CARF Working Paper

CARF-F-387

Solving Backward Stochastic Differential Equations by Connecting the Short-term Expansions

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

First version: 14 June, 2016
This version: 29 June, 2016

CARF is presently supported by Bank of Tokyo-Mitsubishi UFJ, Ltd., Dai-ichi Mutual Life Insurance Company, Nomura Holdings, Inc. and Sumitomo Mitsui Banking Corporation (in alphabetical order). This financial support enables us to issue CARF Working Papers.

Solving Backward Stochastic Differential Equations by Connecting the Short-term Expansions *

Masaaki Fujii[†] & Akihiko Takahashi[‡]

First version: 14 June, 2016
This version: 29 June, 2016

Abstract

In this article, we propose a new numerical computation scheme for Markovian backward stochastic differential equations (BSDEs) by connecting the semi-analytic short-term approximation applied to each time interval, which has a very simple form to implement. We give the error analysis for BSDEs which have generators of quadratic growth with respect to the control variables and bounded terminal conditions. Although the scheme requires higher regularities than the standard method, one can avoid altogether time-consuming Monte Carlo simulation and other numerical integrations for estimating conditional expectations at each space-time node. We provide numerical examples of quadratic-growth (qg) BSDEs as well as standard Lipschitz BSDEs to illustrate the proposed scheme and its empirical convergence rate.

Keywords : asymptotic expansion, discretization, quadratic-growth BSDEs, Lipschitz BSDEs, numerical scheme, BMO-martingales

1 Introduction

The research on backward stochastic differential equations (BSDEs) was initiated by Bismut (1973) [8] for a linear case and followed by Pardoux & Peng (1990) [39] for general non-linear setups. Since then, BSDEs have attracted strong interests among researchers and now exists huge amount of literature. See for example, El Karoui et al. (1997) [22], El Karoui & Mazliak (eds.) (1997) [21], Ma & Yong (2000) [35], Yong & Zhou (1999) [42], Cvitanić & Zhang (2013) [19] and Delong (2013) [20] for excellent reviews and various applications, and also Pardoux & Rascanu (2014) [40] for a recent thorough textbook for BSDEs in the diffusion setup.

In the past decade, there has been significant progress of numerical computation methods for BSDEs. In particular, based on the so-called L^2-*regularity* of the control variables established by Zhang (2001, 2004) [48, 47], now standard backward Monte Carlo schemes for Lipschitz BSDEs have been developed by Bouchard & Touzi (2004) [11], Gobet, Lemor & Warin (2005) [28]. One can find many variants and extensions such as Bouchard & Elie

*All the contents expressed in this research are solely those of the author 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.

(2008) [10] for BSDEs with jumps, Bouchard & Chassagneux (2008) [9] for reflected BSDEs, and Chassagneux & Richou (2016) [15] for the system of reflected BSDEs arising from optimal switching problems. Bender & Denk (2007) [4] proposed a forward scheme free from the linearly growing regression errors existing in the backward schemes; Bender & Steiner (2012) [5] suggested a possible improvement of the scheme [28] by using martingale basis functions for regressions; Crisan & Monolarakis (2014) [18] developed a second-order discretization using the cubature method. A different scheme based on optimal quantization was developed by Bally & Pagès (2003) [6]. See Pagès & Sagna (2015) [38] and references therein for its recent developments. Recently, Chassagneux & Richou (2016) [14] extended the standard backward scheme to quadratic-growth (qg) BSDEs with bounded terminal conditions.

Since the financial crisis in 2009, the importance of BSDEs in the financial industry has grown significantly. This is because BSDEs have become almost unavoidable to understand various valuation adjustments, such as CVA, FVA, KVA etc as well as the optimal risk control under the new regulations. For market developments and related issues, see Brigo, Morini & Pallavicini (2013) [13], Bianchetti & Morini (eds.) (2013) [7] and Crépey & Bielecki (2014) [16]. Despite the above background, the practical use of BSDEs has not been so widespread in the industry. This is partly because that the numerical cost required in the standard Monte Carlo schemes becomes prohibitively high in a practical setup. In order to mitigate the problem, a semi-analytic approximation was proposed by Fujii & Takahashi (2012) [23] and justified in the Lipschitz case by Takahashi & Yamada (2015) [44]. An efficient way of implementation using an interacting particle method by Fujii & Takahashi (2015) [25] has been successfully applied to a large scale credit portfolio by Crépey & Song (2015) [17]. This is an asymptotic approximation around a linear driver motivated by the observation that, for the financial applications, the non-linear part of the driver is typically proportional to an interest rate spread and/or a default intensity which is, at most, of the order of a few percentage points.

However, the non-linear effects may grow and cease to be perturbative when one deals with longer maturities, higher volatilities, or general risk-sensitive control problems for highly concave utility functions. For example, the quadratic-growth terms of the control variables appearing in the exponential utility optimization may give rise to significant non-linearity when the risk-averseness is high. Although one may try higher-order asymptotic expansions, the required analytical calculation soon becomes intractable since one needs very lengthy calculation for each model setup.

In this paper, we aim to achieve the advantages of both the standard Monte Carlo scheme, in terms of generality and scalability, and also the semi-analytic approximation scheme, in terms of the lesser numerical cost. In particular, we follow a discretized backward approximation similar to the one used in the standard scheme and replace time-consuming Monte Carlo simulation estimating the conditional expectation by semi-analytic approximation at each space-time node. For each short time interval, we shall see that a simple low-order asymptotic expansion suffices to achieve necessary accuracy.[1] Although it requires higher regularities, a smaller numerical burden allows us to adopt a finer time partition, which would be too time-consuming in the standard Monte Carlo scheme. Furthermore, we do not have to worry about an appropriate choice of basis functions for regressions, which is a quite delicate problem as clearly explained in [5].

The organization of the paper is as follows: Section 2 explains some important properties

[1] Similar ideas have been applied to stochastic filtering by Fujii (2014) [24] and to European option pricing by Takahashi & Yamada (2016) [45].

of BMO-martingales which prove to be indispensable tools for the following analysis; Section 3 the setup; Section 4 gives the time-discretization and investigates a sequence of qg-BSDEs perturbed in the terminal values; Sections 5 and 6 study the short-term expansion which provides a semi-analytic approximation to the sequence of qg-BSDEs; Section 7 summarizes the computation scheme and gives our main result regarding the total error estimate based on the analysis in the previous sections; Section 8 explains a possible way of implementation using a sparse grid and discusses the sufficient conditions for the convergence; and finally Section 9 and 10 give numerical examples of qg-BSDEs as well as Lipschitz BSDEs, respectively, in order to illustrate the empirical convergence rate of the proposed scheme.

2 Preliminaries

2.1 General setting

Throughout the paper, we fix the terminal time $T > 0$. We work on the filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ carrying a d-dimensional independent standard Brownian motion W. $\mathbb{F} = (\mathcal{F}_t)_{t\in[0,T]}$ is the Brownian filtration satisfying the usual conditions augmented by the $\mathbb{P}$-zero sets.

2.2 Spaces and notation

We denote a generic positive constant by C, which may change line by line and it is sometimes associated with several subscripts (such as $C_{p,K}$) when there is a need to emphasize its dependency on those parameters. $\mathcal{T}_0^T$ denotes the set of all $\mathbb{F}$-stopping times $\tau : \Omega \to [0, T]$. We denote the sup-norm of $\mathbb{R}^k$-valued function $x : [0, T] \to \mathbb{R}^k$, $k \in \mathbb{N}$ by the symbol $||x||_{[a,b]} := \sup\{|x_t|, t \in [a, b]\}$ and write $||x||_t := ||x||_{[0,t]}$.

Let us introduce the following spaces for stochastic processes with $p \geq 2$ and $k \in \mathbb{N}$.

• $\mathcal{S}^p_{[s,t]}(\mathbb{R}^k)$ is the set of $\mathbb{R}^k$-valued adapted processes X satisfying

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

• $\mathcal{S}^\infty_{[s,t]}(\mathbb{R}^k)$ is the set of $\mathbb{R}^k$-valued essentially bounded adapted processes X satisfying

$$||X||_{\mathcal{S}^\infty_{[s,t]}} := \Big|\Big| \sup_{r\in[s,t]} |X_r| \Big|\Big|_\infty < \infty \ .$$

• $\mathcal{H}^p_{[s,t]}(\mathbb{R}^k)$ is the set of $\mathbb{R}^k$-valued progressively measurable processes Z satisfying

$$||Z||_{\mathcal{H}^p_{[s,t]}} := \mathbb{E}\left[\Big(\int_s^t |Z_r|^2 dr\Big)^{\frac{p}{2}}\right]^{\frac{1}{p}} < \infty \ .$$

• $\mathcal{K}^p[s, t]$ is the set of functions (Y, Z) in the space $\mathcal{S}^p_{[s,t]}(\mathbb{R}) \times \mathcal{H}^p_{[s,t]}(\mathbb{R}^{1\times d})$ with the norm defined by

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

We frequently omit the argument $\mathbb{R}^k$ and subscript $[s, t]$ when they are obvious from the context. We use $\big(\Theta_s, s \in [0, T]\big)$ as a collective argument $\Theta_s := (Y_s, Z_s)$ to lighten the notation. We use the following notation for partial derivatives with respect to $x \in \mathbb{R}^d$ such that

$$\partial_x = (\partial_{x^1}, \cdots, \partial_{x^d}) = \Big(\frac{\partial}{\partial x^1}, \cdots, \frac{\partial}{\partial x^d}\Big)$$

and we use $\partial_\theta = (\partial_y, \partial_z)$ for the collective argument Θ.

When there is no confusion, we adopt the so-called Einstein convention assuming the obvious summation of duplicate indexes (such as $i \in \{1, \cdots, d\}$ of x^i) without explicitly using the summation symbol $\sum$. For example, $\partial_{x^i,x^j}\xi(X_T)\partial_x X_T^i \partial_x X_T^j$ assumes the summation about indexes i and j so that it denotes $\sum_{i,j=1}^d \partial_{x^i,x^j}\xi(X_T)\partial_x X_T^i \partial_x X_T^j$.

2.3 BMO-martingale and its properties

Let us introduce BMO-martingales, the associated $\mathcal{H}^2_{BMO}$-space and their properties which play an important role in the following discussions.

Definition 2.1. *A BMO-martingale M is a square integrable martingale with the initial value $M_0 = 0$ and satisfying*

$$||M||^2_{BMO} := \sup_{\tau \in \mathcal{T}_0^T} \Big|\Big| \mathbb{E}\Big[\langle M\rangle_T - \langle M\rangle_\tau | \mathcal{F}_\tau\Big]\Big|\Big|_\infty < \infty \ ,$$

where the supremum is taken over all stopping times $\tau \in \mathcal{T}_0^T$.

We denote the space of BMO-martingales by $BMO(\mathbb{P})$ when the probability measure $\mathbb{P}$ needs to be emphasized.

Definition 2.2. *$\mathcal{H}^2_{BMO}(\mathbb{R}^k)$ is the set of $\mathbb{R}^k$-valued progressively measurable processes Z satisfying*

$$||Z||^2_{\mathcal{H}^2_{BMO}} := \sup_{\tau \in \mathcal{T}_0^T} \Big|\Big| \mathbb{E}\Big[\int_\tau^T |Z_s|^2 ds \Big| \mathcal{F}_\tau\Big]\Big|\Big|_\infty < \infty \ .$$

Note that if $Z \in \mathcal{H}^2_{BMO}(\mathbb{R}^{1\times d})$, we have

$$\Big|\Big|\int_0^\cdot Z_s dW_s\Big|\Big|^2_{BMO} = \sup_{\tau \in \mathcal{T}_0^T} \Big|\Big| \mathbb{E}\Big[\int_\tau^T |Z_s|^2 ds \Big| \mathcal{F}_\tau\Big]\Big|\Big|_\infty = ||Z||^2_{\mathcal{H}^2_{BMO}} < \infty \ ,$$

and hence $Z * W$ is a BMO-martingale. The next result is well-known as *energy inequality.*

Lemma 2.1. *Let Z be in $\mathcal{H}^2_{BMO}$. Then, for any $n \in \mathbb{N}$,*

$$\mathbb{E}\Big[\Big(\int_0^T |Z_s|^2 ds\Big)^n\Big] \le n!\Big(||Z||^2_{\mathcal{H}^2_{BMO}}\Big)^n \ .$$

Proof. See proof of Lemma 9.6.5 in [19]. □

Let $\mathcal{E}(M)$ be a Doléan-Dade exponential of M.

Lemma 2.2. *(Reverse Hölder inequality) Let M be a BMO-martingale. Then, $\big(\mathcal{E}_t(M), t \in [0,T]\big)$ is a uniformly integrable martingale, and for every stopping time $\tau \in \mathcal{T}_0^T$, there exists some positive constant $r^* > 1$ such that the inequality*

$$\mathbb{E}\Big[\mathcal{E}_T(M)^r | \mathcal{F}_\tau\Big] \le C_{r,M}\mathcal{E}_\tau(M)^r \ ,$$

holds for every $1 < r \leq r^*$ with some positive constant $C_{r,M}$ depending only on r and $||M||_{BMO}$.

Proof. See Theorem 3.1 of Kazamaki (1994) [30]. □

Lemma 2.3. *Let M be a square integrable martingale and $\widehat{M} := \langle M \rangle - M$. Then, $M \in BMO(\mathbb{P})$ if and only if $\widehat{M} \in BMO(\mathbb{Q})$ with $d\mathbb{Q}/d\mathbb{P} = \mathcal{E}_T(M)$. Furthermore, $||\widehat{M}||_{BMO(\mathbb{Q})}$ is determined by some function of $||M||_{BMO(\mathbb{P})}$.*

Proof. See Theorem 2.4 and 3.3 in [30]. □

Remark 2.1. *Theorem 3.1 [30] also tells that there exists some decreasing function $\Phi(r)$ with $\Phi(1+) = \infty$ and $\Phi(\infty) = 0$ such that if $||M||_{BMO(\mathbb{P})}$ satisfies*

$$||M||_{BMO(\mathbb{P})} < \Phi(r)$$

then $\mathcal{E}(M)$ allows the reverse Hölder inequality with power r. This implies together with Lemma 2.3, one can take a common positive constant $\bar{r}$ satisfying $1 < \bar{r} \leq r^*$ such that both of the $\mathcal{E}(M)$ and $\mathcal{E}(\hat{M})$ satisfy the reverse Hölder inequality with power $\bar{r}$ under the respective probability measure $\mathbb{P}$ and $\mathbb{Q}$. Furthermore, the upper bound r^* is determined only by $||M||_{BMO(\mathbb{P})}$ (or equivalently by $||M||_{BMO(\mathbb{Q})}$).

3 Setup

Firstly, we introduce the underlying forward process $X_t, t \in [0,T]$:

$$X_t = x_0 + \int_0^t b(r, X_r)dr + \int_0^t \sigma(r, X_r)dW_r \ , \tag{3.1}$$

where $x_0 \in \mathbb{R}^d$, $b : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$, $\sigma : [0,T] \times \mathbb{R}^d \to \mathbb{R}^{d\times d}$.[2]

Assumption 3.1. *(i) For all $t, t' \in [0,T]$ and $x, x' \in \mathbb{R}^d$, there exists a positive constant K such that*

$$|b(t,x) - b(t',x')| + |\sigma(t,x) - \sigma(t',x')| \leq K\big(|t-t'|^{\frac{1}{2}} + |x-x'|\big) \ .$$

(ii) $||b(\cdot,0)||_T + ||\sigma(\cdot,0)||_T \leq K$.
(iii) b and σ are continuously differentiable with arbitrary many times with respect to x and satisfy, for every $m \in \mathbb{N}$,

$$\begin{aligned} &|\partial_x^m b(t,x)| + |\partial_x^m \sigma(t,x)| \leq K \ , \\ &|\partial_x^m b(t,x) - \partial_x^m b(t',x)| + |\partial_x^m \sigma(t,x) - \partial_x^m \sigma(t',x)| \leq K|t-t'|^{1/2} \ , \end{aligned} \tag{3.2}$$

for all $t, t' \in [0,T]$ and $x \in \mathbb{R}^d$.

Let us now introduce a qg-BSDE which is a target of our investigation:

$$Y_t = \xi(X_T) + \int_t^T f(r, X_r, Y_r, Z_r)dr - \int_t^T Z_r dW_r, \quad t \in [0,T] \tag{3.3}$$

where $\xi : \mathbb{R}^d \to \mathbb{R}$, $f : [0,T] \times \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^{1\times d} \to \mathbb{R}$.

[2] Useful standard estimates on the Lipschitz SDEs can be found, for example, in Appendix A of [27].

Assumption 3.2. *(i) f satisfies the quadratic structure condition [2]:*

$$|f(t,x,y,z)| \le l_t + \beta|y| + \frac{\gamma}{2}|z|^2$$

where $\beta \ge 0, \gamma > 0$ are constants, $l : [0,T] \to \mathbb{R}$ is a positive function bounded by a constant K, i.e. $||l||_T \le K$.
(ii) f satisfies, for all $t, t' \in [0,T]$, $y, y' \in \mathbb{R}$, $x, x' \in \mathbb{R}^d$, $z, z' \in \mathbb{R}^{1\times d}$,

$$\begin{aligned}
|f(t,x,y,z) - f(t',x,y,z)| &\le K|t-t'|^{1/2} ,\\
|f(t,x,y,z) - f(t,x,y',z)| &\le K|y-y'| ,\\
|f(t,x,y,z) - f(t,x,y,z')| &\le K\big(1+|z|+|z'|\big)|z-z'| ,\\
|f(t,x,y,z) - f(t,x',y,z)| &\le K\big(1+|y|+|z|^2\big)|x-x'|.
\end{aligned}$$

(iii) ξ is a bounded function satisfying $|\xi(x)| \le K$ for all $x \in \mathbb{R}^d$ and also arbitrary many times continuously differentiable such that

$$|\partial_x^m \xi(x)| \le K$$

for every $m \in \mathbb{N}$ uniformly in $x \in \mathbb{R}^d$.
(iv) the driver f is arbitrary many times differentiable with respect to the spacial variables with continuous derivatives. In particular, we assume that, for all $(t,x,y,z) \in [0,T] \times \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^{1\times d}$,

$$\begin{aligned}
&|\partial_y f(t,x,y,z)| \le K\\
&|\partial_z f(t,x,y,z)| \le K(1+|z|)\\
&|\partial_z^2 f(t,x,y,z)| \le K
\end{aligned}$$

and that all the higher-order derivatives involving y, z are bounded by some constant K. On the other hand, for every $m \in \mathbb{N}$,

$$|\partial_x^m f(t,x,y,z)| \le K(1+|y|+|z|^2) .$$

Note that Chassagneux & Richou (2016) [14] assumes instead the global Lipschitz continuity for the argument x in (ii). We shall discuss some generalization of the terminal function in Section 7.

Remark 3.1. *The smoothness conditions in Assumptions 3.1and 3.2 are required only in the later part of the discussions. The third-order differentiability for the spacial variables is required for the error estimate of the short-term expansions given in Sections 5 and 6. At the last step where we connect the short-term expansions in Section 7, we effectively need $(n \times 2) + 1$-order differentiability, where n is the number of time partitions. On the other hand, we are not assuming differentiability in time t , the non-degeneracy of σ nor uniform Lipschitz continuity of the driver which are required in the so-called four-step scheme (Ma, Protter & Yong (1994) [34]), the finite-difference method for a quasi-linear PDE system.*

It has been well-known since the work of Kobylanski (2000) [32] that there exists a unique solution (Y,Z) for (3.3) such that $Y \in \mathcal{S}^\infty$, $Z \in \mathcal{H}^2_{BMO}$.

Lemma 3.1. *(universal bound) Under Assumptions 3.1 and 3.2, the solution (Y,Z) of (3.3)* *satisfies*

$$||Y||_{\mathcal{S}^\infty} \le e^{\beta T}\Big(||\xi(\cdot)||_{\mathbb{L}^\infty(\mathbb{R}^d)} + T||l||_T\Big) \ ,$$
$$||Z||^2_{\mathcal{H}^2_{BMO}} \le \frac{e^{4\gamma||Y||_{\mathcal{S}^\infty}}}{\gamma^2}\Big(3+6\gamma T(\beta||Y||_{\mathcal{S}^\infty}+||l||_T)\Big) \ .$$

Proof. This follows straightforwardly from the quadratic structure condition [2] which is given by Assumption 3.2 (i). See, for example, Lemma 3.1 and 3.2 in [26]. □

4 A Sequence of qg-BSDEs perturbed in terminals

4.1 Setup

Let us introduce a time partition $\pi : 0 = t_0 < t_1 < \cdots < t_n = T$. We put $h_i := t_i - t_{i-1}$, $|\pi| := \max_{1\le i\le n} h_i$. We denote each interval by $I_i := [t_{i-1}, t_i]$, $i \in \{1, \cdots, n\}$ and assume that there exists some positive constant C such that $|\pi|n \le C$ as well as $|\pi|/h_i \le C$ for every $i \in \{1, \cdots, n\}$. In order to approximate the BSDE (3.3), we introduce a sequence of qg-BSDEs perturbed in the terminal values for each interval $t \in I_i$, $i \in \{1, \cdots, n\}$ in the following way:

$$\overline{Y}^i_t = \widehat{u}^{i+1}(X_{t_i}) + \int_t^{t_i} f(r, X_r, \overline{Y}^i_r, \overline{Z}^i_r)dr - \int_t^{t_i} \overline{Z}^i_r dW_r \ , \tag{4.1}$$

where $\widehat{u}^{i+1} : \mathbb{R}^d \to \mathbb{R}$.

Assumption 4.1. *(i) Each terminal function $\widehat{u}^{i+1}(x)$, $x \in \mathbb{R}^d$ of the period I_i is specified by*

$$\widehat{u}^{n+1}(x) \ := \ \xi(x) \ ,$$
$$\widehat{u}^{i+1}(x) \ := \ \overline{Y}^{i+1,t_i,x}_{t_i} - \delta^{i+1}(x), \quad i \in \{1, \cdots, n-1\}$$

where $\big(\overline{Y}^{i+1,t_i,x}_t, t \in [t_i, t_{i+1}]\big)$ is the solution of (4.1) for the period I_{i+1} corresponding to the underlying process X with the initial data $(t_i, X_{t_i} = x)$ [3]*. The perturbation terms $\delta^{i+1} : \mathbb{R}^d \to \mathbb{R}$, $i \in \{1, \cdots, n\}$ are in $\mathcal{C}^\infty_b$ and absolutely bounded such that, for a given integer $k \ge 3$,*

$$\max_{1\le i\le n} ||\widehat{u}^{i+1}(\cdot)||_{\mathbb{L}^\infty(\mathbb{R}^d)} \le K' \ ,$$
$$\max_{1\le i\le n} ||\partial^m_x \widehat{u}^{i+1}(\cdot)||_{\mathbb{L}^\infty(\mathbb{R}^d)} \le K' \ ,$$

for every $m \in \{1, \cdots, k\}$ with some n-independent positive constant K' satisfying [4]

$$K' \ge e^{\beta T}\Big(||\xi(\cdot)||_{\mathbb{L}^\infty(\mathbb{R}^d)} + T||l||_T\Big) \ .$$

[3] In other words, the underlying forward process is given by $X^{t_i,x}_s = x + \int_{t_i}^s b(r, X^{t_i,x}_r)dr + \int_{t_i}^s \sigma(r, X^{t_i,x}_r)dW_r$, $s \in I_{i+1}$, and hence $\overline{Y}^{i+1,t_i,x}_{t_i}$ is a deterministic function of $x \in \mathbb{R}^d$.

[4] The exact size of K' is somewhat arbitrary if it is big enough not to contradict the true solution of (3.3). This condition is necessary only for making errors from short-term expansions bounded independently from each interval.

(ii) There exists an n-independent positive constant C such that

$$\sum_{i=1}^{n-1} ||\delta^{i+1}(\cdot)||_{\mathbb{L}^\infty(\mathbb{R}^d)} \leq C \ .$$

We use the convention $\delta^{n+1} \equiv 0$ and $\overline{Y}_{t_n}^{n+1} = \xi(X_{t_n})$ in the following.

Remark 4.1. *(ii) in the above assumption is trivial for finite n. However this becomes a constraint when one considers the convergence in the limit $n \to \infty$.*

The classical (as well as variational) differentiability of qg-BSDEs is well-known by the works of Ankirchner et al. (2007) [1], Briand & Confortola (2008) [12] and Imkeller & Reis (2010) [29]. See Fujii & Takahashi (2015) [26] for the extension of these results to qg-BSDEs with Poisson random measures. Using these results, one can show iteratively that $\overline{Y}_{t_i}^{i+1,t_i,x}$ is in $\mathcal{C}_b^\infty$ and bounded when seen as a deterministic function of $x \in \mathbb{R}^d$. Note that the absolute bounds on higher order derivatives $m > k$ may depend on the number of time partitions "n". See also the proof of Proposition 4.1 given below.

4.2 Properties of the solution

Applying the known results of qg-BSDE for each period, one sees that there exists a unique solution $(\overline{Y}^i, \overline{Z}^i) \in \mathcal{S}^\infty_{[t_{i-1},t_i]} \times \mathcal{H}^2_{BMO[t_{i-1},t_i]}$. Applying Lemma 3.1 for each period I_i, one also sees

$$||\overline{Y}^i||_{\mathcal{S}^\infty[t_{i-1},t_i]} \leq ||\overline{Y}||_{\mathcal{S}^\infty} := e^{\beta|\pi|}\Big(K' + |\pi|||l||_T\Big) \ , \tag{4.2}$$

which is bounded uniformly in $i \in \{1, \cdots, n\}$, and so is $||\overline{Z}^i||_{\mathcal{H}^2_{BMO}[t_{i-1},t_i]}$.

Proposition 4.1. *Under Assumptions 3.1, 3.2 and Assumption 4.1 (i), there exists some positive (i,n)-independent constant C such that the process $\overline{Z}^i_t, t \in I_i$ of the solution to the BSDE (4.1) satisfies*

$$|\overline{Z}^i_t| \leq C(1 + |X_t|), \ t \in I_i$$

uniformly in $i \in \{1, \cdots, n\}$.

Proof. We use the representation theorem for the control variable (Theorem 8.5 in [1]) and follow the arguments of Theorem 3.1 in Ma & Zhang (2002) [36]. Let us introduce the parameterized solution $(X^{t,x}, \overline{Y}^{i,t,x}, \overline{Z}^{i,t,x})$ with the initial data $(t,x) \in [t_{i-1}, t_i] \times \mathbb{R}^d$:

$$X^{t,x}_s = x + \int_t^s b(r, X^{t,x}_r)dr + \int_t^s \sigma(r, X^{t,x}_r)dW_r, \tag{4.3}$$

$$\overline{Y}^{i,t,x}_s = \widehat{u}^{i+1}(X^{t,x}_{t_i}) + \int_s^{t_i} f(r, X^{t,x}_r, \overline{Y}^{i,t,x}_r, \overline{Z}^{i,t,x}_r)dr - \int_s^{t_i} \overline{Z}^{i,t,x}_r dW_r, \tag{4.4}$$

$s \in [t, t_i]$ where the classical differentiability of (4.3) and (4.4) with respect to the position x is known [1]. The differential processes $(\partial_x X^{t,x}, \partial_x \overline{Y}^{i,t,x}, \partial_x \overline{Z}^{i,t,x})$ are given by the solutions to the following forward- and backward-SDE (FBSDE):

$$\begin{aligned}
\partial_x X_s^{t,x} &= \mathbb{I} + \int_t^s \partial_x b(r, X_r^{t,x}) \partial_x X_r^{t,x} dr + \int_t^s \partial_x \sigma(r, X_r^{t,x}) \partial_x X_r^{t,x} dW_r, \\
\partial_x \overline{Y}_s^{i,t,x} &= \partial_x \widehat{u}^{i+1}(X_{t_i}^{t,x}) \partial_x X_{t_i}^{t,x} dr + \int_t^{t_i} \Big\{ \partial_x f(r, X_r^{t,x}, \overline{\Theta}_r^{i,t,x}) \partial_x X_r^{t,x} \\
&\quad + \partial_\theta f(r, X_r^{t,x}, \overline{\Theta}_r^{i,t,x}) \partial_x \overline{\Theta}_r^{i,t,x} \Big\} dr - \int_t^{t_i} \partial_x \overline{Z}_r^{i,t,x} dW_r, \qquad (4.5)
\end{aligned}$$

where $\mathbb{I}$ is the $d \times d$ identity matrix and $\partial_x \overline{\Theta}^{i,t,x} = (\partial_x \overline{Y}^{i,t,x}, \partial_x \overline{Z}^{i,t,x})$. Note that $|\partial_y f|$ is bounded and

$$|\partial_z f(r, X_r^{t,x}, \overline{\Theta}_r^{i,t,x})| \leq K\big(1 + |\overline{Z}_r^{i,t,x}|\big)$$

by Assumption 3.2 (iv). By the facts given in (4.2) and the remark that follows, one sees

$$||\partial_z f(\cdot, X_\cdot^{t,x}, \overline{\Theta}_\cdot^{i,t,x})||_{\mathcal{H}^2_{BMO[t,t_i]}} \leq C$$

with some constant C. Thus Corollary 9 in [12] or Theorem A.1 in [26] implies that the BSDE (4.5) has a unique solution satisfying, for any $p \geq 2$,

$$\begin{aligned}
||\partial_x \overline{\Theta}^{i,t,x}||^p_{\mathcal{K}^p[t,t_i]} &\leq C_{p,\bar{q}} \mathbb{E}\Bigg[|\partial_x \widehat{u}^{i+1}(X_{t_i}^{t,x}) \partial_x X_{t_i}^{t,x}|^{p\bar{q}^2} \\
&\quad + \Big(\int_t^{t_i} |\partial_x f(r, X_r^{i,t,x}, \overline{\Theta}_r^{i,t,x}) \partial_x X_r^{i,t,x}| dr\Big)^{p\bar{q}^2} \Bigg]^{\frac{1}{\bar{q}^2}} \\
&\leq C_{p,\bar{q}} \mathbb{E}\Big[||\partial_x X^{t,x}||^{2p\bar{q}^2}_{[t,t_i]} \Big]^{\frac{1}{2\bar{q}^2}} \left(1 + ||\overline{Y}^i||^p_{\mathcal{S}^\infty[t,t_i]} + \mathbb{E}\left[\Big(\int_t^{t_i} |\overline{Z}_r^{i,t,x}|^2 dr\Big)^{2p\bar{q}^2} \right]^{\frac{1}{2\bar{q}^2}} \right)
\end{aligned} \qquad (4.6)$$

where $\bar{q}$ is a positive constant satisfying $q_* \leq \bar{q} < \infty$. Here, $q_* = \frac{r^*}{r^*-1} > 1$ is the conjugate exponent of r^* the upper bound of power with which the Reverse Hölder inequality holds for $\mathcal{E}(\partial_z f * W)$. We have used Assumption 3.2 (iv) and Hölder inequality in the last line.

By the standard estimate of SDE [5], one can show that $||\partial_x X^{t,x}||_{\mathcal{S}^{2p\bar{q}^2}} \leq C$ with some positive constant C that is independent of the initial data (t, x). The boundedness of $\overline{Y}^i$ in (4.2) and the following remark on $\overline{Z}^i$ together with Lemma 2.1 show that the right-hand side of (4.6) is bounded by some positive constant. In particular, one can choose a common constant C for every $i \in \{1, \cdots, n\}$ such that

$$|\partial_x \overline{Y}_t^{i,t,x}| \leq ||\partial_x \overline{Y}^{i,t,x}||_{\mathcal{S}^p[t,t_i]} \leq C$$

uniformly in $(t, x) \in [t_{i-1}, t_i] \times \mathbb{R}^d$. By the representation theorem [1, 36], we have

$$\overline{Z}_t^i = \partial_x \overline{u}^i(t, X_t) \sigma(t, X_t), \;\; t \in [t_{i-1}, t_i]$$

[5] See, for example, Appendix A in [27].

where the function $\partial_x \overline{u}^i : [t_{i-1}, t_i] \times \mathbb{R}^d \to \mathbb{R}^{d\times d}$ is defined by $\partial_x \overline{u}^i(t,x) := \partial_x \overline{Y}_t^{i,t,x}$. Now the Lipschitz property of σ gives the desired result. □

Let us now define a progressively measurable process $\left(\overline{Z}_t, t \in [0,T]\right)$ by

$$\overline{Z}_t := \sum_{i=1}^n \overline{Z}_t^i \mathbf{1}_{\{t_{i-1} \le t < t_i\}}, \ t \in [0,T] \tag{4.7}$$

so that

$$\int_0^T |\overline{Z}_t|^2 dt = \sum_{i=1}^n \int_{t_{i-1}}^{t_i} |\overline{Z}_t^i|^2 dt \ .$$

Proposition 4.2. *Under Assumptions 3.1, 3.2 and 4.1, the process $\left(\overline{Z}_t, \ t \in [0,T]\right)$ defined by (4.7) belongs to $\mathcal{H}^2_{BMO[0,T]}$ satisfying $||\overline{Z}||_{\mathcal{H}^2_{BMO}[0,T]} \le C$ with some n-independent positive constant C.* [6]

Proof. Applying Itô formula to $e^{2\gamma \overline{Y}^i}$, one obtains for $t \in I_i$ that

$$\begin{aligned}\int_t^{t_i} e^{2\gamma \overline{Y}^i_r} 2\gamma^2 |\overline{Z}^i_r|^2 dr &= e^{2\gamma \overline{Y}^i_{t_i}} - e^{2\gamma \overline{Y}^i_t} + \int_t^{t_i} e^{2\gamma \overline{Y}^i_r} 2\gamma f(r, X_r, \overline{Y}^i_r, \overline{Z}^i_r) dr \\ &- \int_t^{t_i} e^{2\gamma \overline{Y}^i_r} 2\gamma \overline{Z}^i_r dW_r \ .\end{aligned}$$

The quadratic structure condition in Assumption 3.2 (i) gives

$$\int_t^{t_i} e^{2\gamma \overline{Y}^i_r} \gamma^2 |\overline{Z}_r|^2 dr \le e^{2\gamma \overline{Y}^i_{t_i}} - e^{2\gamma \overline{Y}^i_t} + \int_t^{t_i} e^{2\gamma \overline{Y}^i_r} 2\gamma \big(l_r + \beta |\overline{Y}^i_r|\big) dr - \int_t^{t_i} e^{2\gamma \overline{Y}^i_r} 2\gamma \overline{Z}^i_r dW_r \ .$$

Since $\overline{Y}^i_{t_i} = \widehat{u}^{i+1}(X_{t_i}) = \overline{Y}^{i+1}_{t_i} - \delta^{i+1}(X_{t_i})$, one obtains

$$\begin{aligned} e^{2\gamma \overline{Y}^i_{t_i}} &= e^{2\gamma \left(\overline{Y}^{i+1}_{t_i} - \delta^{i+1}(X_{t_i})\right)} \\ &= e^{2\gamma \overline{Y}^{i+1}_{t_i}} + e^{2\gamma \overline{Y}^{i+1}_{t_i}} \left(e^{-2\gamma \delta^{i+1}(X_{t_i})} - 1\right) \ .\end{aligned}$$

Since $||\delta^{i+1}(\cdot)||_{\mathbb{L}^\infty(\mathbb{R}^d)} \le C$ uniformly in $i \in \{1, \cdots, n\}$, there exists some positive constant C such that

$$e^{2\gamma \overline{Y}^i_{t_i}} \le e^{2\gamma \overline{Y}^{i+1}_{t_i}} + C e^{2\gamma \overline{Y}^{i+1}_{t_i}} |\delta^{i+1}(X_{t_i})| \ .$$

It follows that, with the choice $t = t_{i-1}$,

$$\begin{aligned}\int_{t_{i-1}}^{t_i} e^{2\gamma \overline{Y}^i_r} \gamma^2 |\overline{Z}^i_r|^2 dr \le \left(e^{2\gamma \overline{Y}^{i+1}_{t_i}} - e^{2\gamma \overline{Y}^i_{t_{i-1}}}\right) + C e^{2\gamma \overline{Y}^{i+1}_{t_i}} |\delta^{i+1}(X_{t_i})| \\ + \int_{t_{i-1}}^{t_i} e^{2\gamma \overline{Y}^i_r} 2\gamma \big(l_r + \beta |\overline{Y}^i_r|\big) dr - \int_{t_{i-1}}^{t_i} e^{2\gamma \overline{Y}^i_r} 2\gamma \overline{Z}^i_r dW_r.\end{aligned}$$

[6] $\overline{Z} \in \mathcal{H}^2_{BMO[0,T]}$ is obvious for finite n since $||\overline{Z}||^2_{\mathcal{H}^2_{BMO}[0,T]} \le \sum_{i=1}^n ||\overline{Z}^i||^2_{\mathcal{H}^2_{BMO}[t_{i-1},t_i]}$.

Thus for any $\tau \in \mathcal{T}_0^T$ and $j := \min\Big(j \in \{1, \cdots, n\} : \tau \le t_j\Big)$,

$$\begin{aligned}
\int_\tau^{t_j} & e^{2\gamma \overline{Y}^j_r}\gamma^2 |\overline{Z}^j_r|^2 dr + \sum_{i=j+1}^n \int_{t_{i-1}}^{t_i} e^{2\gamma \overline{Y}^i_r}\gamma^2 |\overline{Z}^i_r|^2 dr \\
&\le e^{2\gamma \overline{Y}^{n+1}_{t_n}} - e^{2\gamma \overline{Y}^j_\tau} + C\sum_{i=j}^n e^{2\gamma \overline{Y}^{i+1}_{t_i}} |\delta^{i+1}(X_{t_i})| \\
&+ \int_\tau^{t_j} e^{2\gamma \overline{Y}^j_r} 2\gamma\big(l_r + \beta|\overline{Y}^j_r|\big) dr + \sum_{i=j+1}^n \int_{t_{i-1}}^{t_i} e^{2\gamma \overline{Y}^i_r} 2\gamma\big(l_r + \beta|\overline{Y}^i_r|\big) dr \\
&- \int_\tau^{t_j} e^{2\gamma \overline{Y}^j_r} 2\gamma \overline{Z}^j_r dW_r - \sum_{i=j+1}^n \int_{t_{i-1}}^{t_i} e^{2\gamma \overline{Y}^i_r} 2\gamma \overline{Z}^i_r dW_r \ .
\end{aligned}$$

Since $e^{2\gamma \overline{Y}^j_\tau} > 0$ and $\delta^{n+1} \equiv 0$, one obtains

$$\begin{aligned}
&\mathbb{E}\left[\int_\tau^{t_j} e^{2\gamma \overline{Y}^j_r}\gamma^2 |\overline{Z}^j_r|^2 dr + \sum_{i=j+1}^n \int_{t_{i-1}}^{t_i} e^{2\gamma \overline{Y}^i_r}\gamma^2 |\overline{Z}^i_r|^2 dr \Big| \mathcal{F}_\tau\right] \\
&\quad \le \mathbb{E}\left[e^{2\gamma \xi(X_T)} + C\sum_{i=j}^{n-1} e^{2\gamma \overline{Y}^{i+1}_{t_i}} |\delta^{i+1}(X_{t_i})| + \sum_{i=j}^n \int_{t_{i-1}\vee\tau}^{t_i} e^{2\gamma \overline{Y}^i_r} 2\gamma\big(l_r + \beta|\overline{Y}^i_r|\big) dr \Big| \mathcal{F}_\tau\right] \ .
\end{aligned}$$

By Assumption 4.1 (ii) and (4.2), the above inequality implies that there exists some n-independent constant C such that

$$\begin{aligned}
\mathbb{E}\left[\int_\tau^T |\overline{Z}_r|^2 dr \Big| \mathcal{F}_\tau\right] &= \mathbb{E}\left[\int_\tau^{t_j} |\overline{Z}^j_r|^2 dr + \sum_{i=j+1}^n \int_{t_{i-1}}^{t_i} |\overline{Z}^i_r|^2 dr \Big| \mathcal{F}_\tau\right] \\
&\le \frac{e^{4\gamma ||\overline{Y}||_{\mathcal{S}^\infty}}}{\gamma^2}\left(1 + 2\gamma T\big(||l||_T + \beta||\overline{Y}||_{\mathcal{S}^\infty}\big) + C\sum_{i=1}^{n-1} ||\delta^{i+1}(\cdot)||_{\mathbb{L}^\infty(\mathbb{R}^d)}\right) \le C \ ,
\end{aligned}$$

and hence the claim is proved. □

4.3 Error estimates for the perturbed BSDEs in the terminals

Let (Y, Z) be the solution to the BSDE (3.3) and $(\overline{Y}^i, \overline{Z}^i)$ to (4.1). Let us put

$$\begin{aligned}
&\delta Y^i_t := Y_t - \overline{Y}^i_t, \quad \delta Z^i_t := Z_t - \overline{Z}^i_t \\
&\delta f^i(t) := f(t, X_t, Y_t, Z_t) - f(t, X_t, \overline{Y}^i_t, \overline{Z}^i_t) \ ,
\end{aligned}$$

for $t \in I_i$, $i \in \{1, \cdots, n\}$. Then $(\delta Y^i, \delta Z^i)$ follows the BSDE

$$\delta Y^i_t = \delta Y^{i+1}_{t_i} + \delta^{i+1}(X_{t_i}) + \int_t^{t_i} \delta f^i(r) dr - \int_t^{t_i} \delta Z^i_r dW_r \ , \tag{4.8}$$

for $t \in I_i$.

Theorem 4.1. *Under Assumptions 3.1, 3.2 and Assumption 4.1 (i), the inequality*

$$\sum_{i=1}^{n}\int_{t_{i-1}}^{t_i}\mathbb{E}|\delta Z_r^i|^2dr \le C_p\max_{1\le i\le n}\mathbb{E}\left[\sup_{r\in I_i}|\delta Y_r^i|^{2p}\right]^{1/p}+\frac{C}{|\pi|}\sum_{i=1}^{n-1}\mathbb{E}|\delta^{i+1}(X_{t_i})|^2 ,$$

holds for any $p>1$ with some n-independent positive constants C and C_p.

Proof. For each interval I_i, let us define new progressively measurable processes $\big(\beta_r^i, r\in I_i\big)$ and $\big(\gamma_r^i, r\in I_i\big)$ as follows:

$$\beta_r^i := \frac{f(r,X_r,Y_r,Z_r)-f(r,X_r,\overline{Y}_r^i,Z_r)}{\delta Y_r^i}\mathbf{1}_{\delta Y_r^i\neq 0} ,$$

$$\gamma_r^i := \frac{f(r,X_r,\overline{Y}_r^i,Z_r)-f(r,X_r,\overline{Y}_r^i,\overline{Z}_r^i)}{|\delta Z_r^i|^2}\mathbf{1}_{\delta Z_r^i\neq 0}(\delta Z_r^i)^\top .$$

Then, $|\beta^i|\le K$ is a bounded process by the Lipschitz property, and by Proposition 4.1, there exists some (i,n)-independent positive constant C such that

$$|\gamma_r^i|\le K(1+|Z_r|+|\overline{Z}_r^i|)\le C(1+|X_r|) \tag{4.9}$$

for $r\in I_i$, $i\in\{1,\cdots,n\}$. The BSDE (4.8) can now be written as

$$\delta Y_t^i = \delta Y_{t_i}^{i+1}+\delta^{i+1}(X_{t_i})+\int_t^{t_i}\Big(\beta_r^i\delta Y_r^i+\delta Z_r^i\gamma_r^i\Big)dr-\int_t^{t_i}\delta Z_r^i dW_r . \tag{4.10}$$

A simple application of Itô formula gives

$$\begin{aligned}&\mathbb{E}|\delta Y_{t_{i-1}}^i|^2+\int_{t_{i-1}}^{t_i}\mathbb{E}|\delta Z_r^i|^2dr\\ &\qquad =\mathbb{E}\big|\delta Y_{t_i}^{i+1}+\delta^{i+1}(X_{t_i})\big|^2+\int_{t_{i-1}}^{t_i}\mathbb{E}\Big[2\delta Y_r^i\big(\beta_r^i\delta Y_r^i+\delta Z_r^i\gamma_r^i\big)\Big]dr .\end{aligned}$$

By Hölder inequality and (4.9), one obtains with some positive constants C, C_p that

$$\begin{aligned}\frac{1}{2}\int_{t_{i-1}}^{t_i}\mathbb{E}|\delta Z_r^i|^2dr &\le \Big(\mathbb{E}|\delta Y_{t_i}^{i+1}|^2-\mathbb{E}|\delta Y_{t_{i-1}}^i|^2\Big)+C|\pi|\mathbb{E}|\delta Y_{t_i}^{i+1}|^2+\frac{C}{|\pi|}\mathbb{E}|\delta^{i+1}(X_{t_i})|^2\\ &\quad +C\int_{t_{i-1}}^{t_i}\mathbb{E}\Big[|\delta Y_r^i|^2(1+|\gamma_r^i|^2)\Big]dr\\ &\le \Big(\mathbb{E}|\delta Y_{t_i}^{i+1}|^2-\mathbb{E}|\delta Y_{t_{i-1}}^i|^2\Big)+C|\pi|\mathbb{E}|\delta Y_{t_i}^{i+1}|^2+\frac{C}{|\pi|}\mathbb{E}|\delta^{i+1}(X_{t_i})|^2\\ &\quad +C_p|\pi|\mathbb{E}\Big[\sup_{r\in I_i}|\delta Y_r^i|^{2p}\Big]^{1/p}\Big(1+\mathbb{E}\Big[\sup_{r\in I_i}|X_r|^{2q}\Big]^{1/q}\Big)\\ &\le \Big(\mathbb{E}|\delta Y_{t_i}^{i+1}|^2-\mathbb{E}|\delta Y_{t_{i-1}}^i|^2\Big)+C|\pi|\mathbb{E}|\delta Y_{t_i}^{i+1}|^2+\frac{C}{|\pi|}\mathbb{E}|\delta^{i+1}(X_{t_i})|^2\\ &\quad +C_p|\pi|\mathbb{E}\Big[\sup_{r\in I_i}|\delta Y_r^i|^{2p}\Big]^{1/p} ,\end{aligned}$$

where p is an arbitrary constant satisfying $p > 1$ and q (> 1) is its conjugate exponent. Summing up for $i \in \{1, \cdots, n\}$, one obtains

$$\begin{aligned}
&\frac{1}{2}\sum_{i=1}^{n}\int_{t_{i-1}}^{t_i}\mathbb{E}|\delta Z_r^i|^2 dr \\
&\qquad \leq \mathbb{E}|\delta Y_{t_n}^{n+1}|^2 - \mathbb{E}|\delta Y_{t_0}^1|^2 + C|\pi|\sum_{i=1}^{n}\mathbb{E}|\delta Y_{t_i}^{i+1}|^2 \\
&\qquad + \frac{C}{|\pi|}\sum_{i=1}^{n}\mathbb{E}|\delta^{i+1}(X_{t_i})|^2 + C_p|\pi|\sum_{i=1}^{n}\mathbb{E}\left[\sup_{r\in I_i}|\delta Y_r^i|^{2p}\right]^{1/p} .
\end{aligned}$$

Since $\delta Y_{t_n}^{n+1} = \delta^{n+1} = 0$, one gets the desired result. □

Theorem 4.2. *Under Assumptions 3.1, 3.2 and 4.1, there exists some n-independent positive constants $\bar{q} > 1$ and $C_{p,\bar{q}}$ such that*

$$\mathbb{E}\left[\max_{1\leq i\leq n}\sup_{r\in I_i}|\delta Y_r^i|^p\right] \leq C_{p,\bar{q}}\mathbb{E}\left[\Big(\sum_{i=1}^{n-1}|\delta^{i+1}(X_{t_i})|\Big)^{p\bar{q}}\right]^{1/\bar{q}}$$

for any $p > 1$.

Proof. Let us use the same notations β_r^i, γ_r^i defined in Theorem 4.1. We also introduce the process $\Big(\overline{\gamma}_r, r \in [0, T]\Big)$ by

$$\overline{\gamma}_r := \sum_{i=1}^{n}\gamma_r^i \mathbf{1}_{\{t_{i-1}\leq r<t_i\}} .$$

With $\overline{Z}$ defined by (4.7), it satisfies

$$|\overline{\gamma}_r| \leq K(1 + |Z_r| + |\overline{Z}_r|) .$$

By Lemma 3.1 and Proposition 4.2, both of Z and $\overline{Z}$ are in $\mathcal{H}^2_{BMO}$, and so is $\overline{\gamma}$;

$$||\overline{\gamma}||^2_{\mathcal{H}^2_{BMO}} \leq K^2(T + ||Z||^2_{\mathcal{H}^2_{BMO}} + ||\overline{Z}||^2_{\mathcal{H}^2_{BMO}}) \leq C$$

where, in particular, the constant C is n-independent.

From the remark following Definition 2.2, one can show that $\overline{\gamma} * W \in BMO(\mathbb{P})$. Thus the new probability measure $\mathbb{Q}$ can be defined by $d\mathbb{Q}/d\mathbb{P} = \mathcal{E}_T$, where $\mathcal{E}$ is a Doléan-Dade exponential $\mathcal{E}_t := \mathcal{E}\Big(\int_0^T \overline{\gamma}_r^\top dW_r\Big)$. The Brownian motion $W^{\mathbb{Q}}$ under the measure $\mathbb{Q}$ is given by

$$W_t^{\mathbb{Q}} = W_t - \int_0^t \overline{\gamma}_r dr$$

for $t \in [0, T]$. Furthermore, it follows from Lemma 2.2 that there exists a constant r^* satisfying $1 < r^* < \infty$ such that, for every $1 < \bar{r} \leq r^*$, the reverse Hölder inequality of power $\bar{r}$ holds:

$$\frac{1}{\mathcal{E}_\tau}\mathbb{E}\Big[\mathcal{E}_T^{\bar{r}}|\mathcal{F}_\tau\Big]^{1/\bar{r}} \leq C_{\bar{r}} .$$

Here, $\tau \in \mathcal{T}_0^T$ is an arbitrary $\mathbb{F}$-stopping time, $C_{\bar{r}}$ is some positive constant depending only on $\bar{r}$ and the $\mathcal{H}^2_{BMO}$-norm of $\overline{\gamma}$. We put $\bar{q} > 1$ as the conjugate exponent of this $\bar{r}$ in the following. By the last observation, all these constants can be chosen n-independently.

Under the new measure $\mathbb{Q}$, the BSDE (4.10) is given by

$$\delta Y_t^i = \delta Y_{t_i}^{i+1} + \delta^{i+1}(X_{t_i}) + \int_t^{t_i} \beta_r^i \delta Y_r^i dr - \int_t^{t_i} \delta Z_r^i dW_r^{\mathbb{Q}} ,$$

which can be solved as

$$\delta Y_t^i = \mathbb{E}^{\mathbb{Q}}\Big[e^{\int_t^{t_i} \beta_r^i dr}\Big(\delta Y_{t_i}^{i+1} + \delta^{i+1}(X_{t_i})\Big)\Big|\mathcal{F}_t\Big] ,$$

for all $t \in I_i$. Since $|\beta^i| \le K$, one obtains

$$|\delta Y_t^i| \le e^{Kh_i}\mathbb{E}^{\mathbb{Q}}\Big[|\delta Y_{t_i}^{i+1}| + |\delta^{i+1}(X_{t_i})|\Big|\mathcal{F}_t\Big] ,$$

and also

$$|\delta Y_t^i| \le \mathbb{E}^{\mathbb{Q}}\left[e^{K\sum_{j=i}^n h_j}|\delta Y_{t_n}^{n+1}| + \sum_{j=i}^n e^{K\sum_{k=i}^j h_k}|\delta^{j+1}(X_{t_j})|\Big|\mathcal{F}_t\right] ,$$

by iterating the first inequality. Since $\delta Y_{t_n}^{n+1} = \delta^{n+1}(X_{t_n}) = 0$, one concludes

$$|\delta Y_t^i| \le \mathbb{E}^{\mathbb{Q}}\left[\sum_{j=i}^{n-1} e^{K\sum_{k=i}^j h_k}|\delta^{j+1}(X_{t_j})|\Big|\mathcal{F}_t\right]$$

for $t \in I_i$, $i \in \{1, \cdots, n\}$.

The reverse Hölder inequality gives

$$\begin{aligned} |\delta Y_t^i| &\le e^{KT}\mathbb{E}^{\mathbb{Q}}\left[\sum_{j=i}^{n-1}|\delta^{j+1}(X_{t_j})|\Big|\mathcal{F}_t\right] = \frac{e^{KT}}{\mathcal{E}_t}\mathbb{E}\left[\mathcal{E}_T\sum_{j=i}^{n-1}|\delta^{j+1}(X_{t_j})|\Big|\mathcal{F}_t\right] \\ &\le C_{\bar{q}}e^{KT}\mathbb{E}\left[\Big(\sum_{j=i}^{n-1}|\delta^{j+1}(X_{t_j})|\Big)^{\bar{q}}\Big|\mathcal{F}_t\right]^{1/\bar{q}} , \end{aligned}$$

and hence

$$\max_{1\le i\le n}\sup_{t\in I_i}|\delta Y_t^i| \le C_{\bar{q}}\sup_{t\in[0,T]}\mathbb{E}\left[\Big(\sum_{i=1}^{n-1}|\delta^{i+1}(X_{t_i})|\Big)^{\bar{q}}\Big|\mathcal{F}_t\right]^{1/\bar{q}} .$$

Using Jensen and Doob's maximal inequalities, one finally obtains

$$\mathbb{E}\left[\max_{1\le i\le n}\sup_{t\in I_i}|\delta Y_t^i|^p\right] \le C_{p,\bar q}\mathbb{E}\left[\sup_{t\in[0,T]}\mathbb{E}\left[\Big(\sum_{i=1}^{n-1}|\delta^{i+1}(X_{t_i})|\Big)^{\bar q}\Big|\mathcal{F}_t\right]^{p/\bar q}\right]$$

$$\le C_{p,\bar q}\mathbb{E}\left[\sup_{t\in[0,T]}\mathbb{E}\left[\Big(\sum_{i=1}^{n-1}|\delta^{i+1}(X_{t_i})|\Big)^{\bar q}\Big|\mathcal{F}_t\right]^{p}\right]^{1/\bar q} \quad \text{(Jensen's inequality)}$$

$$\le C_{p,\bar q}\mathbb{E}\left[\Big(\sum_{i=1}^{n-1}|\delta^{i+1}(X_{t_i})|\Big)^{p\bar q}\right]^{1/\bar q} \quad \text{(Doob's inequality)}$$

which proves the claim. □

From Theorems 4.1 and 4.2, we obtain the following corollary.

Corollary 4.1. *Under Assumptions 3.1, 3.2 and 4.1, there exist some n-independent positive constants $\bar q > 1$ and $C_{p,\bar q}$ such that*

$$\max_{1\le i\le n}\mathbb{E}\left[\sup_{r\in I_i}|\delta Y_r^i|^{2p}\right]^{\frac{1}{p}} + \sum_{i=1}^n\int_{t_{i-1}}^{t_i}\mathbb{E}|\delta Z_r^i|^2dr \le \frac{C_{p,\bar q}}{|\pi|}\mathbb{E}\left[\Big(\sum_{i=1}^{n-1}|\delta^{i+1}(X_{t_i})|^2\Big)^{p\bar q}\right]^{\frac{1}{p\bar q}},$$

for any $p > 1$.

5 Short-term expansion: Step 1

In the following two sections, we approximate the solution $(\overline{Y}^i, \overline{Z}^i)$ of the BSDE (4.1) semi-analytically and also obtain its error estimate. We need two steps for achieving this goal, which involve the linearization method and the small-variance expansion method for BSDEs proposed in Fujii & Takahashi (2012) [23] and (2015) [27], respectively.

Standing Assumptions for Section 5

We make Assumptions 3.1, 3.2 and Assumption 4.1 (i) (but not (ii)) the standing assumptions for this section.

Let us first introduce the following decomposition of the BSDE (4.1):

$$\overline{Y}_t^{i,[0]} = \widehat{u}^{i+1}(X_{t_i}) - \int_t^{t_i}\overline{Z}_r^{i,[0]}dW_r\ , \tag{5.1}$$

$$\overline{Y}_t^{i,[1]} = \int_t^{t_i} f\big(r, X_r, \overline{Y}_r^{i,[0]}, \overline{Z}_r^{i,[0]}\big)dr - \int_t^{t_i}\overline{Z}_r^{i,[1]}dW_r\ , \tag{5.2}$$

for each interval $t \in I_i$, $i \in \{1, \cdots, n\}$. They are the leading contributions in the linearization method [23, 44].

Lemma 5.1. *For every interval $I_i, i \in \{1, \cdots, n\} \le C$, there exists a unique solution $(\overline{Y}^{i,[0]}, \overline{Z}^{i,[0]})$ to the BSDE (5.1) satisfying, with some (i, n)-independent positive constants*

C and C_p, that

$$||\overline{Y}^{i,[0]}||_{\mathcal{S}^\infty[t_{i-1},t_i]} + ||\overline{Z}^{i,[0]}||_{\mathcal{H}^2_{BMO}[t_{i-1},t_i]} \leq C$$

and

$$||\overline{Z}^{i,[0]}||_{\mathcal{S}^p[t_{i-1},t_i]} \leq C_p$$

for any $p \geq 2$.

Proof. The boundedness $||\overline{Y}^{i,[0]}||_{\mathcal{S}^\infty} \leq C$ follows easily from Assumption 4.1 (i), which then implies $||\overline{Z}^{i,[0]}||_{\mathcal{H}^2_{BMO}} \leq C$. The second claim follows from the similar arguments used in Proposition 4.1 . □

Lemma 5.2. *For every interval I_i, $i \in \{1, \cdots, n\}$, there exists a unique solution $(\overline{Y}^{i,[1]}, \overline{Z}^{i,[1]})$ to the BSDE (5.2) satisfying, with some (i,n)-independent positive constant C_p, that*

$$||\overline{Y}^{i,[1]}||_{\mathcal{S}^p[t_{i-1},t_i]} + ||\overline{Z}^{i,[1]}||_{\mathcal{H}^p[t_{i-1},t_i]} \leq C_p$$

for any $p \geq 2$.

Proof. Since the BSDE is Lipschitz, the existence of a unique solution easily follows. The standard estimate for the Lipschitz BSDEs (see, for example, Appendix B in [27]) and Assumption 3.2 (i) implies

$$\begin{aligned}
\Big|\Big|(\overline{Y}^{i,[1]}, \overline{Z}^{i,[1]})\Big|\Big|^p_{\mathcal{K}^p[t_{i-1},t_i]} &\leq C_p \mathbb{E}\left[\Big(\int_{t_{i-1}}^{t_i} |f(r, X_r, \overline{Y}^{i,[0]}_r, \overline{Z}^{i,[0]}_r)| dr\Big)^p\right] \\
&\leq C_p \mathbb{E}\left[\Big(\int_{t_{i-1}}^{t_i} \big[l_r + \beta|\overline{Y}^{i,[0]}_r| + \frac{\gamma}{2}|\overline{Z}^{i,[0]}_r|^2\big] dr\Big)^p\right] \\
&\leq C_p\Big(||l||^p_T + ||\overline{Y}^{i,[0]}||^p_{\mathcal{S}^p[t_{i-1},t_i]} + ||\overline{Z}^{i,[0]}||^{2p}_{\mathcal{S}^{2p}[t_{i-1},t_i]}\Big) .
\end{aligned}$$

Thus one obtains the desired result by Lemma 5.1. □

We now define the process $(\widetilde{\overline{Y}}^i, \widetilde{\overline{Z}}^i)$ for each interval $t \in I_i$, $i \in \{1, \cdots, n\}$ by

$$\begin{aligned}
\widetilde{\overline{Y}}^i_t &:= \overline{Y}^{i,[0]}_t + \overline{Y}^{i,[1]}_t \\
\widetilde{\overline{Z}}^i_t &:= \overline{Z}^{i,[0]}_t + \overline{Z}^{i,[1]}_t .
\end{aligned}$$

The following property is the main conclusion of this section.

Proposition 5.1. *There exists some (i,n)-independent positive constant C_p such that the inequality*

$$\mathbb{E}\left[||\overline{Y}^i - \widetilde{\overline{Y}}^i||^p_{[t_{i-1},t_i]} + \Big(\int_{t_{i-1}}^{t_i} |\overline{Z}^i_r - \widetilde{\overline{Z}}^i_r|^2 dr\Big)^{\frac{p}{2}}\right] \leq C_p h_i^{3p/2}$$

holds for every interval I_i, $i \in \{1, \cdots, n\}$ with any $p \geq 2$.

Proof. For notational simplicity, let us put

$$\delta Y_t^{i,[0]} := \overline{Y}_t^i - \overline{Y}_t^{i,[0]}, \quad \delta Z_t^{i,[0]} := \overline{Z}_t^i - \overline{Z}_t^{i,[0]}$$
$$\delta Y_t^{i,[1]} := \overline{Y}_t^i - \widetilde{\overline{Y}}_t^i, \quad \delta Z_t^{i,[1]} := \overline{Z}_t^i - \widetilde{\overline{Z}}_t^i$$

for each interval $t \in I_i, i \in \{1, \cdots, n\}$. Then, they are given by the solutions to the following BSDEs respectively:

$$\delta Y_t^{i,[0]} = \int_t^{t_i} f(r, X_r, \overline{Y}_r^i, \overline{Z}_r^i)dr - \int_t^{t_i} \delta Z_r^{i,[0]} dW_r \ ,$$
$$\delta Y_t^{i,[1]} = \int_t^{t_i} \Big(f(r, X_r, \overline{Y}_r^i, \overline{Z}_r^i) - f(r, X_r, \overline{Y}_r^{i,[0]}, \overline{Z}_r^{i,[0]}) \Big) dr - \int_t^{t_i} \delta Z_r^{i,[1]} dW_r \ .$$

By the stability result for the Lipschitz BSDEs (see, for example, Lemma B.2 in [27]), Assumption 3.2 (i), (4.2) and Proposition 4.1, one obtains

$$\begin{aligned} \Big|\Big|(\delta Y^{i,[0]}, \delta Z^{i,[0]})\Big|\Big|^p_{\mathcal{K}^p[t_{i-1},t_i]} &\le C_p \mathbb{E}\left[\Big(\int_{t_{i-1}}^{t_i} \big[l_r + \beta |\overline{Y}_r^i| + \frac{\gamma}{2} |\overline{Z}_r^i|^2 \big] dr \Big)^p \right] \\ &\le C_p h_i^p \left(||l||_T^p + ||\overline{Y}^i||^p_{\mathcal{S}^\infty[t_{i-1},t_i]} + ||\overline{Z}^i||^{2p}_{\mathcal{S}^{2p}[t_{i-1},t_i]} \right) \le C_p h_i^p, \end{aligned} \tag{5.3}$$

with some (i,n)-independent positive constant C_p for $\forall p \ge 2$. Similar analysis for $(\delta Y^{i,[1]}, \delta Z^{i,[1]})$ using Assumption 3.2 (ii) yields

$$\begin{aligned} \Big|\Big|(\delta Y^{i,[1]}, \delta Z^{i,[1]})\Big|\Big|^p_{\mathcal{K}^p[t_{i-1},t_i]} &\le C_p \mathbb{E}\left[\left(\int_{t_{i-1}}^{t_i} \Big[|\delta Y_r^{i,[0]}| + (1 + |\overline{Z}_r^i| + |\overline{Z}_r^{i,[0]}|) |\delta Z_r^{i,[0]}| \Big] dr \right)^p \right] \\ &\le C_p \left(h_i^p ||\delta Y^{i,[0]}||^p_{\mathcal{S}^p[I_i]} + \mathbb{E}\Big[1 + ||\overline{Z}^i||^{2p}_{I_i} + ||\overline{Z}^{i,[0]}||^{2p}_{I_i} \Big]^{\frac{1}{2}} \mathbb{E}\left[\Big(h_i \int_{t_{i-1}}^{t_i} |\delta Z_r^{i,[0]}|^2 dr \Big)^p \right]^{\frac{1}{2}} \right). \end{aligned}$$

By applying Proposition 4.1, Lemma 5.1 and the previous estimate (5.3), one obtains the desired result. □

6 Short-term expansion: Step 2

In the second step of the short-term expansion, we obtain simple analytic approximation for the BSDEs (5.1) and (5.2) while keeping the same order of accuracy given in Proposition 5.1. We use the small-variance expansion method[7] for BSDEs proposed in [27] which renders all the problems into a set of simple ODEs. Furthermore, we shall see that these ODEs can be approximated by a *single-step* Euler method for each interval I_i. This allows us to skip lengthy Monte Carlo simulation for estimating the conditional expectations at each space-time node.

[7] Note that the small-variance asymptotic expansion has been widely applied for the pricing of European contingent claims since the initial attempts by Takahashi (1999) [43] and Kunitomo & Takahashi (2003) [33].

Standing Assumptions for Section 6

Similarly to the last section, we make Assumptions 3.1, 3.2 and Assumption 4.1 (i) (but not (ii)) the standing assumptions for this section.

6.1 Approximation for $(\overline{Y}^{i,[0]}, \overline{Z}^{i,[0]})$

For each interval, we introduce a new parameter ϵ satisfying $\epsilon \in (-c, c)$ with some constant $c > 1$ to perturb (3.1) and (5.1):

$$X_t^\epsilon = X_{t_{i-1}} + \int_{t_{i-1}}^t b(r, X_r^\epsilon)dr + \int_{t_{i-1}}^t \epsilon\sigma(r, X_r^\epsilon)dW_r \ . \tag{6.1}$$

$$\overline{Y}_t^{i,[0],\epsilon} = \widehat{u}^{i+1}(X_{t_i}^\epsilon) - \int_t^{t_i} \overline{Z}_r^{i,[0],\epsilon} dW_r \ . \tag{6.2}$$

for $t \in I_i = [t_{i-1}, t_i]$, $i \in \{1, \cdots, n\}$. Notice that the way ϵ is introduced to X^ϵ, by which we have a different process for each interval I_i.[8] In the following, in order to avoid confusion between the index specifying the interval and the one for the component of $x \in \mathbb{R}^d$, we use the bold Gothic symbols such as $\{\mathbf{i}, \mathbf{j}, \cdots\}$ for the latter, each of which runs through 1 to d.

Lemma 6.1. *The classical derivatives of $(X^\epsilon, \overline{Y}^{i,[0],\epsilon}, \overline{Z}^{i,[0],\epsilon})$ with respect to ϵ*

$$\partial_\epsilon^k X_t^\epsilon := \frac{\partial^k}{\partial\epsilon^k} X_t^\epsilon, \quad \partial_\epsilon^k \overline{Y}^{i,[0],\epsilon} := \frac{\partial^k}{\partial\epsilon^k} \overline{Y}_t^{i,[0],\epsilon}, \quad \partial_\epsilon^k \overline{Z}^{i,[0],\epsilon} := \frac{\partial^k}{\partial\epsilon^k} \overline{Z}_t^{i,[0],\epsilon}$$

for $k = \{1, 2, 3\}$ are given by the solutions to the following forward- and backward-SDEs:

$$\partial_\epsilon X_t^{\epsilon,\mathbf{i}} = \int_{t_{i-1}}^t \partial_{x^\mathbf{j}} b^\mathbf{i}(r, X_r^\epsilon)\partial_\epsilon X_r^{\epsilon,\mathbf{j}} dr + \int_{t_{i-1}}^t \big[\sigma^\mathbf{i}(r, X_r^\epsilon) + \epsilon(\partial_\epsilon X_r^{\epsilon,\mathbf{j}})\partial_{x^\mathbf{j}}\sigma^\mathbf{i}(r, X_r^\epsilon)\big] dW_r \ ,$$

$$\begin{aligned}
\partial_\epsilon^2 X_t^{\epsilon,\mathbf{i}} &= \int_{t_{i-1}}^t \big[\partial_{x^\mathbf{j}} b^\mathbf{i}(r, X_r^\epsilon)\partial_\epsilon^2 X_r^{\epsilon,\mathbf{j}} + \partial^2_{x^\mathbf{j},x^\mathbf{k}} b^\mathbf{i}(r, X_r^\epsilon)\partial_\epsilon X_r^{\epsilon,\mathbf{j}}\partial_\epsilon X_r^{\epsilon,\mathbf{k}}\big] dr \\
&+ \int_{t_{i-1}}^t \big[2(\partial_\epsilon X_r^{\epsilon,\mathbf{j}})\partial_{x^\mathbf{j}}\sigma^\mathbf{i}(r, X_r^\epsilon) + \epsilon(\partial_\epsilon^2 X_r^{\epsilon,\mathbf{j}})\partial_{x^\mathbf{j}}\sigma^\mathbf{i}(r, X_r^\epsilon) \\
&\quad + \epsilon(\partial_\epsilon X_r^{\epsilon,\mathbf{j}})(\partial_\epsilon X_r^{\epsilon,\mathbf{k}})\partial^2_{x^\mathbf{j},x^\mathbf{k}}\sigma^\mathbf{i}(r, X_r^\epsilon)\big] dW_r \ ,
\end{aligned}$$

$$\begin{aligned}
\partial_\epsilon^3 X_t^{\epsilon,\mathbf{i}} &= \int_{t_{i-1}}^t \big[\partial_{x^\mathbf{j}} b^\mathbf{i}(r, X_r^\epsilon)\partial_\epsilon^3 X_r^{\epsilon,\mathbf{j}} + 3\partial^2_{x^\mathbf{j},x^\mathbf{k}} b^\mathbf{i}(r, X_r^\epsilon)\partial_\epsilon^2 X_r^{\epsilon,\mathbf{j}}\partial_\epsilon X_r^{\epsilon,\mathbf{k}} \\
&\quad + \partial^3_{x^\mathbf{j},x^\mathbf{k},x^\mathbf{m}} b^\mathbf{i}(r, X_r^\epsilon)\partial_\epsilon X_r^{\epsilon,\mathbf{j}}\partial_\epsilon X_r^{\epsilon,\mathbf{k}}\partial_\epsilon X_r^{\epsilon,\mathbf{m}}\big] dr + \int_{t_{i-1}}^t \big[3(\partial_\epsilon^2 X_r^{\epsilon,\mathbf{j}})\partial_{x^\mathbf{j}}\sigma^\mathbf{i}(r, X_r^\epsilon) \\
&\quad + 3(\partial_\epsilon X_r^{\epsilon,\mathbf{j}})(\partial_\epsilon X_r^{\epsilon,\mathbf{k}})\partial^2_{x^\mathbf{j},x^\mathbf{k}}\sigma^\mathbf{i}(r, X_r^\epsilon) + \epsilon(\partial_\epsilon^3 X_r^{\epsilon,\mathbf{j}})\partial_{x^\mathbf{j}}\sigma^\mathbf{i}(r, X_r^\epsilon) \\
&\quad + 3\epsilon(\partial_\epsilon^2 X_r^{\epsilon,\mathbf{j}})(\partial_\epsilon X_r^{\epsilon,\mathbf{k}})\partial^2_{x^\mathbf{j},x^\mathbf{k}}\sigma^\mathbf{i}(r, X_r^\epsilon) + \epsilon(\partial_\epsilon X_r^{\epsilon,\mathbf{j}})(\partial_\epsilon X_r^{\epsilon,\mathbf{k}})(\partial_\epsilon X_r^{\epsilon,\mathbf{m}})\partial^3_{x^\mathbf{j},x^\mathbf{k},x^\mathbf{m}}\sigma^\mathbf{i}(r, X_r^\epsilon)\big] dW_r \ ,
\end{aligned}$$

[8] It would be more appropriate to write $X_t^{i,\epsilon}$ to emphasize the dependence on the interval $t \in I_i$, but we have omitted "i" to lighten the notation.

and

$$\partial_\epsilon \overline{Y}_t^{i,[0],\epsilon} = \partial_{x^{\mathbf{j}}} \widehat{u}^{i+1}(X_{t_i}^\epsilon)\partial_\epsilon X_{t_i}^{\epsilon,\mathbf{j}} - \int_t^{t_i} \partial_\epsilon \overline{Z}_r^{i,[0],\epsilon} dW_r \ ,$$

$$\partial_\epsilon^2 \overline{Y}_t^{i,[0],\epsilon} = \partial_{x^{\mathbf{j}}} \widehat{u}^{i+1}(X_{t_i}^\epsilon)\partial_\epsilon^2 X_{t_i}^{\epsilon,\mathbf{j}} + \partial^2_{x^{\mathbf{j}},x^{\mathbf{k}}} \widehat{u}^{i+1}(X_{t_i}^\epsilon)\partial_\epsilon X_{t_i}^{\epsilon,\mathbf{j}}\partial_\epsilon X_{t_i}^{\epsilon,\mathbf{k}} - \int_t^{t_i} \partial_\epsilon^2 \overline{Z}_r^{i,[0],\epsilon} dW_r \ ,$$

$$\begin{aligned}\partial_\epsilon^3 \overline{Y}_t^{i,[0],\epsilon} &= \partial_{x^{\mathbf{j}}} \widehat{u}^{i+1}(X_{t_i}^\epsilon)\partial_\epsilon^3 X_{t_i}^{\epsilon,\mathbf{j}} + 3\partial^2_{x^{\mathbf{j}},x^{\mathbf{k}}} \widehat{u}^{i+1}(X_{t_i}^\epsilon)(\partial_\epsilon^2 X_{t_i}^{\epsilon,\mathbf{j}})(\partial_\epsilon X_{t_i}^{\epsilon,\mathbf{k}}) \\ &\quad + \partial^3_{x^{\mathbf{j}},x^{\mathbf{k}},x^{\mathbf{m}}} \widehat{u}^{i+1}(X_{t_i}^\epsilon)(\partial_\epsilon X_{t_i}^{\epsilon,\mathbf{j}})(\partial_\epsilon X_{t_i}^{\epsilon,\mathbf{k}})(\partial_\epsilon X_{t_i}^{\epsilon,\mathbf{m}}) - \int_t^{t_i} \partial_\epsilon^3 \overline{Z}_r^{i,[0],\epsilon} dW_r \ ,\end{aligned}$$

for $t \in I_i = [t_{i-1}, t_i]$. Einstein convention is used with $\{\mathbf{i}, \mathbf{j}, \cdots\}$ running through 1 to d.

Proof. The classical differentiability can be shown by following the arguments of Theorem 3.1 in [36]. See Section 6 of [27] for more details. □

Lemma 6.2. *For $k = \{1, 2, 3\}$, there exists some (i, n)-independent positive constant $C_{p,k}$ such that the inequality*

$$\mathbb{E}\left[||\partial_\epsilon^k X^\epsilon||^p_{[t_{i-1},t_i]}\right] \le C_{p,k} h_i^{kp/2}$$

holds for every interval I_i, $i \in \{1, \cdots, n\}$ with any $p \ge 2$.

Proof. This can be shown by applying the standard estimates for the Lipschitz SDEs given, for example, in Appendix A of [27]. For $k = 1$,

$$\mathbb{E}\Big[||\partial_\epsilon X^\epsilon||^p_{[t_{i-1},t_i]}\Big] \ \le \ C_p \mathbb{E}\left[\Big(\int_{t_{i-1}}^{t_i} |\sigma(r, X_r^\epsilon)|^2 dr\Big)^{p/2}\right] \le C_p h_i^{p/2} \ .$$

For $k = 2$, one obtains

$$\begin{aligned}\mathbb{E}\Big[||\partial_\epsilon^2 X^\epsilon||^p_{[t_{i-1},t_i]}\Big] &\le C_p \mathbb{E}\left[\Big(\int_{t_{i-1}}^{t_i} |\partial_\epsilon X_r^\epsilon|^2 dr\Big)^p + \Big(\int_{t_{i-1}}^{t_i} \big[|\partial_\epsilon X_r^\epsilon|^2 + |\partial_\epsilon X_r^\epsilon|^4\big] dr\Big)^{\frac{p}{2}}\right] \\ &\le C_p \left(h_i^p \mathbb{E}\Big[||\partial_\epsilon X^\epsilon||^{2p}_{I_i}\Big] + h_i^{p/2}\mathbb{E}\Big[||\partial_\epsilon X^\epsilon||^p_{I_i} + ||\partial_\epsilon X^\epsilon||^{2p}_{I_i}\Big]\right) \le C_p h_i^p \ ,\end{aligned}$$

as desired. One can show the last case $k = 3$ in a similar manner. □

Let introduce the following processes, with $k \in \{0, 1, 2\}$,

$$X_t^{[k]} := \frac{\partial^k}{\partial \epsilon^k} X_t^\epsilon\Big|_{\epsilon=0}, \ \overline{Y}_t^{i,[0],[k]} := \frac{\partial^k}{\partial \epsilon^k} \overline{Y}_t^{i,[0],\epsilon}\Big|_{\epsilon=0}, \quad \overline{Z}_t^{i,[0],[k]} := \frac{\partial^k}{\partial \epsilon^k} \overline{Z}_t^{i,[0],\epsilon}\Big|_{\epsilon=0}$$

and also

$$\widetilde{\overline{Y}}_t^{i,[0]} := \sum_{k=0}^2 \frac{1}{k!} \overline{Y}_t^{i,[0],[k]}, \quad \widetilde{\overline{Z}}_t^{i,[0]} := \sum_{k=0}^2 \frac{1}{k!} \overline{Z}_t^{i,[0],[k]} \tag{6.3}$$

for each interval $t \in I_i$, $i \in \{1, \cdots, n\}$.

Lemma 6.3. *There exists some (i,n)-independent positive constant C_p such that the inequality*

$$\mathbb{E}\left[||\overline{Y}^{i,[0]}-\widetilde{\overline{Y}}^{i,[0]}||^p_{[t_{i-1},t_i]}+\Big(\int_{t_{i-1}}^{t_i}|\overline{Z}_r^{i,[0]}-\widetilde{\overline{Z}}_r^{i,[0]}|^2dr\Big)^{p/2}\right]\le C_p h_i^{3p/2}$$

holds for every interval $I_i, i\in\{1,\cdots,n\}$ with any $p\ge 2$.

Proof. We can use the residual formula of Taylor expansion thanks to the classical differentiability of $\overline{\Theta}^{i,[0],\epsilon}$ with respect to ϵ;

$$\begin{aligned}&\mathbb{E}\left[||\overline{Y}^{i,[0]}-\widetilde{\overline{Y}}^{i,[0]}||^p_{[t_{i-1},t_i]}+\Big(\int_{t_{i-1}}^{t_i}|\overline{Z}_r^{i,[0]}-\widetilde{\overline{Z}}_r^{i,[0]}|^2dr\Big)^{p/2}\right]\\
&\le C_p\mathbb{E}\left[\sup_{r\in I_i}\Big|\frac{1}{2}\int_0^1(1-\epsilon)^2\partial_\epsilon^3\overline{Y}_r^{i,[0],\epsilon}d\epsilon\Big|^p+\left(\int_{t_{i-1}}^{t_i}\Big|\frac{1}{2}\int_0^1(1-\epsilon)^2\partial_\epsilon^3\overline{Z}_r^{i,[0],\epsilon}d\epsilon\Big|^2dr\right)^{p/2}\right]\\
&\le C_p\int_0^1\left(\mathbb{E}\left[||\partial_\epsilon^3\overline{Y}^{i,[0],\epsilon}||^p_{[t_{i-1},t_i]}+\left(\int_{t_{i-1}}^{t_i}|\partial_\epsilon^3\overline{Z}_r^{i,[0],\epsilon}|^2dr\right)^{p/2}\right]\right)d\epsilon\ .\end{aligned}$$

Applying the standard estimates of the Lipschitz BSDEs (see, for example, Appendix B of [27]), the boundedness of $\partial_x^k\widehat{u}^{i+1}$ as well as Lemma 6.2, one obtains

$$\begin{aligned}&\mathbb{E}\left[||\overline{Y}^{i,[0]}-\widetilde{\overline{Y}}^{i,[0]}||^p_{[t_{i-1},t_i]}+\Big(\int_{t_{i-1}}^{t_i}|\overline{Z}_r^{i,[0]}-\widetilde{\overline{Z}}_r^{i,[0]}|^2dr\Big)^{p/2}\right]\\
&\le C_p\int_0^1\left(\mathbb{E}\left[||\partial_\epsilon^3X^\epsilon||^p_{I_i}+||\partial_\epsilon^2X^\epsilon||^p_{I_i}||\partial_\epsilon X^\epsilon||^p_{I_i}+||\partial_\epsilon X^\epsilon||^{3p}_{I_i}\right]\right)d\epsilon\le C_ph_i^{3p/2}\end{aligned}$$

as desired. □

The last lemma implies that it suffices to obtain $(\widetilde{\overline{Y}}^{i,[0]},\widetilde{\overline{Z}}^{i,[0]})$ for our purpose, which is the second order approximation of $(\overline{Y}^{i,[0]},\overline{Z}^{i,[0]})$. Furthermore, as we shall see next, the solution of these BSDEs can be obtained explicitly by simple ODEs thanks to the grading structure introduced by the asymptotic expansion. The relevant system of FBSDEs is summarized below:

$$\begin{aligned}X_t^{[0]}&=X_{t_{i-1}}+\int_{t_{i-1}}^t b(r,X_r^{[0]})dr\ ,\\
X_t^{[1],\mathbf{i}}&=\int_{t_{i-1}}^t\partial_{x^\mathbf{j}}b^\mathbf{i}(r,X_r^{[0]})X_r^{[1],\mathbf{j}}dr+\int_{t_{i-1}}^t\sigma^\mathbf{i}(r,X_r^{[0]})dW_r\ ,\\
X_t^{[2],\mathbf{i}}&=\int_{t_{i-1}}^t\Big(\partial_{x^\mathbf{j}}b^\mathbf{i}(r,X_r^{[0]})X_r^{[2],\mathbf{j}}+\partial^2_{x^\mathbf{j},x^\mathbf{k}}b^\mathbf{i}(r,X_r^{[0]})X_r^{[1],\mathbf{j}}X_r^{[1],\mathbf{k}}\Big)dr\\
&\quad+\int_{t_{i-1}}^t 2X_r^{[1],\mathbf{j}}\partial_{x^\mathbf{j}}\sigma^\mathbf{i}(r,X_r^{[0]})dW_r\ ,\end{aligned}$$

and

$$\overline{Y}_t^{i,[0],[0]} = \widehat{u}^{i+1}(X_{t_i}^{[0]}) - \int_t^{t_i} \overline{Z}_r^{i,[0],[0]} dW_r \ , \tag{6.4}$$

$$\overline{Y}_t^{i,[0],[1]} = \partial_{x^{\mathbf{j}}} \widehat{u}^{i+1}(X_{t_i}^{[0]}) X_{t_i}^{[1],\mathbf{j}} - \int_t^{t_i} \overline{Z}_r^{i,[0],[1]} dW_r \tag{6.5}$$

$$\overline{Y}_t^{i,[0],[2]} = \partial_{x^{\mathbf{j}}} \widehat{u}^{i+1}(X_{t_i}^{[0]}) X_{t_i}^{[2],\mathbf{j}} + \partial^2_{x^{\mathbf{j}},x^{\mathbf{k}}} \widehat{u}^{i+1}(X_{t_i}^{[0]}) X_{t_i}^{[1],\mathbf{j}} X_{t_i}^{[1],\mathbf{k}} - \int_t^{t_i} \overline{Z}_r^{i,[0],[2]} dW_r \ , \tag{6.6}$$

for $t \in I_i, i \in \{1, \cdots, n\}$ with Einstein convention for $\{\mathbf{i}, \mathbf{j}, \cdots\}$.

Definition 6.1. *(Coefficient functions)*
We define the set of functions $\chi : I_i \times \mathbb{R}^d \to \mathbb{R}^d$, $y : I_i \times \mathbb{R}^d \to \mathbb{R}$, $\mathbf{y}^{[1]} : I_i \times \mathbb{R}^d \to \mathbb{R}^d$, $\mathbf{y}^{[2]} : I_i \times \mathbb{R}^d \to \mathbb{R}^d$, $G^{[2]} : I_i \times \mathbb{R}^d \to \mathbb{R}^{d \times d}$, $y_0^{[2]} : I_i \times \mathbb{R}^d \to \mathbb{R}$ *by*

$$\begin{aligned}
&\chi(t,x) := x + \int_{t_{i-1}}^t b(r, \chi(r,x)) dr \ , \\
&y(t,x) := \widehat{u}^{i+1}(\chi(t_i, x)) \ , \\
&\mathbf{y}^{[1]}_{\mathbf{j}}(t,x) := \partial_{x^{\mathbf{j}}} \widehat{u}^{i+1}(\chi(t_i,x)) + \int_t^{t_i} \partial_{x^{\mathbf{j}}} b^{\mathbf{k}}(r, \chi(r,x)) \mathbf{y}^{[1]}_{\mathbf{k}}(r,x) dr \ , \\
&G^{[2]}_{\mathbf{j},\mathbf{k}}(t,x) := \partial^2_{x^{\mathbf{j}},x^{\mathbf{k}}} \widehat{u}^{i+1}(\chi(t_i,x)) \\
&\qquad + \int_t^{t_i} \left\{ \Big(\big[\partial_x b(r, \chi(r,x))\big] G^{[2]}(r,x) \Big)^{\leftrightarrow}_{\mathbf{j},\mathbf{k}} + \partial^2_{x^{\mathbf{j}},x^{\mathbf{k}}} b^{\mathbf{m}}(r, \chi(r,x)) \mathbf{y}^{[2]}_{\mathbf{m}}(r,x) \right\} dr \ , \\
&y_0^{[2]}(t,x) := \int_t^{t_i} \mathrm{Tr}\Big(G^{[2]}(r,x) [\sigma\sigma^\top](r, \chi(r,x)) \Big) dr \ ,
\end{aligned}$$

and $\mathbf{y}^{[2]} = \mathbf{y}^{[1]}$ *for* $(t,x) \in I_i \times \mathbb{R}^d$, $i \in \{1, \cdots, n\}$. *We have used Einstein convention and the notation* $\big([\partial_x b(r,x)]_{\mathbf{i},\mathbf{j}} = \partial_{x^{\mathbf{i}}} b^{\mathbf{j}}(r,x), \mathbf{i}, \mathbf{j} \in \{1, \cdots, d\}\big)$. *We denote the symmetrization by* $A^{\leftrightarrow} := A + A^\top$ *for a* $d \times d$*-matrix* A [9].

Note that the above coefficient functions are given by the ODEs for a given $x \in \mathbb{R}^d$ in each period. The solution of the BSDEs are expressed by these functions in the following way:

Lemma 6.4. *For each period* $t \in I_i, i \in \{1, \cdots, n\}$, *the solutions of the BSDEs (6.4), (6.5) and (6.6) are given by*

$$\overline{Y}_t^{i,[0],[0]} = y(t, X_{t_{i-1}}), \quad \overline{Z}_t^{i,[0],[0]} \equiv 0$$

for the zero-th order,

$$\begin{aligned}
\overline{Y}_t^{i,[0],[1]} &= \mathbf{y}^{[1]}_{\mathbf{j}}(t, X_{t_{i-1}}) X_t^{[1],\mathbf{j}}, \\
\overline{Z}_t^{i,[0],[1]} &= \mathbf{y}^{[1]}_{\mathbf{j}}(t, X_{t_{i-1}}) \sigma^{\mathbf{j}}(t, \chi(t, X_{t_{i-1}})) \ ,
\end{aligned}$$

[9] Hence, $G^{[2]}$ is symmetric matrix valued.

for the first order, and

$$\overline{Y}_t^{i,[0],[2]} = \mathbf{y}_{\mathbf{j}}^{[2]}(t, X_{t_{i-1}}) X_t^{[2],\mathbf{j}} + G_{\mathbf{j},\mathbf{k}}^{[2]}(t, X_{t_{i-1}}) X_t^{[1],\mathbf{j}} X_t^{[1],\mathbf{k}} + y_0^{[2]}(t, X_{t_{i-1}}) \ ,$$
$$\overline{Z}_t^{i,[0],[2]} = 2\Big(\mathbf{y}_{\mathbf{j}}^{[2]}(t, X_{t_{i-1}}) X_t^{[1],\mathbf{k}} \partial_{x^{\mathbf{k}}} \sigma^{\mathbf{j}}(t, \chi(t, X_{t_{i-1}})) + G_{\mathbf{j},\mathbf{k}}^{[2]}(t, X_{t_{i-1}}) X_t^{[1],\mathbf{j}} \sigma^{\mathbf{k}}(t, \chi(t, X_{t_{i-1}}))\Big) \ ,$$

for the second order, where Einstein convention is used.

Proof. This is a special case of the results of Section 8 of [27]. The existence of the unique solution to the BSDEs (6.4), (6.5) and (6.6) is obvious. The expression can be directly checked by applying Itô formula to the suggested forms using the ODEs given in Definition 6.1, and compare the results with the BSDEs. □

Since each interval I_i has a very short span h_i, we expect that we can approximate the above ODEs by just a single-step of Euler method without affecting the order of error given in Lemma 6.3.

Definition 6.2. *(Approximated coefficient functions)*
We define the set of functions; $\overline{\chi} : I_i \times \mathbb{R}^d \to \mathbb{R}^d$, $\overline{y} : I_i \times \mathbb{R}^d \to \mathbb{R}$, $\overline{\mathbf{y}}^{[1]} : I_i \times \mathbb{R}^d \to \mathbb{R}^d$, $\overline{\mathbf{y}}^{[2]} : I_i \times \mathbb{R}^d \to \mathbb{R}^d$, $\overline{G}^{[2]} : I_i \times \mathbb{R}^d \to \mathbb{R}^{d\times d}$, $\overline{y}_0^{[2]} : I_i \times \mathbb{R}^d \to \mathbb{R}$ *by*

$$\begin{aligned}
\overline{\chi}(t,x) &:= x + \Delta(t) b(t_{i-1}, x) \ , \\
\overline{y}(t,x) &:= \widehat{u}^{i+1}(\overline{\chi}(t_i, x)) \ , \\
\overline{\mathbf{y}}_{\mathbf{j}}^{[1]}(t,x) &:= \partial_{x^{\mathbf{j}}} \widehat{u}^{i+1}(\overline{\chi}(t_i,x)) + \delta(t) \partial_{x^{\mathbf{j}}} b^{\mathbf{k}}(t_i, \overline{\chi}(t_i,x)) \partial_{x^{\mathbf{k}}} \widehat{u}^{i+1}(\overline{\chi}(t_i,x)) \ , \\
\overline{G}_{\mathbf{j},\mathbf{k}}^{[2]}(t,x) &:= \partial^2_{x^{\mathbf{j}},x^{\mathbf{k}}} \widehat{u}^{i+1}(\overline{\chi}(t_i,x)) + \delta(t) \Big\{ \Big([\partial_x b(t_i, \overline{\chi}(t_i,x))] \partial^2_{x,x} \widehat{u}^{i+1}(\overline{\chi}(t_i,x)) \Big)^{\leftrightarrow}_{\mathbf{j},\mathbf{k}} \\
&\quad + \partial^2_{x^{\mathbf{j}},x^{\mathbf{k}}} b^{\mathbf{m}}(t_i, \overline{\chi}(t_i,x)) \partial_{x^{\mathbf{m}}} \widehat{u}^{i+1}(\overline{\chi}(t_i,x)) \Big\} \ , \\
\overline{y}_0^{[2]}(t,x) &:= \delta(t) \mathrm{Tr}\Big(\overline{G}^{[2]}(t_i,x) [\sigma\sigma^\top](t_i, \overline{\chi}(t_i,x)) \Big) \ ,
\end{aligned}$$

and $\overline{\mathbf{y}}^{[2]} = \overline{\mathbf{y}}^{[1]}$ *for* $(t,x) \in I_i \times \mathbb{R}^d$, $i \in \{1, \cdots, n\}$. *We have used Einstein convention and the notations* $\Delta(t) := t - t_{i-1}$, $\delta(t) := t_i - t$.

The above functions provide good approximations for the coefficient functions in Definition 6.1 in the following sense:

Lemma 6.5. *There exists some* (i,n)*-independent positive constant* C_p *satisfying*

$$\begin{aligned}
\mathbb{E}\Bigg[&\sup_{t\in I_i} \Big| \chi(t, X_{t_{i-1}}) - \overline{\chi}(t, X_{t_{i-1}}) \Big|^p + \sup_{t\in I_i} \Big| y(t, X_{t_{i-1}}) - \overline{y}(t, X_{t_{i-1}}) \Big|^p \\
&+ \sum_{k=1}^2 \sup_{t\in I_i} \Big| \mathbf{y}^{[k]}(t, X_{t_{i-1}}) - \overline{\mathbf{y}}^{[k]}(t, X_{t_{i-1}}) \Big|^p + \sup_{t\in I_i} \Big| G^{[2]}(t, X_{t_{i-1}}) - \overline{G}^{[2]}(t, X_{t_{i-1}}) \Big|^p \\
&+ \sup_{t\in I_i} \Big| y_0^{[2]}(t, X_{t_{i-1}}) - \overline{y}_0^{[2]}(t, X_{t_{i-1}}) \Big|^p \Bigg] \le C_p h_i^{3p/2} \ ,
\end{aligned}$$

for every interval $I_i, i \in \{1, \cdots, n\}$ *with any* $p \ge 2$.

Proof. See Appendix A. □

We now introduce the processes $(\widehat{Y}_t^{i,[0]}, \widehat{Z}_t^{i,[0]})$ for each period $t \in I_i$. They are defined by $(\widetilde{\overline{Y}}_t^{i,[0]}, \widetilde{\overline{Z}}_t^{i,[0]})$ of (6.3) with the coefficient functions in Definition 6.1 replaced by the approximations in Definition 6.2, i.e.;

$$\begin{aligned}
\widehat{Y}_t^{i,[0]} &:= \overline{y}(t, X_{t_{i-1}}) + (X_t^{[1]})^\top \overline{\mathbf{y}}^{[1]}(t, X_{t_{i-1}}) \\
&+ \frac{1}{2}\Big((X_t^{[2]})^\top \overline{\mathbf{y}}^{[2]}(t, X_{t_{i-1}}) + (X_t^{[1]})^\top \overline{G}^{[2]}(t, X_{t_{i-1}}) X_t^{[1]} + \overline{y}_0^{[2]}(t, X_{t_{i-1}})\Big), \qquad (6.7)\\
\widehat{Z}_t^{i,[0]} &:= \overline{\mathbf{y}}^{[1]\top}(t, X_{t_{i-1}}) \sigma(t, \overline{\chi}(t, X_{t_{i-i}})) \\
&+ \Big((X_t^{[1]})^\top \partial_x \sigma(t, \overline{\chi}(t, X_{t_{i-1}})) \overline{\mathbf{y}}^{[2]}(t, X_{t_{i-1}}) + (X_t^{[1]})^\top \overline{G}^{[2]}(t, X_{t_{i-1}}) \sigma(t, \overline{\chi}(t, X_{t_{i-1}}))\Big), \qquad (6.8)
\end{aligned}$$

where we have used Matrix notation for simplicity. The details of indexing can be checked from those given in Lemma 6.4.

Lemma 6.6. *There exits some (i,n)-independent positive constant C_p such that the inequality*

$$\mathbb{E}\left[||\widetilde{\overline{Y}}^{i,[0]} - \widehat{Y}^{i,[0]}||^p_{[t_{i-1},t_i]} \right] + \mathbb{E}\left[||\widetilde{\overline{Z}}^{i,[0]} - \widehat{Z}^{i,[0]}||^p_{[t_{i-1},t_i]} \right] \le C_p h_i^{3p/2}$$

holds for every interval $I_i, i \in \{1, \cdots, n\}$ for any $p \ge 2$.

Proof. It can be shown easily from Lemmas 6.2 and 6.5. □

Corollary 6.1. *There exits some (i,n)-independent positive constant C_p such that*

$$\mathbb{E}\left[||\overline{Y}^{i,[0]} - \widehat{Y}^{i,[0]}||^p_{[t_{i-1},t_i]} + \Big(\int_{t_{i-1}}^{t_i} |\overline{Z}_r^{i,[0]} - \widehat{Z}_r^{i,[0]}|^2 dr\Big)^{p/2} \right] \le C_p h_i^{3p/2}$$

holds for every interval $I_i, i \in \{1, \cdots, n\}$ with any $p \ge 2$.

Proof. It follows directly from Lemmas 6.3 and 6.6. □

Since $X_{t_{i-1}}^{[1]} = X_{t_{i-1}}^{[2]} = 0$, we have a very simple expression at the starting time t_{i-1} of each period $I_i = [t_{i-1}, t_i]$:

$$\begin{aligned}
\widehat{Y}_{t_{i-1}}^{i,[0]} &= \overline{y}(t_{i-1}, X_{t_{i-1}}) + \frac{1}{2}\overline{y}_0^{[2]}(t_{i-1}, X_{t_{i-1}}), \\
\widehat{Z}_{t_{i-1}}^{i,[0]} &= \overline{\mathbf{y}}^{[1]\top}(t_{i-1}, X_{t_{i-1}}) \sigma(t_{i-1}, X_{t_{i-1}}).
\end{aligned}$$

We have the following continuity property of the approximated solution $(\widehat{Y}^{i,[0]}, \widehat{Z}^{i,[0]})$:

Lemma 6.7. *There exists some (i,n)-independent positive constant C_p such that the inequality*

$$\mathbb{E}\left[\sup_{t\in I_i} \Big| \widehat{Y}_t^{i,[0]} - \widehat{Y}_{t_{i-1}}^{i,[0]} \Big|^p \right] + \mathbb{E}\left[\sup_{t\in I_i} \Big| \widehat{Z}_t^{i,[0]} - \widehat{Z}_{t_{i-1}}^{i,[0]} \Big|^p \right] \le C_p h_i^{p/2} ,$$

holds for every interval $I_i, i \in \{1, \cdots, n\}$ with any $p \geq 2$.

Proof. Since $\overline{y}(t,x) = \overline{y}(t_{i-1}, x)$ for $(t,x) \in I_i \times \mathbb{R}^d$, we have

$$\begin{aligned}\widehat{Y}_t^{i,[0]} - \widehat{Y}_{t_{i-1}}^{i,[0]} &= X_t^{[1]\top}\overline{\mathbf{y}}^{[1]}(t, X_{t_{i-1}}) + \frac{1}{2}\Big(X_t^{[2]\top}\overline{\mathbf{y}}^{[2]}(t, X_{t_{i-1}}) + X_t^{[1]\top}\overline{G}^{[2]}(t, X_{t_{i-1}})X_t^{[1]}\Big) \\ &+ \frac{1}{2}\Big(\overline{y}_0^{[2]}(t, X_{t_{i-1}}) - \overline{y}_0^{[2]}(t_{i-1}, X_{t_{i-1}})\Big) .\end{aligned}$$

The bounds in (A.3), (A.5) and (A.7) as well as the estimates in Lemma 6.2 imply

$$\begin{aligned}\mathbb{E}\left[\sup_{t\in I_i}\Big|\widehat{Y}_t^{i,[0]} - \widehat{Y}_{t_{i-1}}^{i,[0]}\Big|^p\right] &\leq C_p\mathbb{E}\left[||X^{[1]}||_{I_i}^p + ||X^{[2]}||_{I_i}^p + ||X^{[1]}||_{I_i}^{2p} + h_i^p\Big(1 + |X_{t_{i-1}}|^{2p}\Big)\right] \\ &\leq C_p h_i^{p/2}\end{aligned}$$

as desired. Similarly we have

$$\begin{aligned}|\widehat{Z}_t^{i,[0]} - \widehat{Z}_{t_{i-1}}^{i,[0]}| &\leq |\overline{\mathbf{y}}^{[1]}(t, X_{t_{i-1}}) - \overline{\mathbf{y}}^{[1]}(t_{i-1}, X_{t_{i-1}})||\sigma(t, \overline{\chi}(t, X_{t_{i-1}}))| \\ &\quad + |\overline{\mathbf{y}}^{[1]}(t_{i-1}, X_{t_{i-1}})||\sigma(t, \overline{\chi}(t, X_{t_{i-1}})) - \sigma(t_{i-1}, X_{t_{i-1}})| \\ &\quad + |X_t^{[1]}|\Big|\partial_x\sigma(t, \overline{\chi}(t, X_{t_{i-1}}))\overline{\mathbf{y}}^{[2]}(t, X_{t_{i-1}}) + \overline{G}^{[2]}(t, X_{t_{i-1}})\sigma\big(t, \overline{\chi}(t, X_{t_{i-1}})\big)\Big| \\ &\leq C\Delta(t)\Big(1 + |X_{t_{i-1}}|\Big) + C\Big(\Delta(t)^{1/2} + \Delta(t)(1 + |X_{t_{i-1}}|)\Big) + C|X_t^{[1]}|\Big(1 + |X_{t_{i-1}}|\Big) ,\end{aligned}$$

with some positive constant C. Thus, we obtain $\mathbb{E}\left[\sup_{t\in I_i}\Big|\widehat{Z}_t^{i,[0]} - \widehat{Z}_{t_{i-1}}^{i,[0]}\Big|^p\right] \leq C_p h_i^{p/2}$ as desired. $\square$

6.2 Approximation for $(\overline{Y}^{i,[1]}, \overline{Z}^{i,[1]})$

We now want to approximate the remaining BSDE (5.2) appeared in the decomposition of $(\overline{Y}^i, \overline{Z}^i)$. We shall see below that this can be achieved in a very simple fashion. We define the process $(\widehat{Y}^{i,[1]}, \widehat{Z}^{i,[1]})$ by

$$\widehat{Y}_t^{i,[1]} := \delta(t) f\Big(t_{i-1}, X_{t_{i-1}}, \widehat{Y}_{t_{i-1}}^{i,[0]}, \widehat{Z}_{t_{i-1}}^{i,[0]}\Big) , \tag{6.9}$$

$$\widehat{Z}_t^{i,[1]} := 0 , \tag{6.10}$$

for each period $t \in I_i, i \in \{1, \cdots, n\}$. Here, $\delta(t) = t_i - t$ as before.

Lemma 6.8. *There exists some (i,n)-independent positive constant C_p such that the inequality*

$$\mathbb{E}\left[||\overline{Y}^{i,[1]} - \widehat{Y}^{i,[1]}||_{[t_{i-1},t_i]}^p + \Big(\int_{t_{i-1}}^{t_i} |\overline{Z}_r^{i,[1]}|^2 dr\Big)^{p/2}\right] \leq C_p h_i^{3p/2}$$

holds for every interval $I_i, i \in \{1, \cdots, n\}$ with any $p \geq 2$.

Proof. Let us put

$$\delta Y_t^{i,[1]} := \overline{Y}_t^{i,[1]} - \widehat{Y}_t^{i,[1]}, \quad \delta Z_t^{i,[1]} := \overline{Z}_t^{i,[1]}$$

for $t \in I_i$. Then, $(\delta Y^{i,[1]}, \delta Z^{i,[1]})$ is the solution of the following 0-Lipschitz BSDE:

$$\delta Y_t^{i,[1]} = \int_t^{t_i} \delta f(r) dr - \int_t^{t_i} \delta Z_r^{i,[1]} dW_r \ ,$$

where

$$\delta f(r) := f(r, X_r, \overline{Y}_r^{i,[0]}, \overline{Z}_r^{i,[0]}) - f(t_{i-1}, X_{t_{i-1}}, \widehat{Y}_{t_{i-1}}^{i,[0]}, \widehat{Z}_{t_{i-1}}^{i,[0]}) \ .$$

From Assumption 3.2 (ii), it satisfies with the positive constant K that

$$\begin{aligned}
|\delta f(r)| \ \leq\ & |f(r, X_r, \overline{Y}_r^{i,[0]}, \overline{Z}_r^{i,[0]}) - f(t_{i-1}, X_{t_{i-1}}, \overline{Y}_r^{i,[0]}, \overline{Z}_r^{i,[0]})| \\
& + |f(t_{i-1}, X_{t_{i-1}}, \overline{Y}_r^{i,[0]}, \overline{Z}_r^{i,[0]}) - f(t_{i-1}, X_{t_{i-1}}, \widehat{Y}_r^{i,[0]}, \widehat{Z}_r^{i,[0]})| \\
& + |f(t_{i-1}, X_{t_{i-1}}, \widehat{Y}_r^{i,[0]}, \widehat{Z}_r^{i,[0]}) - f(t_{i-1}, X_{t_{i-1}}, \widehat{Y}_{t_{i-1}}^{i,[0]}, \widehat{Z}_{t_{i-1}}^{i,[0]})| \\
\leq\ & K\Big(\Delta(r)^{1/2} + (1 + |\overline{Y}_r^{i,[0]}| + |\overline{Z}_r^{i,[0]}|^2)|X_r - X_{t_{i-1}}|\Big) \\
& + K|\overline{Y}_r^{i,[0]} - \widehat{Y}_r^{i,[0]}| + K(1 + |\overline{Z}_r^{i,[0]}| + |\widehat{Z}_r^{i,[0]}|)|\overline{Z}_r^{i,[0]} - \widehat{Z}_r^{i,[0]}| \\
& + K|\widehat{Y}_r^{i,[0]} - \widehat{Y}_{t_{i-1}}^{i,[0]}| + K(1 + |\widehat{Z}_r^{i,[0]}| + |\widehat{Z}_{t_{i-1}}^{i,[0]}|)|\widehat{Z}_r^{i,[0]} - \widehat{Z}_{t_{i-1}}^{i,[0]}| \ .
\end{aligned}$$

From Lemma 5.1, we know that $||\overline{Y}^{i,[0]}||_{\mathcal{S}^\infty[t_{i-1},t_i]} + ||\overline{Z}^{i,[0]}||_{\mathcal{S}^p[t_{i-1},t_i]} \leq C_p$ for any $p \geq 2$. From (6.7), (6.8), Lemma 6.2, and the boundedness properties of $(\overline{y}, \overline{\mathbf{y}}^{[i]}, \overline{G}^{[2]})$ shown in Appendix A , a similar inequality $||\widehat{Y}^{i,[0]}||_{\mathcal{S}^p[t_{i-1},t_i]} + ||\widehat{Z}^{i,[0]}||_{\mathcal{S}^p[t_{i-1},t_i]} \leq C_p$ holds. The following continuity property of the Lipschitz SDE is also well-known to hold for any $p \geq 2$:

$$\mathbb{E}\left[\sup_{t \in I_i} |X_t - X_{t_{i-1}}|^p\right] \leq C_p h_i^{p/2} \ .$$

Using the above estimates, Corollary 6.1 and Lemma 6.7, one obtains

$$\begin{aligned}
& \mathbb{E}\left[||\delta Y^{i,[1]}||_{I_i}^p + \Big(\int_{t_{i-1}}^{t_i} |\delta Z_r^{i,[1]}|^2 dr\Big)^{p/2}\right] \leq C_p \mathbb{E}\left[\Big(\int_{t_{i-1}}^{t_i} |\delta f(r)| dr\Big)^p\right] \\
& \leq C_p \Bigg\{ h_i^{3p/2} + h_i^p \mathbb{E}\Big[1 + ||\overline{Y}^{i,[0]}||_{I_i}^{2p} + ||\overline{Z}^{i,[0]}||_{I_i}^{4p}\Big]^{\frac{1}{2}} \mathbb{E}\left[\sup_{t \in I_i} |X_t - X_{t_{i-1}}|^{2p}\right]^{\frac{1}{2}} \\
& \quad + h_i^p \left(\mathbb{E}\Big[||\overline{Y}^{i,[0]} - \widehat{Y}^{i,[0]}||_{I_i}^p\Big] + \mathbb{E}\left[\sup_{t \in I_i} |\widehat{Y}_t^{i,[0]} - \widehat{Y}_{t_{i-1}}^{i,[0]}|^p\right]\right) \\
& \quad + \mathbb{E}\Big[1 + ||\overline{Z}^{i,[0]}||_{I_i}^{2p} + ||\widehat{Z}^{i,[0]}||_{I_i}^{2p}\Big]^{\frac{1}{2}} \mathbb{E}\left[\Big(h_i \int_{t_{i-1}}^{t_i} |\overline{Z}_r^{i,[0]} - \widehat{Z}_r^{i,[0]}|^2 dr\Big)^p\right]^{\frac{1}{2}} \\
& \quad + h_i^p \mathbb{E}\Big[1 + ||\widehat{Z}^{i,[0]}||_{I_i}^{2p} + |\widehat{Z}_{t_{i-1}}^{i,[0]}|^{2p}\Big]^{\frac{1}{2}} \mathbb{E}\left[\sup_{t \in I_i} |\widehat{Z}_t^{i,[0]} - \widehat{Z}_{t_{i-1}}^{i,[0]}|^{2p}\right]^{\frac{1}{2}} \Bigg\} \leq C_p h_i^{3p/2} \ ,
\end{aligned}$$

as desired. □

6.3 Summary of short-term expansions

Since there are many steps to follow, let us summarize the result of the short-term expansions. In the last two sections, we have been trying to approximate a BSDE (4.1) i.e., $\overline{Y}_t^i = \widehat{u}^{i+1}(X_{t_i}) + \int_t^{t_i} f(r, X_r, \overline{Y}_r^i, \overline{Z}_r^i)dr - \int_t^{t_i} \overline{Z}_r^i dW_r$. We have obtained the approximated solution $(\widehat{Y}^i, \widehat{Z}^i)$ by

$$\begin{aligned} \widehat{Y}_t^i &:= \widehat{Y}_t^{i,[0]} + \widehat{Y}_t^{i,[1]} , && (6.11) \\ \widehat{Z}_t^i &:= \widehat{Z}_t^{i,[0]} + \widehat{Z}_t^{i,[1]} \ (= \widehat{Z}_t^{i,[0]}), && (6.12) \end{aligned}$$

for every interval $t \in I_i, i \in \{1, \cdots, n\}$, for which the exact expressions can be read from (6.7), (6.8), (6.9) and (6.10). We have the following error estimate:

Theorem 6.1. *Under Assumptions 3.1, 3.2 and Assumption 4.1(i), the process* $(\widehat{Y}^i, \widehat{Z}^i)$ *defined by (6.11) and (6.12) is the short-term approximation of the solution* $(\overline{Y}^i, \overline{Z}^i)$ *of the BSDE (4.1) and satisfies, with some* (i, n)*-independent positive constant* C_p*, that*

$$\mathbb{E}\left[||\overline{Y}^i - \widehat{Y}^i||^p_{[t_{i-1}, t_i]} + \Big(\int_{t_{i-1}}^{t_i} |\overline{Z}_r^i - \widehat{Z}_r^i|^2 dr \Big)^{p/2} \right] \leq C_p h_i^{3p/2} ,$$

for every period $I_i, i \in \{1, \cdots, n\}$ *and* $\forall p \geq 2$.

Proof. It follows directly from Proposition 5.1, Corollary 6.1 and Lemma 6.8. □

Remark 6.1. *If* σ *is non-degenerate, one may first approximate the density of* $X_{t_i}^\epsilon$ *and then* $\overline{Y}_t^{i,[0],\epsilon}$ *by integrating the terminal* $\widehat{u}^{i+1}$ *with the estimated density, which is the standard method for the pricing of European contingent claims proposed in [43, 33]. Although one may significantly relax the smoothness assumptions, it introduces a numerical integration at each space-time node as in the standard Monte Carlo scheme. Thus, we have pursued a much less numerically costly scheme by assuming the higher regularities in this work.*

7 Connecting the sequence of qg-BSDEs

7.1 Connecting procedure

In this section, we complete the approximation procedure by connecting the sequence of qg-BSDEs perturbed in the terminals (4.1). Notice that under the condition $X_{t_{i-1}} = x, x \in \mathbb{R}^d$, the approximated solution

$$\widehat{Y}_{t_{i-1}}^{i,t_{i-1},x} := \widehat{Y}_{t_{i-1}}^i \big|_{X_{t_{i-1}}=x}, \quad \widehat{Z}_{t_{i-1}}^{i,t_{i-1},x} := \widehat{Z}_{t_{i-1}}^i \big|_{X_{t_{i-1}}=x}$$

is given by

$$\begin{aligned} \widehat{Y}_{t_{i-1}}^{i,t_{i-1},x} &= \overline{y}(t_{i-1}, x) + \frac{1}{2}\overline{y}_0^{[2]}(t_{i-1}, x) \\ &+ h_i f\Big(t_{i-1}, x, \overline{y}(t_{i-1}, x) + \frac{1}{2}\overline{y}_0^{[2]}(t_{i-1}, x), \overline{\mathbf{y}}^{[1]\top}(t_{i-1}, x)\sigma(t_{i-1}, x)\Big) , && (7.1) \\ \widehat{Z}_{t_{i-1}}^{i,t_{i-1},x} &= \overline{\mathbf{y}}^{[1]\top}(t_{i-1}, x)\sigma(t_{i-1}, x) , && (7.2) \end{aligned}$$

where

$$\begin{cases} \overline{\chi}(t_i,x) = x + h_i b\big(t_{i-1},x\big)\ , \\ \overline{y}(t_{i-1},x) = \widehat{u}^{i+1}\big(\overline{\chi}(t_i,x)\big)\ , \\ \mathbf{y}^{[1]}(t_{i-1},x) = \Bigg(\ \mathbb{I} + h_i\big[\partial_x b(t_i,\overline{\chi}(t_i,x))\big]\Bigg)\,\partial_x \widehat{u}^{i+1}(\overline{\chi}(t_i,x))\ , \\ \overline{y}_0^{[2]}(t_{i-1},x) = h_i \mathrm{Tr}\Bigg(\ \partial^2_{x,x}\widehat{u}^{i+1}(\overline{\chi}(t_i,x))\big[\sigma\sigma^\top\big](t_i,\overline{\chi}(t_i,x))\Bigg)\ , \end{cases}$$

where $\mathbb{I}$ denotes $d \times d$-identity matrix.

We connect the sequence of qg-BSDEs by the following scheme:

Definition 7.1. *(Connecting Scheme)*

(i) Setting $\widehat{u}^{n+1}(x) = \xi(x)$, $x \in \mathbb{R}^d$.
(ii) Repeating from $i = n$ to $i = 1$ that

- *Calculate the short-term approximation of the BSDE (4.1) by using (7.1) and store the values $\big\{\widehat{Y}^{i,t_{i-1},x'}_{t_{i-1}}\big\}_{x' \in B_i}$ for a finite subset B_i of $\mathbb{R}^d$.*
- *Define the terminal function $\widehat{u}^i(x), x \in \mathbb{R}^d$ for the next period I_{i-1} by*

$$\widehat{u}^i(x) := \mathrm{Interpolation}\Big(\big\{\widehat{Y}^{i,t_{i-1},x'}_{t_{i-1}}\big\}_{x' \in B_i}\Big)(x)$$

where "Interpolation" stands for some smooth interpolating function satisfying the bounds in Assumption 4.1 (i) .

From the definition of δ^i in Assumption 4.1, we have

$$\begin{aligned} \delta^i(x) &= \overline{Y}^{i,t_{i-1},x}_{t_{i-1}} - \widehat{u}^i(x) \\ &:= \delta^i_{SE}(x) + \mathcal{R}^i(x) \end{aligned}$$

where

$$\delta^i_{SE}(x) := \Big(\overline{Y}^{i,t_{i-1},x}_{t_{i-1}} - \widehat{Y}^{i,t_{i-1},x}_{t_{i-1}}\Big)\ , \tag{7.3}$$

$$\mathcal{R}^i(x) := \Big(\widehat{Y}^{i,t_{i-1},x}_{t_{i-1}} - \widehat{u}^i(x)\Big)\ . \tag{7.4}$$

Here, δ^i_{SE} denotes the error of the short-term approximation, and $\mathcal{R}^i$ the interpolation error as well as the regularization effects rendering the approximated function $\widehat{Y}^{i,t_{i-1},x}_{t_{i-1}}$ into the bounds satisfying Assumption 4.1 (i).

Remark 7.1. *As we have mentioned in Remark 3.1, repeating the above connecting procedure effectively introduces higher order derivatives through the short-term expansions. This is why we need the smoothness conditions given in Assumptions 3.1 and 3.2.*

7.2 Total error estimate

Before going to give the main result, we need the following lemma:

Lemma 7.1. *Under Assumptions 3.1 and 3.2, the solution $Y_t, t \in [0,T]$ of the BSDE (3.3) satisfies the continuity property $\mathbb{E}\left[\sup_{s\le u\le t}\big|Y_u - Y_s\big|^p\right] \le C_p\big|t-s\big|^{p/2}$ for any $0 \le s \le t \le T$ and $p \ge 2$ with some positive constant C_p.*

Proof. Using the Burkholder-Davis-Gundy inequality and Assumption 3.2 (i), one obtains

$$\begin{aligned}\mathbb{E}\left[\sup_{s\le u\le t}|Y_u - Y_s|^p\right] &\le C_p\mathbb{E}\left[\Big(\int_s^t |f(r,X_r,Y_r,Z_r)|dr\Big)^p + \Big(\int_s^t |Z_r|^2 dr\Big)^{p/2}\right] \\ &\le C_p\mathbb{E}\left[\Big(\int_s^t \big[l_r + \beta|Y_r| + \frac{\gamma}{2}|Z_r|^2\big]dr\Big)^p + \Big(\int_s^t |Z_r|^2 dr\Big)^{p/2}\right] .\end{aligned}$$

The boundedness of $l, |Y|$ and the fact that $|Z_t| \le C(1+|X_t|)$ with some constant C prove the claim. □

Let us now provide the main result of the paper:

Theorem 7.1. *Define the piecewise constant process $(Y_t^\pi, Z_t^\pi), t\in[0,T]$ by*

$$Y_t^\pi := \widehat{u}^i(X_{t_{i-1}}) , \tag{7.5}$$

$$Z_t^\pi := \overline{\mathbf{y}}^{[1]\top}(t_{i-1},X_{t_{i-1}})\sigma(t_{i-1},X_{t_{i-1}}) , \tag{7.6}$$

for $t_{i-1} \le t < t_i$, $i \in \{1,\cdots,n\}$ and $Y_{t_n}^\pi = \xi(X_{t_n})$, $Z_{t_n}^\pi = 0$, where the $\widehat{u}^i$ and $\overline{\mathbf{y}}^{[1]}$ are those determined by the connecting scheme in Definition 7.1. Then, under Assumptions 3.1, 3.2 and 4.1, there exist some n-independent positive constants $\bar{q} > 1$ and $C_{p,\bar{q}}$ such that the inequality

$$\begin{aligned}&\left(\max_{1\le i\le n}\mathbb{E}\left[\big|\big|Y - Y^\pi\big|\big|^{2p}_{[t_{i-1},t_i]}\right]^{1/p} + \sum_{i=1}^n\int_{t_{i-1}}^{t_i}\mathbb{E}\Big[\big|Z_t - Z_t^\pi\big|^2\Big]dt\right)^{1/2} \\ &\le C_{p,\bar{q}}\sqrt{|\pi|} + C_{p,\bar{q}}\sqrt{n}\,\mathbb{E}\left[\Big(\sum_{i=1}^n\big|\mathcal{R}^i(X_{t_{i-1}})\big|^2\Big)^{p\bar{q}}\right]^{\frac{1}{2p\bar{q}}}\end{aligned}$$

holds for any $p > 1$.

Proof. One obtains, by simple manipulation, that

$$\begin{aligned}&\max_{1\le i\le n}\mathbb{E}\left[\big|\big|Y - Y^\pi\big|\big|^{2p}_{[t_{i-1},t_i]}\right]^{1/p} + \sum_{i=1}^n\int_{t_{i-1}}^{t_i}\mathbb{E}\big|Z_t - Z_t^\pi\big|^2 dt \\ &\le C_p\left(\max_{1\le i\le n}\mathbb{E}\Big[\big|Y_{t_{i-1}} - \overline{Y}^i_{t_{i-1}}\big|^{2p}\Big]^{1/p} + \sum_{i=1}^n\int_{t_{i-1}}^{t_i}\mathbb{E}\big|Z_t - \overline{Z}^i_t\big|^2 dt\right) \\ &+C_p\left(\max_{1\le i\le n}\mathbb{E}\Big[\big|\overline{Y}^i_{t_{i-1}} - \widehat{u}^i(X_{t_{i-1}})\big|^{2p}\Big]^{1/p} + \sum_{i=1}^n\int_{t_{i-1}}^{t_i}\mathbb{E}\big|\overline{Z}^i_t - \widehat{Z}^i_t\big|^2 dt\right) \\ &+C_p\left(\max_{1\le i\le n}\mathbb{E}\left[\sup_{t\in I_i}\big|Y_t - Y_{t_{i-1}}\big|^{2p}\right]^{1/p} + \sum_{i=1}^n\int_{t_{i-1}}^{t_i}\mathbb{E}\big|\widehat{Z}^i_t - \widehat{Z}^i_{t_{i-1}}\big|^2 dt\right) .\end{aligned}$$

It follows, by applying Corollary 4.1, Theorem 6.1, Lemmas 6.7 and 7.1, and expressions (7.3) and (7.4), that

$$
\begin{aligned}
&\max_{1\le i\le n}\mathbb{E}\left[||Y-Y^{\pi}||^{2p}_{[t_{i-1},t_i]}\right]^{1/p}+\sum_{i=1}^{n}\int_{t_{i-1}}^{t_i}\mathbb{E}|Z_t-Z_t^{\pi}|^2dt\\
&\le \frac{C_{p,\bar q}}{|\pi|}\mathbb{E}\left[\Big(\sum_{i=1}^{n-1}|\delta^{i+1}(X_{t_i})|^2\Big)^{p\bar q}\right]^{\frac{1}{p\bar q}}\\
&\quad +C_p\left(\max_{1\le i\le n}\mathbb{E}\Big[|\delta^i(X_{t_{i-1}})|^{2p}\Big]^{1/p}+\sum_{i=1}^{n}h_i^3\right)+C_p|\pi|\\
&\le C_p|\pi|+\frac{C_{p,\bar q}}{|\pi|}\mathbb{E}\left[\Big(\sum_{i=1}^{n}|\delta^{i}(X_{t_{i-1}})|^2\Big)^{p\bar q}\right]^{\frac{1}{p\bar q}}\\
&\le C_p|\pi|+\frac{C_{p,\bar q}}{|\pi|}\left\{n^{p\bar q-1}\sum_{i=1}^{n}\mathbb{E}\left[\big|\delta^i_{SE}(X_{t_{i-1}})\big|^{2p\bar q}\right]+\mathbb{E}\left[\Big(\sum_{i=1}^{n}|\mathcal{R}^i(X_{t_{i-1}})|^2\Big)^{p\bar q}\right]\right\}^{\frac{1}{p\bar q}}\\
&\le C_{p,\bar q}|\pi|+C_{p,\bar q}n\mathbb{E}\left[\Big(\sum_{i=1}^{n}|\mathcal{R}^i(X_{t_{i-1}})|^2\Big)^{p\bar q}\right]^{\frac{1}{p\bar q}},
\end{aligned}
$$

which proves the desired result. □

Remark 7.2. *One may want to replace the* $\{X_{t_i}\}$ *in (7.5) and (7.6) by their Euler approximation:*

$$X^{\pi}_{t_i}:=X^{\pi}_{t_{i-1}}+b(t_{i-1},X^{\pi}_{t_{i-1}})h_i+\sigma(t_{i-1},X^{\pi}_{t_{i-1}})(W_{t_i}-W_{t_{i-1}}),\ \ i\in\{1,\cdots,n\}$$

with $X^{\pi}_{t_0}=X_{t_0}=x_0$*. Using the well-known results (see, for example, [31])*

$$\mathbb{E}\left[\max_{0\le i\le n}|X^{\pi}_{t_i}|^p\right]\le C_p\ ,\qquad \max_{1\le i\le n}\mathbb{E}\left[\sup_{t\in I_i}|X_t-X^{\pi}_{t_{i-1}}|^p\right]\le C_p|\pi|^{p/2}$$

for any $p\ge 2$*, one can show that the same estimate in Theorem 7.1 holds also in this case.*

Remark 7.3. Comments on the Lipschitz BSDEs and the terminal function
The proposed scheme can be equally applicable to the standard Lipschitz BSDEs with the smooth driver f *and also with the smooth terminal function* $\xi(x)$ *of linear growth. Except Proposition 4.2, which is not necessary anymore (and so is Assumption 4.1 (ii)), all the relevant results can be shown with slightly sharper estimates by following the arguments similar to those of Theorem 3.1 in Bouchard & Touzi (2004) [11] with additional perturbation terms in the terminals.*

In financial applications, the terminal function may depend on the average of the security price. This situation can be put in the current framework by adding another underlying process defined by $X'_t=\frac{1}{t}\int_0^t X_r dr,\quad t>0$ *and* $X'_0:=X_0$*. The situation with multiple payoffs such as* $\xi=\sum_{j=1}^m\xi_l(X_{T_j}),0<T_1<\cdots<T_m\le T$ *can easily be handled by making the time partition* π *contain* $\{T_j\}_{1\le j\le m}$ *and modifying the connecting procedure given in Definition 7.1 at these points accordingly.*

8 An example of implementation

8.1 Initial preparation by truncation

In this section, using a specific scheme of implementation, we discuss the interpolation error and the associated conditions for the convergence. In order to make the interpolation-step easy to handle, it is useful to limit the relevant space for X to some compact subset of $\mathbb{R}^d$. This procedure is involved implicitly or explicitly in many of the existing numerical methods at the stage regressing the conditional expectations by a given set of basis functions.

Let us introduce a smooth truncation function $\varphi_M : \mathbb{R} \to \mathbb{R}$ such that

$$\varphi_M(x) = \begin{cases} x & \text{for} \quad |x| \le M \\ \text{sign}(x)(M+1) & \text{for} \quad |x| > M+2 \end{cases} ,$$

bounded $|\varphi_M| \le M+1$ and has derivatives of all orders absolutely bounded by 1. We then consider the truncated BSDE:

$$Y_t^M = \xi_M(X_T) + \int_t^T f_M(r, X_r, Y_r^M, Z_r^M)dr - \int_t^T Z_r^M dW_r \ , \quad t \in [0,T] \tag{8.1}$$

with the definitions

$$\begin{aligned} &\xi_M(x) := \xi(\varphi_M(x)) \ , \\ &f_M(r,x,y,z) := f(r, \varphi_M(x), y, z) \ , \end{aligned}$$

for all $(r,x,y,z) \in [0,T] \times \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^{1\times d}$. It clearly satisfies Assumption 3.2 and hence also Lemma 3.1. With regard to the difference relative to the solution (Y,Z) of the original BSDE (3.3), we have the following result:

Theorem 8.1. *Under Assumptions 3.1 and 3.2, there exists some positive constant $\bar{q}$ (>1) and $C_{p,\bar{q}}$ such that the inequality*

$$\mathbb{E}\left[||Y - Y^M||_T^p + \Big(\int_0^T |Z_r - Z_r^M|^2 dr \Big)^{\frac{p}{2}} \right] \le C_{p,\bar{q}}(1 + |x_0|^{p+l})/M^l$$

holds for any $p \ge 2$ and $l > 0$.

Proof. Let us put the processes

$$\begin{aligned} &\delta Y_r^M := Y_r - Y_r^M, \quad \delta Z_r^M := Z_r - Z_r^M \\ &\beta_r^M := \frac{f(r, X_r, Y_r, Z_r) - f(r, X_r, Y_r^M, Z_r)}{\delta Y_r^M} \mathbf{1}_{\delta Y_r^M \neq 0} \ , \\ &\gamma_r^M := \frac{f(r, X_r, Y_r^M, Z_r) - f(r, X_r, Y_r^M, Z_r^M)}{|\delta Z_r^M|^2} \mathbf{1}_{\delta Z_r^M \neq 0} (\delta Z_r^M)^\top \ , \\ &\delta f^M(r) := f(r, X_r, Y_r^M, Z_r^M) - f(r, \varphi_M(X_r), Y_r^M, Z_r^M) \ , \end{aligned}$$

for $r \in [0,T]$ and $\delta\xi^M(x) := \xi(x) - \xi(\varphi_M(x))$ for $x \in \mathbb{R}^d$. Then, $(\delta Y^M, \delta Z^M)$ is a unique solution to the following BSDE:

$$\delta Y_t^M = \delta\xi^M(X_T) + \int_t^T \Big\{\beta_r^M \delta Y_r^M + \delta Z_r^M \gamma_r^M + \delta f^M(r)\Big\} dr - \int_t^T \delta Z_r^M dW_r \ .$$

Here, $|\beta^M|$ is bounded and $\gamma^M \in \mathcal{H}^2_{BMO}$ whose norm is bounded independently from M. Thus, with some constant $\bar{q} > 1$ depending only on $||\gamma^M||_{\mathcal{H}^2_{BMO}}$, one obtains from Lemma 2.1 and Theorem A.1 [26] that

$$\begin{aligned}
&\mathbb{E}\Big[||\delta Y^M||_T^p + \Big(\int_0^T |\delta Z_r^M|^2 dr\Big)^{p/2}\Big] \\
&\quad \le C_{p,\bar{q}} \mathbb{E}\Big[|\delta\xi^M(X_T)|^{p\bar{q}^2} + \Big(\int_0^T |\delta f^M(r)| dr\Big)^{p\bar{q}^2}\Big]^{1/\bar{q}^2} \\
&\quad \le C_{p,\bar{q}} \mathbb{E}\Big[\Big(1 + \Big(\int_0^T \big[1 + |Y_r^M| + |Z_r^M|^2\big] dr\Big)^{p\bar{q}^2}\Big) ||X - \varphi_M(X)||_T^{p\bar{q}^2}\Big]^{1/\bar{q}^2} \\
&\quad \le C_{p,\bar{q}} \mathbb{E}\Big[||X \mathbf{1}_{\{|x|>M\}}||_T^{2p\bar{q}^2}\Big]^{1/2\bar{q}^2} \le C_{p,\bar{q}} \mathbb{E}\Big[||X||_T^{2p\bar{q}^2} \Big(\frac{||X||_T}{M}\Big)^{2l\bar{q}^2}\Big]^{1/2\bar{q}^2} \\
&\quad \le C_{p,\bar{q}} (1 + |x|^{p+l})/M^l \ ,
\end{aligned}$$

for any $p \ge 2$ and $l > 0$ as desired. □

From the last theorem, one can make the difference $(Y - Y^M, Z - Z^M)$ arbitrary small easily by taking large enough M. Thus we may treat the x-truncated BSDE (8.1) as the target of analysis. In this case, $u^M : [0,T] \times \mathbb{R}^d \to \mathbb{R}$ defined by $u^M(t,x) := Y_t^{M,t,x}$ must satisfy

$$u^M(t,x) = u^M(t, \varphi_M(x)) \tag{8.2}$$

for all $(t,x) \in [0,T] \times \mathbb{R}^d$. Since the solution of (8.1) must be equal to the solution of (3.3) with the forward process $X_t, t \in [0,T]$ replaced by $\varphi_M(X_t), t \in [0,T]$. Therefore, in this case, one can concentrate on the interpolation in the compact set $|x| \le M$ and smoothly connect to the constant function outside.

8.2 Interpolation and discussions on convergence

Since we cannot explicitly estimate the effects of the regularization step in Definition 7.1, let us first assume that the bounds of Assumption 4.1 (i) are satisfied by $\widehat{Y}_{t_{j-1}}^{j,t_{j-1},x}$ for every $j \in \{i, \cdots, n\}$ and concentrate on the interpolation problem at the time t_{i-1} within a compact set $|x| \le M$. There exists a very interesting result on high dimensional polynomial interpolation on sparse grids. By Theorem 8 (as well as Remark 9) of Barthelmann et al. [3], it is known that there exists an interpolating function satisfying the following uniform estimates on the compact set:

$$\Big|\Big|\widehat{Y}_{t_{i-1}}^{i,t_{i-1},\cdot} - \mathcal{A}^{q,d}\big(\widehat{Y}_{t_{i-1}}^{i,t_{i-1},\cdot}\big)\Big|\Big|_\infty \le C_{q,d} N_{(q,d)}^{-k} (\log(N_{(q,d)}))^{(k+1)(d-1)} \ . \tag{8.3}$$

Here, $\mathcal{A}^{q,d}(f) : \mathbb{R}^d \to \mathbb{R}$ is an interpolating *polynomial* function of degree q $(\ge d)$ for the function $f : \mathbb{R}^d \to \mathbb{R}$ based on the Smolyak algorithm. The interpolating function is uniquely determined by the values of $f(x_i), x_i \in H(q,d)$ where $H(q,d)$ is the sparse grid whose number of nodes is give by $N_{(q,d)}$. The power k on the right-hand side of (8.3) comes from the fact that $f(x) = \widehat{Y}^{i,t_{i-1},x}_{t_{i-1}}$ has bounded derivatives up to the k-th order. $C_{q,d}$ is some positive constant depending only on (q,d) and $||\partial_x^m f||_\infty$ of $m = \{0,\cdots,k\}$. The sparse grid $H(q,d)$ is the set of points on which the Chebyshev polynomials take the extrema. For details, see [3, 41] and references therein. The sparse grid method looks very attractive since (8.3) has only weak dependency on the dimension d.

Unfortunately, however, we do not know the error estimate similar to (8.3) regarding on its derivatives, and hence we cannot be sure that if $\mathcal{A}^{q,d}(\widehat{Y}^{i,t_{i-1},\cdot}_{t_{i-1}})$ satisfies the bounds in Assumption 4.1 (i) when the derivatives are calculated analytically from it. Thus, let us write $\partial_\Delta \mathcal{A}^{q,d}(\widehat{Y}^{i,t_{i-1},\cdot}_{t_{i-1}})$ and $\partial^2_\Delta \mathcal{A}^{q,d}(\widehat{Y}^{i,t_{i-1},\cdot}_{t_{i-1}})$ as approximated the first and the second order derivatives using the finite difference method with central difference scheme of size Δ, which is also numerically more efficient. Then, using (8.3), we have

$$\left|\left|\partial_x \widehat{Y}^{i,t_{i-1},\cdot}_{t_{i-1}} - \partial_\Delta \mathcal{A}^{q,d}(\widehat{Y}^{i,t_{i-1},\cdot}_{t_{i-1}})\right|\right|_\infty \leq C\left(\Delta^2 + \frac{C_{q,d}N^{-k}_{(q,d)}(\log(N_{(q,d)}))^{(k+1)(d-1)}}{\Delta}\right), \tag{8.4}$$

$$\left|\left|\partial^2_x \widehat{Y}^{i,t_{i-1},\cdot}_{t_{i-1}} - \partial^2_\Delta \mathcal{A}^{q,d}(\widehat{Y}^{i,t_{i-1},\cdot}_{t_{i-1}})\right|\right|_\infty \leq C\left(\Delta^2 + \frac{C_{q,d}N^{-k}_{(q,d)}(\log(N_{(q,d)}))^{(k+1)(d-1)}}{\Delta^2}\right), \tag{8.5}$$

where C is some constant depending only on the bounds K' since $\widehat{Y}^{i,t_{i-1},x}_{t_{i-1}}$ is assumed to satisfy the bounds in Assumption 4.1 (i).

From the shape of the relevant BSDEs (6.4), (6.5) and (6.6), and also from the standard estimate for the Lipschitz BSDEs (see, Lemma B.2 of [27], for example), we can show that the same error size of Corollary 6.1 (and hence also of Theorem 6.1) for the next period $t \in I_{i-1}$ is maintained as long as the following inequalities are satisfied:

$$\begin{aligned} N^{-k}_{(q,d)}(\log(N_{(q,d)}))^{(k+1)(d-1)} &\leq Ch^{3/2}_{i-1}, \\ \left(\Delta^2 + \frac{N^{-k}_{(q,d)}(\log(N_{(q,d)}))^{(k+1)(d-1)}}{\Delta}\right)h^{1/2}_{i-1} &\leq Ch^{3/2}_{i-1}, \\ \left(\Delta^2 + \frac{N^{-k}_{(q,d)}(\log(N_{(q,d)}))^{(k+1)(d-1)}}{\Delta^2}\right)h_{i-1} &\leq Ch^{3/2}_{i-1}, \end{aligned}$$

with some positive constant C. Note that the factors $h^{1/2}_{i-1}$ and h_{i-1} in the last two inequalities come from the norms of $X^{[1]}$ and $(X^{[1]})^2$ in the BSDEs (6.5) and (6.6). The above conditions can be achieved, for example, by choosing Δ and the order of sparse grid $H(q,d)$ such that

$$\begin{aligned} &\Delta = C|\pi|^{1/2} \\ &N_{(q,d)} > Cn^{3/(2k)}. \end{aligned}$$

Once this is done, we can perform the short-term expansions based on $\mathcal{A}^{q,d}(\widehat{Y}^{i,t_{i-1},\cdot}_{t_{i-1}})$, $\partial_\Delta \mathcal{A}^{q,d}(\widehat{Y}^{i,t_{i-1},\cdot}_{t_{i-1}})$, and $\partial^2_\Delta \mathcal{A}^{q,d}(\widehat{Y}^{i,t_{i-1},\cdot}_{t_{i-1}})$. [10]

In each period, thanks to the assumed regularities given by Assumptions 3.1 and 3.2, we know that the approximated solution by the short-term expansion (7.1) is smooth and bounded for each period in the set $|x| \le M$. [11] The difficult problem that still remains is that if the uniform bounds of Assumption 4.1 (i) is maintained as the limit $n \to \infty$. As long as this non-divergence condition holds true, the above discussions using the sparse grid imply that one obtains $||\mathcal{R}^i||_\infty \le C|\pi|^{3/2}$ and hence also $||\delta^i||_\infty \le C|\pi|^{3/2}$. Since Assumption 4.1 (ii) is now satisfied as the limit $n \to \infty$, one can get the converging result of the order of $|\pi|^{1/2}$ by Theorem 7.1.

One can expect from the approximation scheme in Section 7, the issue of bounded derivatives as $n \to \infty$ is closely related to the stability of discrete solution technique of PDEs. As long as the numerical experiments we tried, we have not encountered the divergence of derivatives by choosing

$$\Delta = C\sqrt{|\pi|} \tag{8.6}$$

where C is the constant of the order of X's volatility. When a derivative blowups, a slightly bigger factor C makes it bounded in every example. However, getting the sufficient conditions for the non-divergence in a general setup (and also with different ways of estimating the derivatives) seems quite involved. We leave this problem for further research.

9 Numerical examples (qg-BSDEs)

In the remainder of the paper, we demonstrate our computation scheme and its empirical convergence rate using illustrative models. For simplicity, we use a full grid (instead of a sparse grid) at each time step so that we can approximate all the relevant derivatives by the central difference scheme. In this case, since there is no interpolation involved, the numerical solution is guaranteed to converge the true solution as long as there is no blowup of the derivatives (Assumption 4.1 (i)).

9.1 A solvable qg-BSDE

Let us first consider the following model with $d = 2$ similar to those studied in [14]:

$$X_t = x_0 + \int_0^t \begin{pmatrix} b_1 X_s^1 \\ b_2 X_s^2 \end{pmatrix} ds + \int_0^t \begin{pmatrix} \sigma_1 X_s^1 & 0 \\ 0 & \sigma_2 X_s^2 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ \rho & \sqrt{1-\rho^2} \end{pmatrix} dW_s \ , \tag{9.1}$$

$$Y_t = \xi(X_T) + \int_t^T \frac{a}{2}|Z_s|^2 ds - \int_t^T Z_s dW_s \ , \tag{9.2}$$

where $b_i, \sigma_i,\ i \in \{1,2\},\ \rho \in [-1,1]$ and a are all constants. For this example, by using a exponential transformation $\big(e^{aY_t}, t \in [0,T]\big)$, we obtain a closed form solution:

$$Y_t = \frac{1}{a} \log\Big(\mathbb{E}\Big[\exp\big(a\xi(X_T)\big)\Big|\mathcal{F}_t\Big]\Big) \ . \tag{9.3}$$

[10] The interpolation function may not necessarily be smooth as discussed above, since we only need sufficient accuracy to calculate the derivatives based on the finite difference method. Hence, one may adopt a simpler sparse-grid interpolation using piecewise linear functions. See [46] and also [37] for interesting applications.

[11] By the linear growth property of σ, the approximated solution (7.1) has a quadratic growth in x from $h_i f(\cdots)$ term. Thus we need to scale, at least, $|\pi|M^2 = \mathcal{O}(1)$.

The expectation can be evaluated semi-analytically by integrating over the density of X. We use

$$\xi(x) = 3\sum_{\mathbf{i}=1}^{2} \sin^2(x^{\mathbf{i}}) \tag{9.4}$$

as the terminal value function, and set $x_0 = (1,1)^\top$, $T = 1$, $b_1 = b_2 = 0.05$, $\rho = 0.3$. We have tested the following five sets of parameters σ_i and a:

$$\begin{aligned} \text{set}_1 &= \Big\{\sigma_1 = \sigma_2 = 0.5, a = 1.0\Big\}, \quad \text{set}_2 = \Big\{\sigma_1 = \sigma_2 = 0.5, a = 2.0\Big\} \\ \text{set}_3 &= \Big\{\sigma_1 = \sigma_2 = 0.5, a = 3.0\Big\}, \quad \text{set}_4 = \Big\{\sigma_1 = \sigma_2 = 1.0, a = 3.0\Big\} \\ \text{set}_5 &= \Big\{\sigma_1 = \sigma_2 = 0.5, a = 4.0\Big\} \end{aligned} \tag{9.5}$$

by changing the number of partitions from $n = 1$ to $n = 300$. In Figure 1, we have plotted $\log_{10}$(relative error) against the $\log_{10}(n)$ for $\text{set}_i, i \in \{1, \cdots, 5\}$, where the relative error is defined by

$$\frac{\text{estimated } Y_0 \text{ by the proposed scheme} - \text{the value obtained from (9.3)}}{\text{the value obtained from (9.3)}}.$$

We scaled Δ of the central difference scheme according to (8.6). One observes that the convergence is more stable for smaller a. Interestingly, for all the results, the empirical convergence rate is close to or slightly higher than 1. This faster than expected convergence may be due to the following reasons; for the converging cases, the quantity in Assumption 4.1 (ii) is not only bounded by some constant C but rather bounded by decreasing sequence C_n, which is expected to make the constant in Theorem 4.2 also decreasing sequence with respect to n.

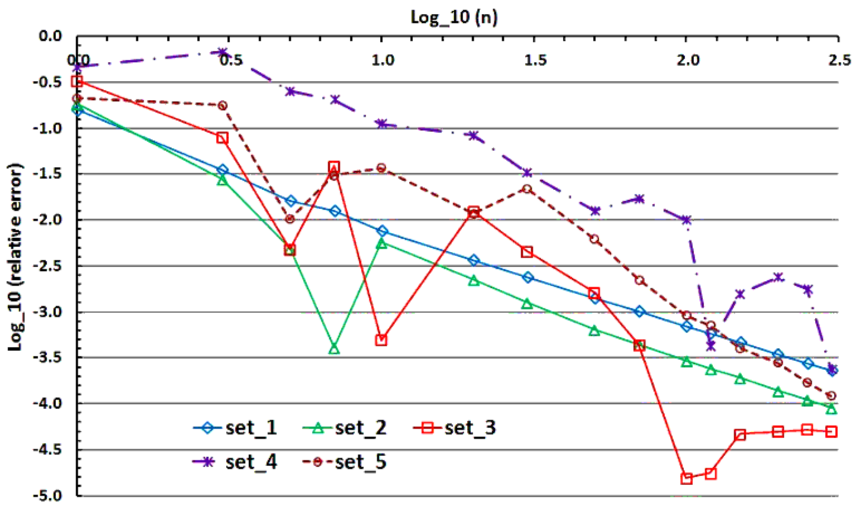

Figure 1: Empirical convergence of the proposed scheme for (9.2) with $\text{set}_i, i \in \{1, \cdots, 5\}$.

9.2 Stochastic quadratic coefficient

As a next example, let us make the coefficient of the quadratic term $|Z|^2$ stochastic as

$$Y_t = \xi(X_T) + \int_t^T \frac{a}{2}\sin\big(X_s^{\mathbf{1}} + X_s^{\mathbf{2}}\big)|Z_s|^2 ds - \int_t^T Z_s dW_s \ , \tag{9.6}$$

while keeping the same dynamics for X. We have chosen $x_0 = (1,1)^\top$, $T = 1.0$, $b_1 = b_2 = 0.05$, $\rho = 0.3$, $\sigma_1 = \sigma_2 = 0.5$ as the common parameters and tested the two cases; $\text{set}_1 := \{a = 1.0\}$ and $\text{set}_2 := \{a = 3.0\}$. The terminal function ξ is the same as the one used in the last model. In Figure 2, we have plotted the estimated Y_0 by changing $n = 1$ to 300. As expected, we observe that the setup with the bigger "a" needs a finer partition to converge.

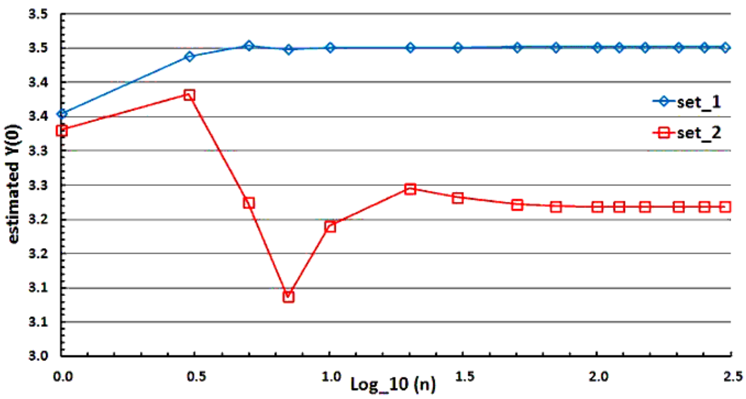

Figure 2: Empirical convergence of the proposed scheme for (9.6).

9.3 A truncation of the driver

As we have emphasized, it is crucial to have bounded derivatives given in Assumption 4.1 (i) for the proposed scheme to converge.[12] For the qg-BSDE (9.2), if we increase the coefficient "a" while keeping the factor C in (8.6) unchanged, we have observed that these derivatives (and hence the estimate of Y) diverge. In the remainder, instead of making C larger, let us study the truncation of the driver f so that it has a global Lipschitz constant N following the scaling rule (see Section 2.1 of [14])

$$N \propto n^\alpha, \quad 0 < \alpha < 1 \ . \tag{9.7}$$

The error estimates for the qg-BSDEs under this truncation have been studied by Imkeller & Reis (2010) [29] (Theorem 6.2) and applied to the backward numerical scheme by Chassagneux & Richou (2016) [14]. This truncation does not affect the theoretical bound on the convergence rate of Theorem 7.1, which is also the case for the scheme studied in [14].

We have chosen the constant C of (8.6) so that it marginally works for the set_3 in (9.5) without any truncation and tested the following four cases;

$$a = 6, \quad a = 8, \quad a = 10, \quad a = 12, \tag{9.8}$$

[12] At the moment, we do not have an explicit estimate of errors arising from the regularization that becomes necessary when the bounds of derivatives are breached.

with the other parameters are set equal to $x_0 = (1,1)^\top$, $T = 1$, $b_1 = b_2 = 0.05$, $\rho = 0.3$, $\sigma_1 = \sigma_2 = 0.5$ and the same terminal function (9.4). We have adopted the scaling factor $\alpha = 1/4$ for the truncation, which is the one used in the numerical examples given in [14]. In Figure 3, we have plotted the $\log_{10}$(relative error) against the $\log_{10}(n)$ changing the number of partitions from $n = 1$ to $n = 500$.

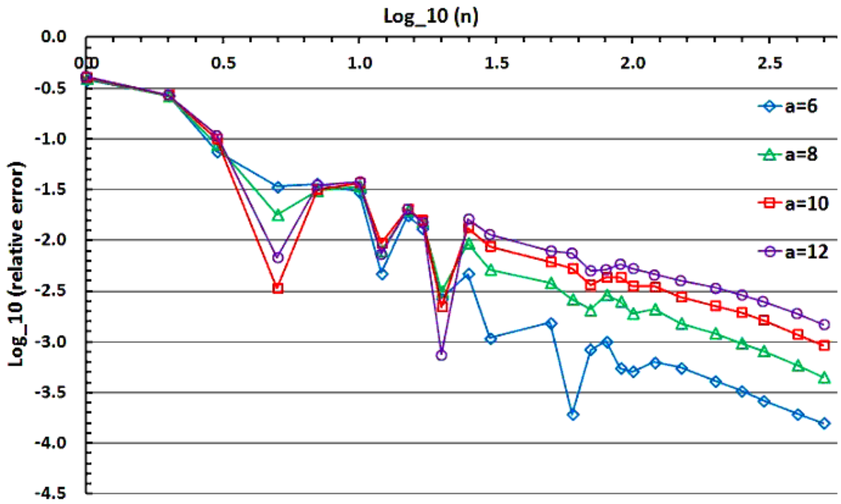

Figure 3: Empirical convergence of the proposed scheme for (10.2) with a truncated driver so that the Lipschitz constant scales as $N \propto n^{1/4}$.

Except for coarse partitions $n \lesssim 30$, the truncation of the driver yields quite stable convergence even for these extremely large quadratic coefficients. If there is no truncation, the calculation fails to converge in every example with the scaling factor C of (8.6) fixed as explained. We find no significant change in the empirical convergence rate, and it is close to one. All of these findings look consistent with the results implied by [14]. There seems a deep relation among the non-divergence of derivatives (Assumption 4.1 (i)), the scaling rule of central difference scheme (8.6), the scaling rule of the Lipschitz constant ($N \propto n^\alpha$), and the stability of the proposed scheme. This interesting problem requires further research.

10 Numerical examples (Lipschitz BSDEs)

10.1 Linear BSDE

As another consistency test, let us consider the following linear BSDE with $d = 2$:

$$Y_t = \xi(X_T) + \int_t^T \Big\{ C(X_s)Y_s + \gamma(X_s)Z_s \Big\} ds - \int_t^T Z_s dW_s \ , \tag{10.1}$$

where the process X follows the dynamics given in (9.1) and $\xi(x) := |x|^2 \exp\big(-0.1|x|^2\big)$, $C(x) := \cos(x^{\mathbf{1}}) + \sin(x^{\mathbf{2}})$, $\gamma(x) := \begin{pmatrix} \cos(x^{\mathbf{1}}) \\ \sin(x^{\mathbf{2}}) \end{pmatrix}$. Since γ is bounded, we can define the new measure $\mathbb{Q}$ by

$$\frac{d\mathbb{Q}}{d\mathbb{P}} = \mathcal{E}\Big(\int_0^T \gamma(X_s)^\top dW_s \Big) \ .$$

Under the new measure, we can show that the solution is given by

$$Y_t = \mathbb{E}^{\mathbb{Q}}\left[\exp\Bigl(\int_t^T C(X_s)ds\Bigr)\xi(X_T)\Big|\mathcal{F}_t\right] \ .$$

We have estimated Y_0 based on the above formula by simulating X under the measure $\mathbb{Q}$ with the parameters $\{x_0 = (1,1)^\top,\ T = 1,\ b_1 = b_2 = 0.05,\ \rho = 0.3,\ \sigma_1 = \sigma_2 = 0.5\}$. With 200,000 paths (half of which are antipathetic) of the step size $dt = 0.001$, we have obtained $Y_0 \simeq 8.934$ with the standard deviation 0.0057 from Monte Carlo simulation. We have estimated Y_0 by the proposed scheme with $n = 1$ to 300 and plotted $\log_{10}$(relative error) against $\log_{10}(n)$ in Figure 4. Here, the relative errors are calculated by treating $Y_0 = 8.934$ as the true solution. One observes a smooth convergence of the estimated Y_0.

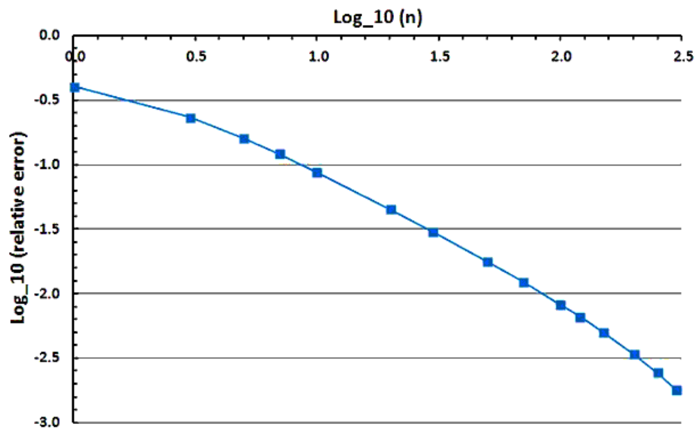

Figure 4: Empirical convergence of the proposed scheme for (10.1).

10.2 Option pricing with different interest rates

Finally, let us consider a very popular valuation problem of European options under two different interest rates, r for investing and R ($\neq r$) for borrowing. This problem has been often used for testing the numerical schemes for BSDEs.

Let us assume the dynamics of the security price as

$$X_t = x_0 + \int_0^t \mu X_s ds + \int_0^t \sigma X_s dW_s \ ,$$

where $d = 1$ and μ, σ are positive constants. For the option payoff $\Phi(X_T)$ at the expiry T, the option price Y_t implied by the self-financing replication is given by

$$Y_t = \Phi(X_T) - \int_t^T \left\{ rY_s + \frac{\mu - r}{\sigma} Z_s - \Bigl(Y_s - \frac{Z_s}{\sigma}\Bigr)^- (R - r) \right\} ds - \int_t^T Z_s dW_s \ . \qquad (10.2)$$

Although the BSDE is not smooth anymore, explicit mollifications for the payoff function and the driver may not be necessary as long as we use a finite difference scheme to approximate the derivatives. [13] Firstly, we study cases where the payoff function is equal to that of a call

[13] We have tested every case with the mollified functions. We have found no meaningful difference in the

option: $\Phi(x) = (x - K)^+$, where $K > 0$ is the strike price. As suggested by [28], this example provides a very interesting test since the price must be exactly equal to that of Black-Scholes model with interest rate R. This is because the replicating portfolio consists of the long-only position and hence the investor must always borrow money to fund her position. We have chosen the common parameters as $\{r = 0.01,\ R = 0.06,\ \mu = 0.06,\ X_0 = 100\}$ and tested the following five sets of (T, σ) [14] with $n = 10$ to $n = 3000$ in Figure 5:

$$\begin{aligned}
&\text{set}_1 = \{T = 1,\ K = 106,\ \sigma = 0.3\}\ , \text{set}_2 = \{T = 1,\ K = 166,\ \sigma = 0.3\} \\
&\text{set}_3 = \{T = 1,\ K = 106,\ \sigma = 1.0\}\ , \text{set}_4 = \{T = 1,\ K = 306,\ \sigma = 1.0\} \\
&\text{set}_5 = \{T = 1,\ K = 106,\ \sigma = 2.0\}.
\end{aligned}$$

The Black-Scholes price for each set is given by $\text{BS} = \{11.9999, 1.1171, 38.3459, 11.6619, 68.2964\}$ respectively. Although the relative errors for OTM options are slightly higher, the convergence rate to the exact BS prices is close to 1 for every case. It is a bit striking that we do not see any deterioration in convergence rate in spite of the non-smooth functions and rather high volatilities.

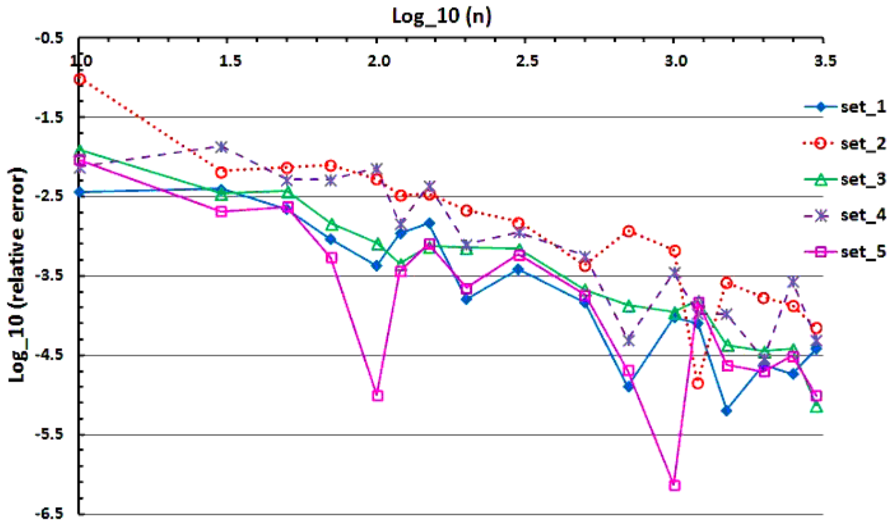

Figure 5: Empirical convergence of the proposed scheme for (10.2) for call options.

Next, let us consider a call-spread case: $\Phi(x) = (x - K_1)^+ - 2(x - K_2)^+$. This is exactly the same setup studied in [28] and hence we can test the consistency between our scheme and the standard regression-based Monte Carlo simulation. Let us choose the same parameter sets as in [28]:

$$\{r = 0.01,\ R = 0.06,\ \mu = 0.05,\ X_0 = 100,\ T = 0.25,\ \sigma = 0.2,\ K_1 = 95,\ K_2 = 105\} \quad (10.3)$$

The result of [28] suggests that $Y_0 = 2.96 \pm 0.01$ or $Y_0 = 2.95 \pm 0.01$ with one standard deviation dependent on the choice of basis functions for the regressions. In Figure 6, we have compared the estimated Y_0 from our scheme and the one in [28]. The dotted lines represent

empirical convergence.

[14] $K = 106$ is close to *at the money forward* for $T = 1$ with 6% interest rate. The bigger strikes correspond to 2σ *out of the money*.

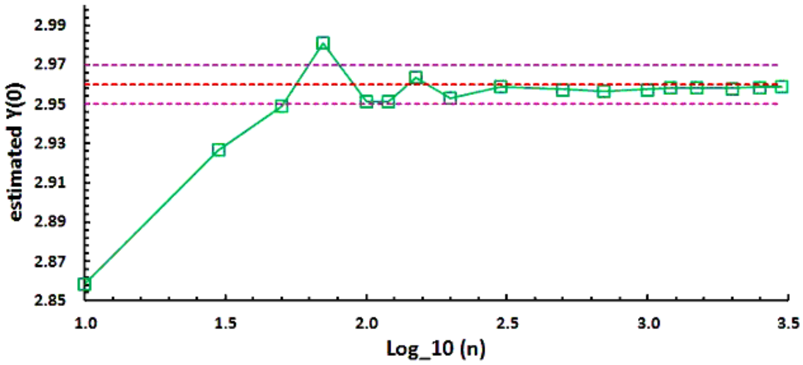

Figure 6: Empirical convergence of the proposed scheme for (10.2) for a call spread.

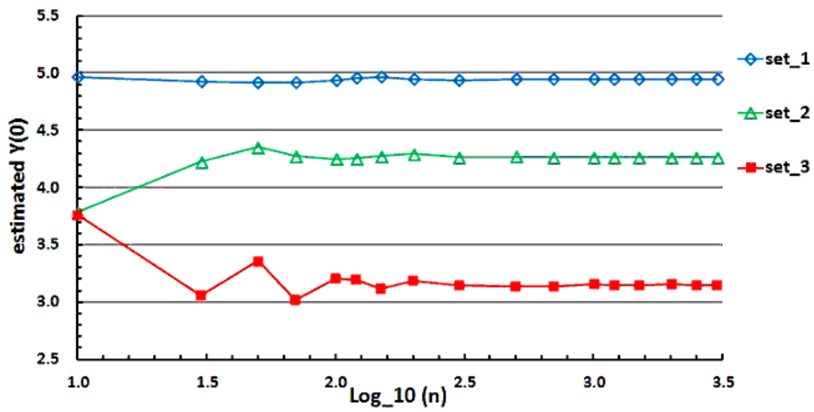

Figure 7: Empirical convergence of the proposed scheme for (10.2) for a call spread with $T = 1.0$ and higher volatilities.

2.96 ± 0.01 for ease of comparison. In our scheme, Y_0 converges toward 2.959. In fact, the improvement of the regression method of [28] using martingale basis functions proposed by Bender & Steiner (2012) [5] suggests 2.96 which is perfectly consistent with our result. We have also tested the convergence with a longer maturity and higher volatilities for the final payoff $\Phi(x) = (x - K_1)^+ - (x - K_2)^+$. We have used $\{r = 0.01,\ R = 0.06,\ \mu = 0.05,\ X_0 = 100,\ K_1 = 95,\ K_2 = 105\}$ as before, but with longer maturity $T = 1.0$ and $\mathrm{set}_1 := \{\sigma = 0.3\}$, $\mathrm{set}_2 := \{\sigma = 0.5\}$ and $\mathrm{set}_3 := \{\sigma = 1.0\}$. From Figure 7, one observes smooth convergence for all the cases. The decrease in price for higher volatilities is natural from the fact that K_2 is closer to the *at-the-money-forward* point and hence the short position has higher sensitivity on the volatility.

An example with a large Lipschitz constant

Bender & Steiner [5] have tested an extreme scenario with a parameter set (10.3) replaced by $R = 3.01$. In this case, the the non-linearity of the driver has a Lipschitz constant $(R-r)/\sigma = 15$. Their experiments suggest that the standard method of [28] fails to converge for this example. Their improved method with martingale basis functions (see Table 3 in [5]) gives $Y_0 \simeq 6.47$ with $n = 128$ and $Y_0 \simeq 6.44$ with the finest partition $n = 181$.

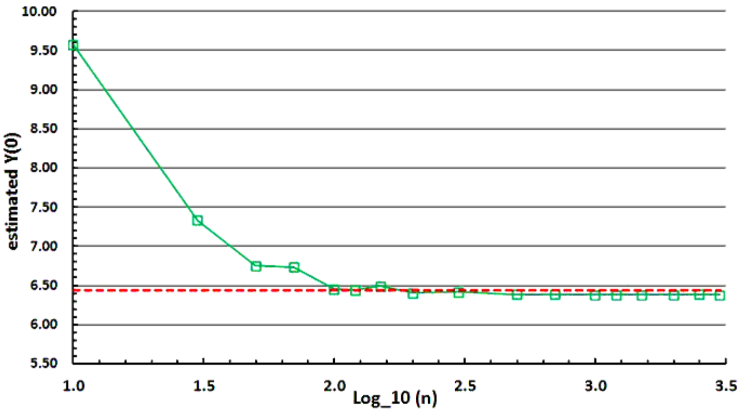

Figure 8: Empirical convergence of the proposed scheme for (10.2) with $R = 3.01$.

In Figure 8, we have plotted estimated Y_0 from our scheme with $n = 10$ to $n = 3000$. The dotted line corresponds to the value 6.44 given in [5]. In our scheme, Y_0 seems to converge 6.38. In particular, with the same discretization $n = 181$, our scheme yields $Y_0 \simeq 6.43$ showing a nice consistency. Note that the method [5] requires to change the basis functions based on the law of X. Every result of this subsection studying (10.2) is indicating that one may relax the smoothness conditions in Assumption 3.2, which will be studied in future research projects.

A Proof for Lemma 6.5

In this section, we give the proof for Lemma 6.5 skipped in the main text. We make Assumptions 3.1, 3.2 and Assumption 4.1 (i) the standing assumptions.

Proof. Firstly, let us consider $(\chi, \overline{\chi})$. Using 1/2-Hölder continuity in t, the global Lipschitz and linear growth properties of b in x, we have

$$\begin{aligned} |\chi(t,x) - \overline{\chi}(t,x)| &\le \int_{t_{i-1}}^t |b(r,\chi(r,x)) - b(t_{i-1},x)|dr \\ &\le K\int_{t_{i-1}}^t \Big[\Delta(r)^{1/2} + |\chi(r,x) - \overline{\chi}(r,x)| + \Delta(r)|b(t_{i-1},x)|\Big]dr \\ &\le C(1+|x|h_i^{1/2})h_i^{3/2} + K\int_{t_{i-1}}^t |\chi(r,x)-\overline{\chi}(r,x)|dr \end{aligned}$$

and hence by Gronwall inequality,

$$\sup_{t\in I_i} |\chi(t,x) - \overline{\chi}(t,x)| \le e^{Kh_i} C(1+|x|\sqrt{h_i})h_i^{3/2} \ .$$

Thus

$$\mathbb{E}\Big[\sup_{t\in I_i} |\chi(t,X_{t_{i-1}}) - \overline{\chi}(t,X_{t_{i-1}})|^p\Big] \le C_p h_i^{3p/2}\Big(1 + h_i^{p/2}\mathbb{E}\Big[|X_{t_{i-1}}|^p\Big]\Big) \le C_p h_i^{3p/2} \quad (A.1)$$

with some (i,n)-independent positive constant C_p.

Since $|\partial_x \widehat{u}^{i+1}| \le K'$ by Assumption 4.1 (i),

$$\begin{aligned}|y(t,x) - \overline{y}(t,x)| &= |\widehat{u}^{i+1}(\chi(t_i,x)) - \widehat{u}^{i+1}(\overline{\chi}(t_i,x))| \\ &\le K'|\chi(t_i,x) - \overline{\chi}(t_i,x)| \ .\end{aligned}$$

Thus from (A.1),

$$\mathbb{E}\left[\sup_{t\in I_i}|y(t,X_{t_{i-1}}) - \overline{y}(t,X_{t_{i-1}})|^p\right] \le C_p h_i^{3p/2} \tag{A.2}$$

with some (i,n)-independent positive constant C_p.

Let us now consider

$$\overline{\mathbf{y}}^{[1]}(t,x) = \partial_x \widehat{u}^{i+1}(\overline{\chi}(t_i,x)) + \delta(t)\big[\partial_x b(t_i,\overline{\chi}(t_i,x))\big]\partial_x \widehat{u}^{i+1}(\overline{\chi}(t_i,x)) \ .$$

Since both $|\partial_x \widehat{u}^{i+1}|$ and $|\partial_x b|$ are bounded, it is easy to see

$$\sup_{(t,x)\in I_i\times\mathbb{R}^d}|\overline{\mathbf{y}}^{[1]}(t,x)| \le C \tag{A.3}$$

with some positive constant C. For $t\in I_i$ with a given $x\in\mathbb{R}^d$, we have

$$\begin{aligned}\mathbf{y}^{[1]}(t,x) - \overline{\mathbf{y}}^{[1]}(t,x) &= \partial_x \widehat{u}^{i+1}(\chi(t_i,x)) - \partial_x \widehat{u}^{i+1}(\overline{\chi}(t_i,x)) \\ &\quad + \int_t^{t_i}\Big(\partial_x b(r,\chi(r,x))\mathbf{y}^{[1]}(r,x) - \partial_x b(t_i,\overline{\chi}(t_i,x))\overline{\mathbf{y}}^{[1]}(t_i,x)\Big)dr \ .\end{aligned}$$

From (A.3), 1/2-Hölder continuity and global Lipschitz property of $\partial_x b$, we obtain

$$\begin{aligned}&|\mathbf{y}^{[1]}(t,x) - \overline{\mathbf{y}}^{[1]}(t,x)| \\ &\le |\partial_x \widehat{u}^{i+1}(\chi(t_i,x)) - \partial_x \widehat{u}^{i+1}(\overline{\chi}(t_i,x))| + \int_t^{t_i}\Bigg\{\ |\partial_x b(r,\chi(r,x))||\mathbf{y}^{[1]}(r,x) - \overline{\mathbf{y}}^{[1]}(r,x)| \\ &\quad + |\partial_x b(r,\chi(r,x))||\overline{\mathbf{y}}^{[1]}(r,x) - \overline{\mathbf{y}}^{[1]}(t_i,x)| + |\partial_x b(r,\chi(r,x)) - \partial_x b(t_i,\overline{\chi}(t_i,x))||\overline{\mathbf{y}}^{[1]}(t_i,x)|\Bigg\}\, dr \\ &\le K'|\chi(t_i,x) - \overline{\chi}(t_i,x)| + K\int_t^{t_i}|\mathbf{y}^{[1]}(r,x) - \overline{\mathbf{y}}^{[1]}(r,x)|dr \\ &\quad + Ch_i^2 + C\int_t^{t_i}\Big(\delta(r)^{1/2} + |\chi(r,x) - \overline{\chi}(r,x)| + |\overline{\chi}(r,x) - \overline{\chi}(t_i,x)|\Big)dr \\ &\le K\int_t^{t_i}|\mathbf{y}^{[1]}(r,x) - \overline{\mathbf{y}}^{[1]}(r,x)|dr + Ch_i^{3/2}\big(1 + |x|\sqrt{h_i}\big) \ .\end{aligned}$$

Thus the backward Gronwall inequality (see, for example, Corollary 6.62 in [40]) gives

$$\sup_{t\in I_i}|\mathbf{y}^{[1]}(t,x) - \overline{\mathbf{y}}^{[1]}(t,x)| \le Ch_i^{3/2}(1 + |x|\sqrt{h_i})e^{Kh_i} \ ,$$

and hence

$$\mathbb{E}\left[\sup_{t\in I_i}|\mathbf{y}^{[1]}(t,X_{t_{i-1}})-\overline{\mathbf{y}}^{[1]}(t,X_{t_{i-1}})|^p\right]\le C_p h_i^{3p/2}\ , \tag{A.4}$$

with some (i,n)-independent constant C_p as desired.

By the boundedness of $|\partial_x^m \widehat{u}^{i+1}(x)|$ and $|\partial_x^m b|$ with $m\in\{1,2\}$, it is easy to see that $|\overline{G}^{[2]}|$ is also bounded

$$\sup_{(t,x)\in I_i\times\mathbb{R}^d}|\overline{G}^{[2]}(t,x)|\le C \tag{A.5}$$

with some positive constant C. Similar analysis done for $\overline{\mathbf{y}}^{[1]}$ using (A.5), 1/2-Hölder and Lipschitz continuities for $\partial_x b,\partial_x^2 b$, the backward Gronwall inequality yields

$$\sup_{t\in I_i}|G^{[2]}(t,x)-\overline{G}^{[2]}(t,x)|\le Ch_i^{3/2}(1+|x|\sqrt{h_i})\ ,$$

and hence

$$\mathbb{E}\left[\sup_{t\in I_i}|G^{[2]}(t,X_{t_{i-1}})-\overline{G}^{[2]}(t,X_{t_{i-1}})|^p\right]\le C_p h_i^{3p/2} \tag{A.6}$$

with some (i,n)-independent positive constant C_p as desired.

Finally, we consider

$$\overline{y}_0^{[2]}(t,x)=\delta(t)\mathrm{Tr}\Big(\overline{G}^{[2]}(t_i,x)[\sigma\sigma^\top]\big(t_i,\overline{\chi}(t_i,x)\big)\Big)\ .$$

From (A.5) and the linear-growth property of σ,

$$|\overline{y}_0^{[2]}(t,x)|\le C\delta(t)\big(1+|x|^2\big)\ , \tag{A.7}$$

is satisfied for every $(t,x)\in I_i\times\mathbb{R}^d$ with some positive constant C. We have

$$y_0^{[2]}(t,x)-\overline{y}_0^{[2]}(t,x)=\int_t^{t_i}\mathrm{Tr}\Big(G^{[2]}(r,x)[\sigma\sigma^\top](r,\chi(r,x))-\overline{G}^{[2]}(t_i,x)[\sigma\sigma^\top](t_i,\overline{\chi}(t_i,x))\Big)dr$$

and thus

$$\begin{aligned}
&|y_0^{[2]}(t,x)-\overline{y}_0^{[2]}(t,x)|\\
&\quad\le\int_t^{t_i}\mathrm{Tr}\Bigg\{\Big(|G^{[2]}(r,x)-\overline{G}^{[2]}(r,x)|+|\overline{G}^{[2]}(r,x)-\overline{G}^{[2]}(t_i,x)|\Big)\big|[\sigma\sigma^\top](r,\chi(r,x))\big|\\
&\qquad\qquad+\big|\overline{G}^{[2]}(t_i,x)\big|\Big|[\sigma\sigma^\top](r,\chi(r,x))-[\sigma\sigma^\top](t_i,\overline{\chi}(t_i,x))\Big|\Bigg\}dr\\
&\quad\le Ch_i^{3/2}(1+|x|)+Ch_i^2|x|^2(1+h_i|x|)
\end{aligned}$$

with some (i, n)-independent constant C. Thus we obtain, for any $p \geq 2$,

$$\mathbb{E}\left[\sup_{t \in I_i} |y_0^{[2]}(t, X_{t_{i-1}}) - \overline{y}_0^{[2]}(t, X_{t_{i-1}})|^p\right] \leq C_p h_i^{3p/2} \tag{A.8}$$

as desired. From (A.1), (A.2), (A.4), (A.6), (A.8) and $\mathbf{y}^{[2]} = \mathbf{y}^{[1]}$ we obtain the claim of Lemma 6.5. □

Acknowledgement

The research is partially supported by Center for Advanced Research in Finance (CARF).

References

[1] Ankirchner, S., Imkeller, P. and Dos Reis, G., 2007 *Classical and Variational Differentiability of BSDEs with Quadratic Growth*, Electronic Journal of Probability, Vol. 12, 1418-1453.

[2] Barrieu, P. and El Karoui, N., 2013, *Monotone stability of quadratic semimartingales with applications to unbounded general quadratic BSDEs*, The Annals of Probability, Vol. 41, No. 3B, 1831-1863.

[3] Barthelmann, V., Novak, E. and Ritter, K., 2000, *High dimensional polynomial interpolation on sparse grids*, Advances in Computational Mathematics 12, 273-288.

[4] Bender, C. and Denk, R., 2007, *A forward scheme for backward SDEs*, Stochastic Processes and their Applications, 117, 1793-1812.

[5] Bender, C. and Steiner, J., 2012, *Least-squares Monte Carlo for Backward SDEs*, Numerical Methods in Finance (edited by Carmona et al.), 257-289, Springer, Berlin.

[6] Bally, V., and Pagès, G., 2003, *A quantization algorithm for solving discrete time multidimensional optimal stopping problems*, Bernoulli, 6, 1003-1049.

[7] Bianchetti, M and M Morini (eds.) (2013), Interest Rate Modeling after the Financial Crisis, *Risk Books, London.*

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

[9] Bouchard, B. and Chassagneux, J-F., 2008, *Discrete-time approximation for continuously and discretely reflected BSDEs*, Stochastic Processes and their Applications, 118, 2269-2293.

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

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

[12] Briand, P. and Confortola, F., 2008, *BSDEs with stochastic Lipschitz condition and quadratic PDEs in Hilbert spaces*, Stochastic Processes and their Applications, 118, pp. 818-838.

[13] Brigo, D, M Morini and A Pallavicini (2013), Counterparty Credit Risk, Collateral and Funding, *Wiley, West Sussex.*

[14] Chassagneux, J.F. and Richou, A., 2016, *Numerical Simulation of Quadratic BSDEs*, Annals of Applied Probabilities, Vol. 26, No. 1, 262-304.

[15] Chassagneux, J.F. and Richou, A., 2016, *Rate of convergence for the discrete-time approximation of reflected BSDEs arising in switching problems*, working paper, arXiv:1602:00015.

[16] Crépey, S, T Bielecki with an introductory dialogue by D Brigo (2014), Counterparty Risk and Funding, *CRC press, NY.*

[17] Crépey, S. and Song, S., 2015, *Counterparty Risk and Funding: Immersion and Beyond*, Working paper, hal-00989062.

[18] Crisan, D. and Monolarakis, K., 2014, *Second order discretization of backward SDEs and simulation with the cubature method*, The annals of Applied Probability, Vol. 24, No. 2, 652-678.

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

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

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

[22] El Karoui, N., Peng, S. and Quenez, M.C., 1997, *Backward stochastic differential equations in Finance*, Mathematical Finance, Vol. 7, No. 1, 1-71.

[23] 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).

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

[25] Fujii, M. and Takahashi, A., 2015, *Perturbative expansion technique for non-linear FBSDEs with interacting particle method*, Asia-Pacific Financial Markets, Vol. 22, 3, 283-304.

[26] Fujii, M. and Takahashi, A., 2015, *Quadratic-exponential growth BSDEs with Jumps and their Malliavin's Differentiability*, Working paper, CARF-F-376. Available in arXiv.

[27] Fujii, M. and Takahashi, A., 2015, *Asymptotic Expansion for Forward-Backward SDEs with Jumps*, Working paper, CARF-F-372. Available in arXiv.

[28] 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, Vol. 15, No. 3, 2172-2202.

[29] Imkeller, P. and Dos Reis, G., 2010, *Path regularity and explicit convergence rate for BSDEs with truncated quadratic growth*, Stochastic Processes and their Applications, 120, 348-379. *Corrigendum for Theorem 5.5, 2010, 120, 2286-2288.*

[30] Kazamaki, N., 1994, *Continuous exponential martingales and BMO*, Lecture Notes in Mathematics, vol. 1579, Springer-Verlag, Berlin.

[31] Kloeden, P. and Platen, E., 1992, *Numerical solution of stochastic differential equations*, Applications of Mathematics (New York) 23. Springer, Berlin.

[32] Kobylanski, M., 2000, *Backward stochastic differential equations and partial differential equations with quadratic growth*, The annals of probability, Vol. 28, No. 2, 558-602.

[33] Kunitomo, N. and Takahashi, A., 2003, *On Validity of the Asymptotic Expansion Approach in Contingent Claim Analysis*, Annals of Applied Probability, 13, No.3, 914-952.

[34] Ma, J., Protter, P. and Yong, J., 1994, *Solving forward-backward stochastic differential equations explicitly-a four step scheme*, Probability Theory and Related Fields, 98, 339-359.

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

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

[37] Ma, X. and Zabaras, N., 2009, *An adaptive hierarchical sparse grid collocation algorithm for the solution of stochastic differential equations*, Journal of Computational Physics, 228, 3084-3113.

[38] Pagès, G. and Sagna, A., 2015, *Improved error bounds for quantization based numerical schemes for BSDE and nonlinear filtering*, arXiv:1510:01048.

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

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

[41] Sauer, T., 1995, *Polynomial interpolation of minimal degree*, Numerische Mathematik, 78, 59-85.

[42] Yong, J. and Zhou, X.Y., 1999, *Stochastic Controls: Hamiltonian systems and HJB equations*, Springer, NY.

[43] Takahashi, A., 1999, *An Asymptotic Expansion Approach to Pricing Contingent Claims,* Asia-Pacific Financial Markets, 6, 115-151.

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

[45] Takahashi, A. and Yamada, T., 2016, *A weak approximation with asymptotic expansion and multidimensional Malliavin weights*, Annals of Applied Probability, Vol. 26, No. 2, 818-856.

[46] Zhang, G., Gunzburger, M. and Zhao, W., 2013, *A sparse-grid method for multi-dimensional backward stochastic differential equations*, Journal of Computational Mathematics, 31, pp. 221-248.

[47] Zhang, J., 2004, *A numerical scheme for BSDEs*, The Annals of Applied Probability, Vol 14, No. 1, 459-488.

[48] Zhang, J., 2001, *Some fine properties of backward stochastic differential equations*, Ph.D Thesis, Purdue University.