

Proof. The existence of the unique solution $(\widetilde{Y},\widetilde{Z},\widetilde{\psi})$ and $(Y^{n,m},Z^{n,m},\psi^{n,m})$ in $\mathcal{K}^p$ for $\forall p \ge 2$ is clear from Lemma B.2. We have, for $\forall p \ge 2$,

$$\Big|\Big|(\delta Y^{n,m},\delta Z^{n,m},\delta\psi^{n,m})\Big|\Big|^p_{\mathcal{K}^p[t,T]} \le C_p\mathbb{E}\Big[|\delta\xi^{n,m}|^p + \Big(\int_t^T |\delta f^{n,m}(r)|dr\Big)^p\Big]$$

by the stability result, where $\delta\xi^{n,m} := \widetilde{\xi}(\widetilde{X}_T) - \xi_m(X_T^n)$ and

$$\begin{aligned} \delta f^{n,m}(r) &:= \widetilde{f}\Big(r, \widetilde{X}_r, \widetilde{Y}_r, \widetilde{Z}_r, \int_{\mathbb{R}_0} \rho(z)\widetilde{\psi}_r(z)\nu(dz)\Big) \\ &\quad - f_m\Big(r, X_r^n, \widetilde{Y}_r, \widetilde{Z}_r, \int_{\mathbb{R}_0} \rho(z)\widetilde{\psi}_r(z)\nu(dz)\Big) . \end{aligned}$$

Firstly, let us fix m. Since $\partial_x \xi_m$ and $\partial_x f_m$ are bounded, the result of Theorem C.1 yields

$$\lim_{n\to\infty}\Big|\Big|(\delta Y^{n,m}, \delta Z^{n,m}, \delta\psi^{n,m})\Big|\Big|^p_{\mathcal{K}^p[t,T]} \le C_p \mathbb{E}\left[|\delta\xi^m|^p + \Big(\int_t^T |\delta f^m(r, \widetilde{\Theta}_r)|dr\Big)^p\right]$$

with $\delta\xi^m := \widetilde{\xi}(\widetilde{X}_T) - \xi_m(\widetilde{X}_T)$ and $\delta f^m(r, \widetilde{\Theta}_r) := (\widetilde{f} - f_m)\Big(r, \widetilde{X}_r, \widetilde{Y}_r, \widetilde{Z}_r, \int_{\mathbb{R}_0} \rho(z)\widetilde{\psi}_r(z)\nu(dz)\Big)$. Since $\widetilde{\Theta} \in \mathbb{S}^p \times \mathcal{K}^p$ for $\forall p \ge 2$ and $(\widetilde{f}, f_m)$ have the linear growth in (y, z, u) and the polynomial growth in x with proportional coefficients independent of m, passing to the limit $m \to \infty$ yields the desired result from the pointwise convergence of the mollified functions and the dominated convergence theorem. Notice also that one can achieve the same convergence with the flipped order of limits $\lim_{n\to\infty}\lim_{m\to\infty}$ by using the fact that $(X^n_s, s \in [t,T])_{n\in\mathbb{N}}$ is almost surely uniformly convergent to $(\widetilde{X}_s, s \in [t,T])$ by taking an appropriate subsequence if necessary. □

Theorems C.1 and C.2 imply that one can work on the process Θ^n defined by the smooth coefficients $(b_n, \sigma_n, \gamma_n, \xi_n, f_n)$ as an arbitrary accurate approximation in the $\mathbb{S}^p \times \mathcal{K}^p$ sense of the original one $\widetilde{\Theta}$, which only satisfies Assumptions C.1 and C.2. In fact, we can weaken the assumptions further. There is no difficulty to add discontinuities to $\widetilde{\xi}$ and $\widetilde{f}$ with respect to x as long as they are all Lebesgue points. If we only assume, in addition to the polynomial growth condition, that $(\widetilde{\xi}, \widetilde{f})$ is Borel measurable, then (ξ_m, f_m) converges to $(\widetilde{\xi}, \widetilde{f})$ only dx-a.e. (and hence $(\widetilde{\xi}, \widetilde{f})$ does not have Lebesgue points everywhere) in general. As long as the forward process $\widetilde{X}$ has no mass on this null set in dx, the same conclusion will hold.

Acknowledgement

The research is partially supported by Center for Advanced Research in Finance (CARF) and JSPS KAKENHI (Grant Number 25380389).

References

[1] Applebaum, D., 2009, *Lévy Processes and Stochastic Calculus (2nd edition)*, Cambridge studies in advanced mathematics, Cambridge University Press, LN.

[2] Barles, G., Buckdahn, R. and Pardoux, E., 1997, *Backward stochastic differential equations and integral-partial differential equations*, Stochastics and Stochastic Reports, Vol. 60, 57-83.

[3] Bender, C. and Denk, R., 2007, *A forward scheme for backward SDE*, Stochastic Processes and their Applications, 117, 12, 1793-1832.

[4] Bichteler, K., Gravereaux, J. and Jacod, J., 1987, *Malliavin calculus for processes with jumps*, Stochastics Monographs, Gordon and Breach science publishers, LN.

[5] Bismut, J.M., 1973, *Conjugate convex functions in optimal stochastic control*, J. Math. Anal. Apl. 44, 384-404.

[6] Bouchard, B. and Elie, R., 2008, *Discrete-time approximation of decoupled Forward-Backward SDE with Jumps*, Stochastic Processes and their Applications, 118, 53-75.

[7] Bouchard, B. and Touzi, N., 2004, *Discrete-time approximation and Monte-Carlo simulation of backward stochastic differential equations*, Stochastic Processes and their Applications, 111, 2, 175-206.

[8] Crépey, S. and Bielecki, T.R. with an introductory dialogue by Damiano Brigo, 2014, *Counterparty Risk and Funding*, CRC press, Taylor & Francis Group, NY.

[9] Crepey, S. and Matoussi, A., 2008, *Reflected and doubly reflected BSDEs with jumps: A priori estimates and comparison*, The Annals of Applied Probability, 18, 5, 2041-2069.

[10] Crépey, S. and Song, S., 2015, *Counterparty Risk and Funding: Immersion and Beyond*, Working paper.

[11] Cvitanić, J. and Zhang, J., 2013, *Contract theory in continuous-time methods*, Springer,Berlin.

[12] Delong, L., 2013, *Backward Stochastic Differential Equations with Jumps and Their Actuarial and Financial Applications*, Springer-Verlag, LN.

[13] Delong, L. and Imkeller, P., 2010, *On Malliavin's differentiability of BSDEs with time delayed generators driven by Brownian motions and Poisson random measures*, Stochastic Processes and their Applications, 120, 9, 1748-1775.

[14] Di Nunno, G., Oksendal, B. and Proske, F., 2009, *Malliavin Calculus for Lévy Processes with Applications to Finance*, Springer, NY.

[15] Dzhaparidze, K. and Valkeila, E., 1990, *On the Hellinger type distances for filtered experiments*, Probab. Theory Related Fields, 85, 1, 105-117.

[16] Elie, R., 2006, *Contrôle stochastique et méthodes numériques en finance mathématique*, Ph.D Thesis, University Paris-Dauphine.

[17] El Karoui, N. and Mazliak, L. (eds.), 1997, *Backward stochastic differential equations*, Addison Wesley Longman Limited, U.S..

[18] El Karoui, N., Peng, S. and Quenez, M.-C., 1997, *Backward stochastic differential equations in finance*, Mathematical Finance, 7(1), 1-71.

[19] Fujii, M., 2014, *Momentum-space approach to asymptotic expansion for stochastic filtering*, Annals of the Institute of Statistical Mathematics, Vol. 66, 93-120.

[20] Fujii, M., 2015, *A polynomial scheme of asymptotic expansion for backward SDEs and option pricing*, Quantitative Finance, in press.

[21] Fujii, M., 2015, *Optimal position management for a market-maker with stochastic price impacts*, Working paper, CARF-F-360.

[22] Fujii, M. and Takahashi, A., 2012, *Analytical approximation for non-linear FBSDEs with perturbation scheme*, International Journal of Theoretical and Applied Finance, 15, 5, 1250034 (24).

[23] Fujii, M. and Takahashi, A., 2015b, *Perturbative Expansion Technique for Non-linear FBSDEs with Interacting Particle Method*, Asia-Pacific Financial Markets, 22, 3, 283-304.

[24] Fujii, M. and Takahashi, A., 2015, *Solving Backward Stochastic Differential Equations by Connecting the Short-term Expansions*, Working paper, CARF-F-387.

[25] Fujii, M. and Takahashi, A., 2017, *Quadratic-exponential growth BSDEs with Jumps and their Malliavin's Differentiability*, Stochastic processes and their applications, in press. https://doi.org/10.1016/j.spa.2017.09.002.

[26] Gikhman, I.I. and Skorohod, A.V., 1972, *Stochastic differential equations*, Springer, Berlin.

[27] Gobet, E., Lemor, J.P. and Warin, X., 2005, *A regression-based Monte Carlo method to solve backward stochastic differential equations*, The Annals of Applied Probability, 15, 3, 2172-2202.

[28] Hamadène, S. and Jeanblanc, M., 2007, *On the stating and stopping problem: Application in reversible investments*, Mathematics of Operations Research, Vol. 32, No. 1, pp. 182-192.

[29] Henry-Labordère, P., 2012, *Cutting CVA's complexity*, Risk Magazine, July Issue, 67-73.

[30] He, S., Wang, J. and Yan, J., *Semimartingale Theory and Stochastic Calculus*, CRC Press, London, UK.

[31] Igari, S., 1996, *Introduction of Real Analysis*, Iwanami publisher, in Japanese.

[32] Kazamaki, N., 1979, *A sufficient condition for the uniform integrability of exponential martingales*, Math. Rep. Toyama Univ. 2, 1-11. MR-0542374.

[33] Kruse, T. and Popier, T., 2015, *BSDEs with monotone generator driven by Brownian and Poisson noises in a general filtration*, Stochastics, 88(4):491-539, 2016

[34] Lejay, A., Mordecki, E. and Torres, S., 2014, *Numerical approximation of backward stochastic differential equations with jumps*, HAL: inria-00357992.

[35] Leoni, G., 2009, *A First Course in Sobolev Spaces*, Graduate Studies in Mathematics, Vol. 105, American Mathematical Society, US.

[36] Lim, A.E., *Quadratic Hedging and Mean-Variance portfolio selection with random parameters in an incomplete market*, Mathematics of Operations Research, Vol. 29, No. 1, 132-161.

[37] Liptser, R.Sh. and Shiryayev, A.N., 1989, *Theory of Martingales*, Mathematics and Its Applications, Kluwer Academic Publishers, LN.

[38] Ma, J. and Yong, J., 2000, *Forward-backward stochastic differential equations and their applications*,Springer, Berlin.

[39] Ma, J. and Zhang, J., 2002, *Representation theorems for backward stochastic differential equations*, The annals of applied probability, 12, 4, 1390-1418.

[40] Mania, M. and Tevzadze, R., 2003, *Backward stochastic PDE and imperfect hedging*, International Journal of Theoretical and Applied Finance, Vol. 6, 7, 663-692.

[41] Morlais, M-A., 2010, *A new existence result for quadratic BSDEs with jumps with application to the utility maximization problem*, Stochastic processes and their applications, 120, 1966-1995.

[42] Pardoux, E. and Peng, S., 1990, *Adapted solution of a backward stochastic differential equations*, Systems Control Lett., 14, 55-61.

[43] Pardoux, E. and Peng, S., 1992, *Backward stochastic differential equations and quasilinear parabolic partial differential equations*, Lecture Notes in Control and Inform. Sci. 176, 200-217, Springer, New York.

[44] Pardoux, E. and Rascanu, A., 2014, *Stochastic Differential Equations, Backward SDEs, Partial Differential Equations*, Springer International Publishing, Switzerland.

[45] Petrou, E., 2008, *Malliavin calculus in Lévy spaces and applications to finance*, Electric Journal of Probability, 13, 27, 852-879.

[46] Pham, H., 2010, *Stochastic control under progressive enlargement of filtrations and applications to multiple defaults risk management*, Stochastic processes and their Applications, 120, 1795-1820.

[47] Protter, P.E., 2005, *Stochastic Integration and Differential Equations*, Springer, NY.

[48] Quenez, M.C. and Sulem, A., 2013, *BSDEs with jumps, optimization and applications to dynamic risk measures*, Stochastic Processes and their Applications, 123, 8, 3328-3357.

[49] Royer, M., 2006, *Backward stochastic differential equations with jumps and related non-linear expectations*, Stochastic process and their applications, 116, 1358-1376.

[50] Takahashi, A., 2015, *Asymptotic Expansion Approach in Finance*, in P Fritz, J Gatheral, A Gulisashvili, A Jacquier and J Teichmann (eds.), *Large Deviations and Asymptotic Methods in Finance*, pp. 345-411, Springer.

[51] Takahashi, A. and Yamada, T., 2015, *On error estimates for asymptotic expansions with Malliavin weights: Application to stochastic volatility model*, Mathematics of Operations Research, Vol. 40, No. 3, pp. 513-541.

[52] Takahashi, A. and Yamada, T., 2015, *A weak approximation with asymptotic expansion and multidimensional Malliavin weights*, Annals of Applied Probability, in press.

[53] Takahashi, A. and Yamada, T., 2015, *An asymptotic expansion of forward-backward SDEs with a perturbed driver*, Forthcoming in International Journal of Financial Engineering.

[54] Touzi, N., 2013, *Optimal stochastic control, stochastic target problems, and backward SDEs*, Springer, New York.

[55] Zhang, J., 2004, *A numerical scheme for bsdes*, The annals of applied probability, 14, 1, 459-488.