RUNGE–KUTTA SCHEMES FOR BACKWARD STOCHASTIC DIFFERENTIAL EQUATIONS

BY JEAN-FRANÇOIS CHASSAGNEUX[1] AND DAN CRISAN[2]

Imperial College London

We study the convergence of a class of Runge–Kutta type schemes for backward stochastic differential equations (BSDEs) in a Markovian framework. The schemes belonging to the class under consideration benefit from a certain stability property. As a consequence, the overall rate of the convergence of these schemes is controlled by their local truncation error. The schemes are categorized by the number of intermediate stages implemented between consecutive partition time instances. We show that the order of the schemes matches the number p of intermediate stages for $p \leq 3$. Moreover, we show that the so-called order barrier occurs at $p = 3$, that is, that it is not possible to construct schemes of order p with p stages, when $p > 3$. The analysis is done under sufficient regularity on the final condition and on the coefficients of the BSDE.

1. Introduction. Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ be a filtered probability space endowed with an $(\mathcal{F}_t)_{t \geq 0}$-adapted Brownian motion $(W_t)_{t \geq 0}$. On $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ we consider the triplet $(X, Y, Z) = \{(X_t, Y_t, Z_t), t \in [0, T]\}$ of $(\mathcal{F}_t)_{t \geq 0}$-adapted stochastic processes satisfying the following equations:

$$(1.1) \qquad X_t = X_0 + \int_0^t b(X_s) \, ds + \int_0^t \sigma(X_s) \, dW_s,$$

$$(1.2) \qquad Y_t = g(X_T) + \int_t^T f(Y_t, Z_t) \, dt - \int_t^T Z_t \, dW_t.$$

System (1.1)–(1.2) is called a (decoupled) forward-backward stochastic differential equation (FBSDE).

The process X, called the forward component of the FBSDE, is a d-dimensional diffusion satisfying a stochastic differential equation (SDE) with Lipschitz-continuous coefficients $b : \mathbb{R}^d \to \mathbb{R}^d$ and $\sigma : \mathbb{R}^d \to \mathbb{R}^d \times \mathbb{R}^d$.

The pair of processes (Y, Z) satisfy the backward stochastic differential equation (BSDE) (1.2). The process Y is a one-dimensional stochastic process with final condition $Y_T = g(X_T)$, where $g : \mathbb{R}^d \to \mathbb{R}$ is a differentiable function with continuous and bounded first derivative [i.e., $g \in C_b^1(\mathbb{R}^d)$]. The process $Z =$

$(Z^1, \ldots, Z^d)$ is a d-dimensional process, written, by convention, as a row vector. The function $f : \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}$ referred to as "the driver," is assumed to be Lipschitz continuous.[3],[4]

The existence and uniqueness of solutions of system (1.1)–(1.2) was first addressed by Pardoux and Peng in [16]. Since then, a large number of papers have been dedicated to the study of FBSDEs. In particular, it is well known that under the Lipschitz-continuity assumption of the coefficients, the following estimate holds true:

$$(1.3) \quad \mathbb{E}\left[\sup_{t \in [0,T]} |X_t|^p\right] + \mathbb{E}\left[\sup_{t \in [0,T]} |Y_t|^2 + \int_0^T |Z_s|^2 \, \mathrm{d}s\right] < \infty \qquad \forall p > 0.$$

Moreover, Pardoux and Peng showed in [15] that

$$Y_t = u(t, X_t), \qquad Z_t = \nabla u^\top(t, X_t)\sigma(X_t), \qquad t \in [0, T],$$

where $u \in C^{1,2}([0, T] \times \mathbb{R}^d)$ is the solution of the final value Cauchy problem

$$(1.4) \quad L^{(0)}u(t, x) = -f\big(u(t, x), \nabla u^\top(t, x)\sigma(x)\big), \qquad t \in [0, T), x \in \mathbb{R}^d,$$

$$(1.5) \quad u(T, x) = g(x), \qquad x \in \mathbb{R}^d$$

with $L^{(0)}$ defined to be the second order differential operator

$$(1.6) \qquad\qquad L^{(0)} = \partial_t + \sum_{i=1}^d b_i \partial_{x_i} + \frac{1}{2} \sum_{i=1}^d a_{ij} \partial_{x_i} \partial_{x_j}$$

and $a = (a_{ij}) = \sigma\sigma^\top$.

There is a vast literature dedicated to the approximation of solutions to stochastic differential equations. In particular, obtaining approximations of the distribution of the forward component X has been largely resolved in the last thirty years. One can refer to [9] and the references therein for a systematic study of numerical methods for approximating X. Such methods are classical by now. More recently, Kusuoka, Lyons, Ninomiya and Victoir [10–14] developed several numerical algorithms for approximating X based on Chen's iterated integrals expansion. These new algorithms generate an approximation of the solution of the SDE in the form of the empirical distribution of a cloud of particles with deterministic trajectories.

By comparison, there are very few numerical methods for approximating the backward component. In this paper, we introduce a large class of numerical schemes for approximating solutions of BSDEs. These schemes are based on the well-known Runge–Kutta methods for ODEs and include new high order schemes as well as existing low order schemes such as the classical extension of the Euler scheme to BSDEs; see, for example, [1, 2, 4, 6].

---

[3]These assumptions will be strengthened in the following section.

[4]For the reader's convenience, we only consider drivers depending on Y and Z; however, the results and the analysis provided here apply to drivers depending also on X.

The approximations presented below are associated to an arbitrary, but fixed, partition π of the interval $[0, T]$, $\pi = \{t_0 = 0 < \cdots < t_i < t_{i+1} < \cdots < t_n = T\}$. We denote $h_i = t_{i+1} - t_i$, $i = 0, \ldots, n-1$ and $|\pi| = \max_i h_i$. Let (Y_i, Z_i) be the approximation of (Y_{t_i}, Z_{t_i}) for $i = 1, \ldots, n$. The construction of the approximating process is done in a recursive manner, backwards in time. We describe in the following the salient features of the class of approximations considered in this paper.

DEFINITION 1.1.

(i) The terminal condition is given by the pair $(Y_n, Z_n) = (g(X_T), \nabla g^\top(X_T) \times \sigma(X_T))$.

(ii) For $i \leq n-1$, the transition from (Y_{i+1}, Z_{i+1}) to (Y_i, Z_i) involves q stages, with $q \geq 1$. Given $q+1$ positive coefficients $0 =: c_1 < c_2 \leq \cdots \leq c_j \leq \cdots c_q \leq c_{q+1} := 1$, we introduce the intermediate "instances" of computation $t_{i,j} := t_{i+1} - c_j h_i$, and define $(Y_{i,j}, Z_{i,j})$, $j = 1, \ldots, q+1$ as follows: by convention, $(Y_{i,1}, Z_{i,1}) = (Y_{i+1}, Z_{i+1})$ and $(Y_{i,q+1}, Z_{i,q+1}) = (Y_i, Z_i)$. Then, for $1 < j \leq q$,

$$
(1.7) \qquad Y_{i,j} = \mathbb{E}_{t_{i,j}}\left[Y_{i+1} + c_j h_i \sum_{k=1}^{j} a_{jk} f(Y_{i,k}, Z_{i,k})\right],
$$

$$
(1.8) \qquad Z_{i,j} = \mathbb{E}_{t_{i,j}}\left[H_j^i Y_{i+1} + h_i \sum_{k=1}^{j-1} \alpha_{jk} H_{j,k}^i f(Y_{i,k}, Z_{i,k})\right].
$$

Finally, the approximation at step (i) is given by

$$
(1.9) \qquad Y_i = \mathbb{E}_{t_i}\left[Y_{i+1} + h_i \sum_{j=1}^{q+1} b_j f(Y_{i,j}, Z_{i,j})\right],
$$

$$
(1.10) \qquad Z_i = \mathbb{E}_{t_i}\left[H_{q+1}^i Y_{i+1} + h_i \sum_{j=1}^{q} \beta_j H_{q+1,j}^i f(Y_{i,j}, Z_{i,j})\right].
$$

The coefficients $(a_{jk})_{1 \leq j,k \leq q}$, $(\alpha_{jk})_{1 \leq j,k \leq q}$, $(b_j)_{1 \leq j \leq q+1}$ and $(\beta_j)_{1 \leq j \leq q}$ take their values in $\mathbb{R}$ with a_{1j}, α_{1j}, $1 \leq j \leq q$ and a_{jk}, α_{jk}, $1 \leq j < k \leq q$ set to 0. Moreover, the following holds:

$$
(1.11) \qquad \sum_{k=1}^{j} a_{jk} = \sum_{k=1}^{j-1} \alpha_{jk} \mathbf{1}_{\{c_k < c_j\}} = c_j, \qquad j \leq q.
$$

The random variables H_j^i, $H_{j,k}^i$, $k \leq j$ are $\mathcal{F}_{t_{i,j}}$-measurable, for all $j \leq q+1$, $i < n$ and have the property that, for all $1 \leq k < j \leq q+1$, $i < n$,

$$
(1.12) \quad \mathbb{E}_{t_{i,j}}[H_j^i] = \mathbb{E}_{t_{i,j}}[H_{j,k}^i] = 0 \quad \text{and} \quad \mathbb{E}[h_i |H_j^i|^2] + \mathbb{E}[h_i |H_{j,k}^i|^2] \leq \Lambda,
$$

where Λ is a positive constant which does not depends on π.

Observe that Y_n, Z_n belong to $L^2(\mathcal{F}_{t_n})$, where for $t \in [0, T]$, $L^2(\mathcal{F}_t)$ is the space of $\mathcal{F}_t$-measurable random variables U such that $\mathbb{E}[|U|^2] < \infty$. This is an immediate consequence of estimates (1.3) and the fact that $g \in C_b^1$. Moreover, an easy (backward) induction proves that the schemes are well defined for $|\pi|$ small enough and that Y_i, Z_i belong to $L^2(\mathcal{F}_{t_i})$ for all $i \leq n$.

In the sequel, we will refer to the schemes defined above by specifying the H-coefficients and using the following tableau for the other coefficients:

$$
\begin{array}{c|cccc|ccc}
c_1 = 0 & a_{11} & \cdots & a_{1q} & 0 & \alpha_{11} & \cdots & \alpha_{1q} \\
\vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\
c_j & a_{j1} & \cdots & a_{jq} & 0 & \alpha_{j1} & \cdots & \alpha_{jq} \\
\vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\
c_q & a_{q1} & \cdots & a_{qq} & 0 & \alpha_{q1} & \cdots & \alpha_{qq} \\
\hline
c_{q+1} = 1 & b_1 & \cdots & b_q & b_{q+1} & \beta_1 & \cdots & \beta_q
\end{array}.
$$

This notation is a natural extension of the classical notation used in the ODEs framework; see, for example, [3].

If the scheme is explicit for the last stage, that is, $b_{q+1} = 0$, we will omit this column in the coefficients tableau. We will also generally omit the "0" coefficients in the tableau and use "*" to denote a coefficient whose value is arbitrary.

Finally, let us also introduce for later use

$$(1.13) \qquad \tilde{\alpha}_{jk} = \alpha_{jk}\mathbf{1}_{\{c_k < c_j\}} \quad \text{and} \quad \tilde{\beta}_j = \beta_j\mathbf{1}_{\{c_j < 1\}}.$$

1.1. *General formulation of one-step schemes.* It is convenient to rewrite the approximations defined above in a more general setting as follows.

DEFINITION 1.2 (One-step scheme).

(i) The terminal condition is given by a pair $(Y_n, Z_n) \in L^2(\mathcal{F}_T)$.
(ii) For $i \leq n - 1$, the transition from (Y_{i+1}, Z_{i+1}) to (Y_i, Z_i) is given by

$$(1.14) \qquad \begin{cases} Y_i = \mathbb{E}_{t_i}[Y_{i+1} + h_i \Phi_i^Y(t_{i+1}, Y_{i+1}, Z_{i+1}, h_i)], \\ Z_i = \mathbb{E}_{t_i}[H_{q+1}^i Y_{i+1} + h_i \Phi_i^Z(t_{i+1}, Y_{i+1}, Z_{i+1}, h_i)], \end{cases}$$

where Φ_i^Y, Φ_i^Z are functions from $\mathbb{R}_+ \times L^2(\mathcal{F}_{t_{i+1}}) \times L^2(\mathcal{F}_{t_{i+1}}) \times \mathbb{R}_+^*$ to $L^2(\mathcal{F}_{t_{i+1}})$, $0 \leq i \leq n - 1$.

REMARK 1.1. In the case of the scheme given in Definition 1.1, the functions Φ_i^Y, Φ_i^Z depend implicitly of the coefficients $(a_{jk})_{1 \leq j,k \leq q}$, $(\alpha_{jk})_{1 \leq j,k \leq q}$, $(b_j)_{1 \leq j \leq q+1}$ and $(\beta_j)_{1 \leq j \leq q}$ and the random variables $(H_j^i)_{1 \leq j \leq q+1}$, $(H_{j,k}^i)_{1 \leq j,k \leq q}$.

1.1.1. *Order of convergence.* The global error we investigate here is given by the pair $(\mathcal{E}_Y(\pi), \mathcal{E}_Z(\pi))$, where

$$\mathcal{E}_Y(\pi) := \max_{0 \leq i \leq n} \mathbb{E}\big[|Y_{t_i} - Y_i|^2\big],$$

$$\mathcal{E}_Z(\pi) := \sum_{i=0}^{n-1} h_i \mathbb{E}\big[|Z_{t_i} - Z_i|^2\big].$$

To control these errors we will use *the local truncation error* for the pair (Y, Z) defined as

$$(1.15) \quad \eta_i := \eta_i^Y + \eta_i^Z, \qquad (\eta_i^Y, \eta_i^Z) := \left(\frac{1}{h_i^2} \mathbb{E}\big[|Y_{t_i} - \hat{Y}_{t_i}|^2\big], \mathbb{E}\big[|Z_{t_i} - \hat{Z}_{t_i}|^2\big] \right)$$

with

$$(1.16) \quad \begin{cases} \hat{Y}_{t_i} := \mathbb{E}_{t_i}\big[Y_{t_{i+1}} + h_i \Phi_i^Y(t_{i+1}, Y_{t_{i+1}}, Z_{t_{i+1}}, h_i)\big], \\ \hat{Z}_{t_i} := \mathbb{E}_{t_i}\big[H_{q+1}^i Y_{t_{i+1}} + h_i \Phi_i^Z(t_{i+1}, Y_{t_{i+1}}, Z_{t_{i+1}}, h_i)\big]. \end{cases}$$

The *global truncation error* for a given grid π is given by

$$\mathcal{T}(\pi) := \mathcal{T}_Y(\pi) + \mathcal{T}_Z(\pi),$$

$$(1.17)$$

$$(\mathcal{T}_Y(\pi), \mathcal{T}_Z(\pi)) := \left(\sum_{i=0}^{n-1} h_i \eta_i^Y, \sum_{i=0}^{n-1} h_i \eta_i^Z \right),$$

where $\mathcal{T}_Y$ is the global truncation error for Y, and $\mathcal{T}_Z$ is the global truncation error for Z defined as above.

The main results of the paper refer to the rate of convergence of the various approximations belonging to the class described in Definition 1.1.

DEFINITION 1.3. An approximation is said to have a *global truncation error* of order m if we have

$$\mathcal{T}(\pi) \leq C|\pi|^{2m}$$

for all sufficiently smooth[5] solutions to (1.4)–(1.5) and all partitions π with sufficiently small mesh size.

REMARK 1.2. Observe that we consider the sum of the global truncation error for the Y component and the Z component to define the *order* of an approximation. It is clear that if one considers BSDEs where the driver f depends only on Y and is only interested in the error on the Y part, it would be more judicious to use only $\mathcal{T}_Y$ in the definition of the order of the method. But our goal here is to deal with the most general case, where f depends on both Y and Z.

---

[5]The required regularity assumptions will be stated in the theorems below.

1.1.2. *Stability.*   To connect the truncation error with the global approximation error, we introduce the notion of L^2-stability for the schemes given in Definition 1.2. By *stability* we mean—roughly speaking—that the outcome of the scheme is "reasonably" modified if we "reasonably" perturb the scheme.

We thus introduce a perturbed scheme,

$$
(1.18) \quad
\begin{cases}
\tilde{Y}_i = \mathbb{E}_{t_i}[\tilde{Y}_{i+1} + h_i \Phi^Y(t_i, h_i, \tilde{Y}_{i+1}, \tilde{Z}_{i+1}) + \zeta_i^Y], \\
\tilde{Z}_i = \mathbb{E}_{t_i}[H^i_{q+1}\tilde{Y}_{i+1} + h_i \Phi^Z(t_i, h_i, \tilde{Y}_{i+1}, \tilde{Z}_{i+1}, h_i) + \zeta_i^Z],
\end{cases}
$$

where ζ_i^Y, ζ_i^Z belongs to $L^2(\mathcal{F}_{t_{i+1}})$, for all $i < n$ and with terminal values $\tilde{Y}_n$ and $\tilde{Z}_n$ belonging to $L^2(\mathcal{F}_T)$.

For $0 \le i \le n$, we denote $\delta Y_i := Y_i - \tilde{Y}_i$ and $\delta Z_i := Z_i - \tilde{Z}_i$ and consider the following definition of stability.

DEFINITION 1.4 (L^2-Stability).   The scheme given in Definition 1.2 is said to be L^2-stable if

$$
\max_i \mathbb{E}[|\delta Y_i|^2] + \sum_{i=0}^{n-1} h_i \mathbb{E}[|\delta Z_i|^2]
$$

$$
\le C\left(\mathbb{E}[|\delta Y_n|^2 + h_{n-1}|\delta Z_n|^2] + \sum_{i=0}^{n-1} h_i \mathbb{E}\left[\frac{1}{h_i^2}|\mathbb{E}_{t_i}[\zeta_i^Y]|^2 + |\mathbb{E}_{t_i}[\zeta_i^Z]|^2\right] \right)
$$

for all sequences ζ_i^Y, ζ_i^Z of $L^2(\mathcal{F}_{t_{i+1}})$-random variables and terminal values (Y_n, Z_n), $(\tilde{Y}_n, \tilde{Z}_n)$ belonging to $L^2(\mathcal{F}_T)$.

Under a reasonable assumption on the functions Φ_i^Y and Φ_i^Z, $i \le n-1$, introduced in (1.14), we are able to prove the stability of the schemes given in Definition 1.2.

THEOREM 1.1 (Sufficient condition for L^2-stability).   *Assume that, for some given grid π and for $i \le n-1$, we have*

$$
(1.19) \quad
\begin{aligned}
&\mathbb{E}_{t_i}\left[|\Phi_i^Y(t_{i+1}, U, V, h_i) - \Phi_i^Y(t_{i+1}, \tilde{U}, \tilde{V}, h_i)|^2\right] \\
&\quad \le C\left(\frac{1}{h_i}\left(\mathbb{E}_{t_i}[|\delta U|^2] - |\mathbb{E}_{t_i}[\delta U]|^2\right) + \mathbb{E}_{t_i}[|\delta U|^2 + |\delta V|^2]\right),
\end{aligned}
$$

$$
(1.20) \quad
\begin{aligned}
&\mathbb{E}_{t_i}\left[|\Phi_i^Z(t_{i+1}, U, V, h_i) - \Phi_i^Z(t_{i+1}, \tilde{U}, \tilde{V}, h_i)|^2\right] \\
&\quad \le \frac{C}{h_i}\left(\frac{1}{h_i}\left(\mathbb{E}_{t_i}[|\delta U|^2] - |\mathbb{E}_{t_i}[\delta U]|^2\right) + \mathbb{E}_{t_i}[|\delta U|^2 + |\delta V|^2]\right),
\end{aligned}
$$

where $U, V, \tilde{U}, \tilde{V}$ belong to $L^2(\mathcal{F}_{t_{i+1}})$, $\delta U := U - \tilde{U}$ and $\delta V := V - \tilde{V}$, then the scheme in Definition 1.2 is L^2-stable.

The following proposition connects the truncation error with the approximation error.

PROPOSITION 1.1. *Assume that the functions* Φ_i^Y *and* Φ_i^Z *satisfy* (1.19)–(1.20) *and* $(Y_n, Z_n) = (g(X_T), \nabla g^\top(X_T)\sigma(X_T))$. *Then there exists a constant* C *independent of the partition* π *such that*

$$(1.21) \qquad \mathcal{E}_Y(\pi) + \mathcal{E}_Z(\pi) \leq C\mathcal{T}(\pi).$$

The proofs of Theorem 1.1 and Proposition 1.1 are postponed to the Appendix.

1.1.3. *Convergence results.* As an application of Definitions 1.3 and 1.4, and Proposition 1.1, we state the following general convergence results (the proofs are postponed to the Appendix):

PROPOSITION 1.2. *If the method is of order* m *and* Φ_i^Y *and* Φ_i^Z *satisfy* (1.19)–(1.20) *and* $(Y_n, Z_n) = (g(X_T), \nabla g^\top(X_T)\sigma(X_T))$, *then there exists a constant* C *independent of the partition* π *such that*

$$(1.22) \qquad \mathcal{E}_Y(\pi) + \mathcal{E}_Z(\pi) \leq C|\pi|^{2m}.$$

Let us conclude this section with the main case of interest for us here, namely the Runge–Kutta schemes given in Definition 1.1.

THEOREM 1.2. (i) *For the schemes given in Definition* 1.1, *if* f *is Lipschitz-continuous, we have that the functions* Φ_i^Y *and* Φ_i^Z *satisfy* (1.19)–(1.20) *provided* $|\pi|$ *is small enough. As a result, the schemes are* L^2-*stable.*
(ii) *Moreover, if the method is of order* m, *then we have*

$$(1.23) \qquad \mathcal{E}_Y(\pi) + \mathcal{E}_Z(\pi) \leq C|\pi|^{2m},$$

provided $|\pi|$ *is small enough.*

REMARK 1.3. In this paper, we are only interested in obtaining an upper bound for the global approximation error $\mathcal{E}_Y(\pi) + \mathcal{E}_Z(\pi)$, in terms of $|\pi|$. An asymptotic expansion of this error in term of $|\pi|$ would also be of interest as it may lead to the use of Romberg–Richardson's extrapolation method. This work is left for future research.

1.2. *Order of convergence of Runge–Kutta methods.* It is a nontrivial task to classify the approximations belonging to the class described by Definition 1.1 through their order of convergence. The order of convergence of a particular scheme depends on several factors. First, it will depend on the number of intermediate steps it uses. Moreover, up to a certain level, the higher the smoothness

of the pair (u, f), the better the order is. However, there is a level of smoothness beyond which the order of approximation cannot typically be improved. This level is identified below through the condition $(\mathbf{H}r)_p$, where $p = 1, 2, \ldots$ is the number of intermediate steps required by the approximation. We show below that, provided the underlying framework satisfies a certain nondegeneracy condition called $(\mathbf{H}o)_p$, the order of the approximation cannot be improved through additional smoothness. This is achieved by identifying the leading order term in the expansion of the error of the approximation. However, should this leading order term be equal to zero, the order of the approximation will be higher. The analysis of the leading error term tells us that, for example, if the driver satisfies the additional constraint $f^z = 0$ (i.e., it is independent of Z, f^z denoting the partial derivative of f with respect to z), then there are two-stage schemes of order three. However, if $f^z \neq 0$, then two-stage schemes will typically have order two.

1.2.1. *Smoothness and nondegeneracy assumptions.* We study the order of the methods given in Definition 1.1 using Itô–Taylor expansions [9]. This requires the smoothness of the value function u. In order to state precisely these assumptions, we recall some notations of Chapter 5 (see Section 5.4) in [9].

Let

$$\mathcal{M} := \{\oslash\} \cup \bigcup_{m=1}^{\infty} \{0, \ldots, d\}^m$$

be the set of multi-indices with entries in $\{0, \ldots, d\}$ endowed with the measure ℓ of the length of a multi-index $[\ell(\oslash) = 0$ by convention$]$.

We introduce the concatenation operator $*$ on $\mathcal{M}$ for multi-indices with finite length $\alpha = (\alpha_1, \ldots, \alpha_p)$, $\beta = (\beta_1, \ldots, \beta_q)$ then $\alpha * \beta = (\alpha_1, \ldots, \alpha_p, \beta_1, \ldots, \beta_q)$.

For a multi-index α with positive finite length, we write $-\alpha$ (resp., $\alpha-$) the multi-index obtained by deleting the first (resp., last) component of α. On the set $\mathcal{M}$, let $n(\alpha)$ be the number of zero in a multi-index α with finite length.

Given a multi-index α, we denote by α^+ the multi-index obtained from α by deleting all its zero components.

For $j \in \{0, 1, \ldots, d\}$, we denote by $(j)_m$ the multi-index with length m and whose entries are all equal to j.

A nonempty subset $\mathcal{A} \subset \mathcal{M}$ is called a hierarchical set if

$$\sup_{\alpha} \ell(\alpha) < \infty \quad \text{and} \quad -\alpha \in \mathcal{A} \qquad \forall \alpha \in \mathcal{A} \setminus \{\oslash\}.$$

For any hierarchical $\mathcal{A}$ set, we consider the remainder set $\mathcal{B}(\mathcal{A})$ given by

$$\mathcal{B}(\mathcal{A}) := \{\alpha \in \mathcal{M} \setminus \mathcal{A} | -\alpha \in \mathcal{A}\}.$$

We will use in the sequel the following sets of multi-indices, for $n \geq 0$:

$$\mathcal{A}_n := \{\alpha | \ell(\alpha) \leq n\}$$

and observe that $\mathcal{B}(\mathcal{A}_n) = \mathcal{A}_{n+1} \setminus \mathcal{A}_n$.

For $j \in \{1, \ldots, d\}$, we consider the operators

$$L^{(j)} = \sum_{k=1}^{d} \sigma^{kj} \partial_{x_k}.$$

For a multi-index $\alpha = (\alpha_1, \ldots, \alpha_p)$, the iteration of these operators has to be understood in the following sense:

$$L^\alpha := L^{(\alpha_1)} \circ \cdots \circ L^{(\alpha_p)}.$$

By convention, $L^\varnothing$ is the identity operator; recall also the definition of the operator $L^{(0)}$ given in (1.6). One can observe that $L^{\alpha * \beta} = L^\alpha \circ L^\beta$.

Let C_b^k be the set of all k-times continuously differentiable functions with all partial derivatives bounded. For a multi-index with finite length α, we consider the set $\mathcal{G}^\alpha$ of all functions $v : [0, T] \times \mathbb{R}^d \to \mathbb{R}$ for which $L^\alpha v$ is well defined and continuous. We also introduce $\mathcal{G}_b^\alpha$ the subset of all functions $v \in \mathcal{G}^\alpha$ such that the function $L^\alpha v$ is bounded. For $v \in \mathcal{G}^\alpha$, we denote $L^\alpha v$ by v^α.

Finally, for $n \geq 1$, we define the set $\mathcal{G}_b^n$ of function v such that $v \in \mathcal{G}_b^\alpha$ for all $\alpha \in \mathcal{A}_n \setminus \{\varnothing\}$.

We are now ready to state the smoothness assumption on the value function u we shall use:

   $(\mathbf{H}r)_1$ The value function u belongs to $\mathcal{G}_b^2$ and $f \in C_b^1$.
   $(\mathbf{H}r)_2$ The value function u belongs to $\mathcal{G}_b^3$ and $f \in C_b^2$.
   $(\mathbf{H}r)_3$ The value function u belongs to $\mathcal{G}_b^4$ and $f \in C_b^3$.
   $(\mathbf{H}r)_4$ The value function u belongs to $\mathcal{G}_b^5$ and $f \in C_b^5$.

Instead of making assumptions on the coefficient b and σ, we shall use in the sequel the following "nondegeneracy" assumption when stating the necessary order conditions:

   $(\mathbf{H}o)_1$ There exists some function $g \in \mathcal{G}_b^2$ such that

$$\mathbb{P}\big(g^{(0)}(X_T) \neq 0\big) \neq 0.$$

   $(\mathbf{H}o)_2$ There exists some function $g \in \mathcal{G}_b^3$ such that

$$\mathbb{P}\big(g^\alpha(X_T) \neq 0\big) \neq 0$$

for $\alpha = (0), (0, 0)$ and $(j, 0)$ for some $j \in \{1, \ldots, d\}$. (Note that g may be different for each α.)

   $(\mathbf{H}o)_3$ There exists some function $g \in \mathcal{G}_b^4$ such that

$$\mathbb{P}\big(g^\alpha(X_T) \neq 0\big) \neq 0$$

for $\alpha = (0), (0, 0)$ and $(j_1, 0), (j_2, 0, 0)$ for some $(j_1, j_2) \in \{1, \ldots, d\}^2$. (Note that g may be different for each α.)

Moreover, for any triplet $(v_1, v_2, v_3) \neq (0, 0, 0)$ we have

$$\mathbb{P}\left(\left(v_1 g^{(0,0,0)} + v_2 f^y g^{(0,0)} + v_3 \sum_{\ell=1}^d f^{z^\ell} g^{(\ell,0,0)}\right)(X_T) \neq 0\right) \neq 0.$$

$(\mathbf{H}o)_4$ There exists some function $g \in \mathcal{G}_b^4$ such that

$$\mathbb{P}(g^\alpha(X_T) \neq 0) \neq 0$$

for $\alpha = (0), (0,0), (j_1, 0), (j_2, 0, 0)$ for some $(j_1, j_2) \in \{1, \ldots, d\}^2$. (Note that g may be different for each α.)

Moreover, we have for pairs $(v_1, v_3) \neq (0, 0)$, $(v_2, v_4) \neq (0, 0)$,

$$\mathbb{P}\left(\left(v_1 g^{(0,0,0)} + v_3 \sum_{j=1}^d {}^j v_g\right)(X_T) \neq 0\right) \neq 0,$$

$$\mathbb{P}\left(\left(v_2 g^{(\ell,0,0,0)} + v_4 f^{z^\ell} \sum_{j=1}^d g^{(j,0,0)}\right)(X_T) \neq 0\right) \neq 0$$

for $1 \leq \ell \leq d$ and for any $(v_1, v_2, v_3, v_4) \neq (0, 0, 0, 0)$ we have

$$\mathbb{P}\left(\left(v_1 g^{(0,0,0,0)} + v_2 \sum_{j=1}^d {}^j v_g^{(0)} + v_3 \sum_{j=1}^d {}^j w_g + v_4 \sum_{\ell=1}^d \sum_{j=1}^d f^{z^\ell} {}^j v_g^{(\ell)}\right)(X_T) \neq 0\right) \neq 0,$$

where we defined ${}^j v_g := f^{z^j} g^{(j,0,0)}$ and ${}^j w_g := f^{z^j} g^{(j,0,0,0)}$, $1 \leq j \leq d$.

REMARK 1.4.   If the Hörmander condition holds true, then all conditions $(\mathbf{H}o)_p$ are satisfied as the distribution of X_T has a smooth positive density with respect to the Lebesgue measure.

1.2.2. *Description of the H-coefficients.*   We now specify the class of random variables H used in the Definition 1.1 of the numerical schemes.

DEFINITION 1.5.   (i) For $m \geq 0$, we denote by $\mathcal{B}_{[0,1]}^m$ the set of bounded measurable functions $\psi : [0, 1] \to \mathbb{R}$ satisfying

$$\int_0^1 \psi(u)\,du = 1 \quad \text{and} \quad \text{if } m \geq 1, \qquad \int_0^1 \psi(u)u^k\,du = 0, \qquad 1 \leq k \leq m.$$

(ii) Let $(\psi^\ell)_{1 \leq \ell \leq d} \in \mathcal{B}_{[0,1]}^m$, for $t \in [0, T]$ and $h > 0$ such that $t + h \leq T$, we define

$$H_{t,h}^\psi := \left(\frac{1}{h} \int_t^{t+h} \psi^\ell\left(\frac{u-t}{h}\right) dW_u^\ell\right)_{1 \leq \ell \leq d},$$

which is a row vector.

By convention, we set $H_{t,0}^\psi = 0$.

For a discussion on the choice of the above coefficients, we refer to Remark 2.1 and Section 2.2.

1.2.3. *One-stage schemes.* We study here the order of the following family of schemes:

$$Y_i = \mathbb{E}_{t_i}\big[Y_{i+1} + h_i b_1 f(Y_{i+1}, Z_{i+1}) + h_i b_2 f(Y_i, Z_i)\big],$$

$$Z_i = \mathbb{E}_{t_i}\big[H_{t_i,h_i}^{\psi_1} Y_{i+1} + h_i \beta_1 H_{t_i,h_i}^{\phi_1} f(Y_{i+1}, Z_{i+1})\big],$$

where $\psi_1, \phi_1 \in \mathcal{B}_{[0,1]}^0$.

THEOREM 1.3. (i) *Assume that* $(\mathbf{H}r)_1$ *holds and that* $\psi_1, \phi_1 \in \mathcal{B}_{[0,1]}^0$. *For* $|\pi|$ *small enough, the above scheme is at least of order* 1 *if*

$$1 = b_1 + b_2.$$

Moreover, under $(\mathbf{H}o)_1$, *this condition is also necessary.*

(ii) *Assume that* $(\mathbf{H}r)_1$ *holds and that* $\psi_1 \in \mathcal{B}_{[0,1]}^1$, $\phi_1 \in \mathcal{B}_{[0,1]}^0$. *For* $|\pi|$ *small enough, the above scheme is at least of order* 2 *if*

$$b_1 = b_2 = \tfrac{1}{2} \quad and \quad \beta_1 = 1.$$

Moreover, under $(\mathbf{H}o)_2$, *this condition is also necessary.*

COROLLARY 1.1. *The above conditions lead to the following tableaux*:

$$\begin{array}{c|c|c} 0 & 0 & 0 \\ \hline 1 & 1 & * \end{array} \quad and \quad \begin{array}{c|c c|c} 0 & 0 & 0 & 0 \\ \hline 1 & 0 & 1 & * \end{array}$$

for the explicit Euler scheme and, respectively, the implicit version and to the tableau

$$\begin{array}{c|c c|c} 0 & 0 & 0 & 0 \\ \hline 1 & \tfrac{1}{2} & \tfrac{1}{2} & 1 \end{array}$$

for the Crank–Nicholson scheme.

REMARK 1.5. (i) The case of the Euler scheme has been widely studied in the literature. Generally speaking, as soon as f is Lipschitz-continuous, the method has been shown to be convergent. Under weak regularity assumption on the coefficient g, the order $\tfrac{1}{2}$ can be retrieved; see, for example, [2, 5, 7, 8, 17]. The order 1 convergence has been first proved in [6] for the general case when f depends on Z; see the references therein for the other cases.

(ii) The Crank–Nicholson scheme of step (ii) has been studied in the general case in [4]. It is proved there to be of order 2.

(iii) To the best of our knowledge, the necessary parts contained in Theorem 1.3 are new.

1.2.4. *Two-stage schemes.* We analyze here the order of the following family of schemes:

DEFINITION 1.6.

$$Y_{i,2} = \mathbb{E}_{t_{i,2}}\big[Y_{i+1} + a_{21}h_i f(Y_{i+1}, Z_{i+1})\big] + a_{22}h_i f(Y_{i,2}, Z_{i,2}),$$

$$Z_{i,2} = \mathbb{E}_{t_{i,2}}\big[H^{\psi_2}_{t_{i,2},c_2h_i}Y_{i+1} + c_2 h_i H^{\phi_2}_{t_{i,2},c_2h_i} f(Y_{i+1}, Z_{i+1})\big]$$

and

$$Y_i = \mathbb{E}_{t_i}\big[Y_i + h_i b_1 f(Y_{i+1}, Z_{i+1}) + h_i b_2 f(Y_{i,2}, Z_{i,2})\big] + h_i b_3 f(Y_i, Z_i),$$

$$Z_i = \mathbb{E}_{t_i}\big[H^{\psi_3}_{t_i,h_i}Y_{i+1} + \beta_1 H^{\phi_3}_{t_i,h_i} h_i f(Y_{i+1}, Z_{i+1}) + \beta_2 H^{\phi_3}_{t_i,(1-c_2)h_i} h_i f(Y_{i,2}, Z_{i,2})\big],$$

where $\phi_2, \phi_3, \psi_2, \psi_3 \in \mathcal{B}^0$.

The following results concern implicit schemes (for the Y part).

THEOREM 1.4.   (i) *Assume that* $(\mathbf{Hr})_3$ *holds,* $\psi_2, \psi_3 \in \mathcal{B}^2_{[0,1]}$, $\phi_2, \phi_3 \in \mathcal{B}^1_{[0,1]}$, $f^z = 0$ *and* $c_2 < 1$. *For* $|\pi|$ *small enough, the following conditions are sufficient to obtain at least an order* 3 *scheme*

$$b_1 = \frac{1}{2} - \frac{1}{6c_2}, \qquad b_2 = \frac{1}{6c_2(1-c_2)}, \qquad b_3 = \frac{1}{2} - \frac{1}{6(1-c_2)},$$

$$a_{21} = \frac{c_2}{2}, \qquad \beta_1 = 1 - \frac{1}{2c_2}, \qquad \beta_2 = \frac{1}{2c_2}.$$

(ii) *If, moreover,* $(\mathbf{Ho})_3$ *holds, these conditions are also necessary.*

(iii) (*Implicit order barrier*) *If* $f^z \neq 0$ *and* $(\mathbf{Ho})_3$ *holds, there is no order* 3 *methods in the class of the schemes given in Definition* 1.1 *with only two stages.*

COROLLARY 1.2.   (i) *For* $0 < c_2 < 1$, *the above conditions lead to the following tableau:*

$$
\begin{array}{c|ccc|cc}
0 & & & & * & \\
c_2 & \frac{c_2}{2} & \frac{c_2}{2} & 0 & c_2 & * \\
\hline
1 & \frac{1}{2} - \frac{1}{6c_2} & \frac{1}{6c_2(1-c_2)} & \frac{2-3c_2}{6(1-c_2)} & 1 - \frac{1}{2c_2} & \frac{1}{2c_2}
\end{array}.
$$

(ii) *Observe that if* $c_2 = \frac{2}{3}$, *then* $b_3 = 0$ *and the tableau has the following explicit form*:

$$
\begin{array}{c|cc|cc}
0 & 0 & 0 & * & 0 \\
\frac{2}{3} & \frac{1}{3} & \frac{1}{3} & \frac{2}{3} & * \\
\hline
1 & \frac{1}{4} & \frac{3}{4} & \frac{1}{4} & \frac{3}{4}
\end{array}.
$$

Part (iii) of the last theorem tells us that it is generally not possible to get an order 3 scheme with a two-stage scheme, even if it is implicit, as soon as we have $f^z \neq 0$. This result differs from the ODE case. This fact is not surprising since the schemes we consider are always explicit for the Z part. The explicit feature of the scheme and the related error, somehow propagates through f^z. This will also be the case for schemes with a higher number of stages. Since we are particularly interested in BSDEs with general drivers, we see then that there is no advantage in using implicit scheme instead of explicit ones. As a result, we concentrate from now on in studying explicit schemes only.

The next result concerns then explicit schemes and exhibits the similarity with the ODEs framework.

THEOREM 1.5. (i) *Assume that* $(\mathbf{H}r)_2$ *holds and* $\psi_2, \psi_3 \in \mathcal{B}^1_{[0,1]}$, ϕ_2, $\phi_3 \in \mathcal{B}^0_{[0,1]}$.
The scheme given in Definition 1.1 *is at least of order* 2 *if*

$$b_1 = 1 - \frac{1}{2c_2} \quad and \quad b_2 = \frac{1}{2c_2},$$

$$\beta_1 + \beta_2 \mathbf{1}_{\{c_2 < 1\}} = 1.$$

(ii) *Moreover, if* $(\mathbf{H}o)_2$ *holds, then the above conditions are necessary.*

It is easily checked that the above conditions leads to the following tableau: For $0 < c_2 \leq 1$,

0	0	0	0	0
c_2	c_2	0	c_2	$*$
1	$1 - \frac{1}{2c_2}$	$\frac{1}{2c_2}$	β_1	$1 - \beta_1$

with $\beta_1 = 1$ if $c_2 = 1$.

1.2.5. *Three-stage schemes.* We analyze next the order of the following family of schemes:

DEFINITION 1.7.

(1.24) $\quad Y_{i,2} = \mathbb{E}_{t_{i,2}}[Y_{i+1} + h_i c_2 f(Y_{i+1}, Z_{i+1})],$

(1.25) $\quad Z_{i,2} = \mathbb{E}_{t_{i,2}}[H^{\psi_2}_{t_{i,2},c_2 h_i} Y_{i+1} + h_i c_2 H^{\phi_2}_{t_{i,2},c_2 h_i} f(Y_{i+1}, Z_{i+1})],$

(1.26) $\quad Y_{i,3} = \mathbb{E}_{t_{i,3}}[Y_{i+1} + h_i a_{31} f(Y_{i+1}, Z_{i,k}) + h_i a_{32} f(Y_{i,2}, Z_{i,2})],$

(1.27)
$$Z_{i,3} = \mathbb{E}_{t_{i,3}}\big[H^{\psi_3}_{t_{i,3},c_3 h_i} Y_{i+1} + h_i \alpha_{31} H^{\phi_3}_{t_{i,3},c_3 h_i} f(Y_{i+1}, Z_{i+1})$$
$$+ h_i \tilde{\alpha}_{32} H^{\phi_3}_{t_{i,3},(c_3-c_2)h_i} f(Y_{i,2}, Z_{i,2})\big].$$

The approximation at step (i) is given by

$$(1.28) \quad Y_i = \mathbb{E}_{t_i}\big[Y_{i+1} + h_i\big(b_1 f(Y_{i+1}, Z_{i+1})$$
$$+ b_2 f(Y_{i,2}, Z_{i,2}) + b_3 f(Y_{i,3}, Z_{i,3})\big)\big],$$

$$Z_i = \mathbb{E}_{t_i}\big[H^{\psi_4}_{t_i,h_i} Y_{i+1}$$
$$(1.29) \quad + h_i\big(\beta_1 H^{\phi_4}_{t_i,h_i} f(Y_{i+1}, Z_{i+1}) + \tilde{\beta}_2 H^{\phi_4}_{t_i,(1-c_2)h_i} f(Y_{i,2}, Z_{i,2})$$
$$+ \tilde{\beta}_3 H^{\phi_4}_{t_i,(1-c_3)h_i} f(Y_{i,3}, Z_{i,3})\big)\big]$$

with $\psi_2, \psi_3, \psi_4 \in B^2_{[0,1]}$, $\phi_2, \phi_3, \phi_4 \in B^1_{[0,1]}$.

THEOREM 1.6. (i) *Assume that* $(\mathbf{H}r)_3$ *holds. The scheme given in Definition* 1.1 *is at least of order 3 if* $c_2 \neq 1$, $c_2 \neq c_3$, *and the following conditions hold true*:

$$b_1 + b_2 + b_3 = 1, \qquad b_2 c_2 + b_3 c_3 = \tfrac{1}{2},$$
$$b_2 c_2^2 + b_3 c_3^2 = \tfrac{1}{3}, \qquad b_3 a_{32} c_2 = b_3 \alpha_{32} c_2 = \tfrac{1}{6}$$

and

$$\beta_1 + \beta_2 + \beta_3 \mathbf{1}_{\{c_3 < 1\}} = 1,$$
$$\beta_2 c_2 + \beta_3 c_3 \mathbf{1}_{\{c_3 < 1\}} = \tfrac{1}{2}.$$

(ii) *Moreover, if* $(\mathbf{H}o)_3$ *holds, then the above conditions are necessary.*

REMARK 1.6. (i) If $c_2 = 1$, then $c_3 = 1$ and $\tilde{\beta}_2 = \tilde{\beta}_3 = 0$. Thus the approximation for Z reads

$$Z_i = \mathbb{E}_{t_i}\big[H^{\psi_4}_{t_i,h_i} Y_{i+1} + h_i \beta_1 H^{\phi_4}_{t_i,h_i} f(Y_{i+1}, Z_{i+1})\big].$$

As shown in last section, this approximation leads generally to an order 2 scheme only setting $\beta_1 = 1$.
   (ii) If $c_3 = c_2$, we obtain an order 2 scheme only as well.

Using [3] we get that

COROLLARY 1.3. (i) *Assume that* $c_2 \neq \tfrac{2}{3}$, $c_3 \notin \{c_2, \tfrac{2}{3}, 1\}$. *Then the above conditions lead to the following tableau*:

0	0	0	0
c_2	c_2	0	0
c_3	$\frac{c_3(3c_2-3c_2^2-c_3)}{c_2(2-3c_2)}$	$\frac{c_3(c_3-c_2)}{c_2(2-3c_2)}$	0
1	$\frac{-3c_3+6c_2c_3+2-3c_2}{6c_2c_3}$	$\frac{3c_3-2}{6c_2(c_3-c_2)}$	$\frac{2-3c_2}{6c_3(c_3-c_2)}$

0	0	0	0
c_2	c_2	$*$	0
c_3	$\frac{c_3(3c_2-3c_2^2-c_3)}{c_2(2-3c_2)}$	$\frac{c_3(c_3-c_2)}{c_2(2-3c_2)}$	$*$
1	β_1	$\frac{2c_3-1}{2(c_3-c_2)}-\frac{c_3}{c_3-c_2}\beta_1$	$\frac{c_3(1-2c_2)}{2c_3(c_3-c_2)}+\frac{c_2}{c_3-c_2}\beta_1$

.

(ii) *If $c_3 = 1$ and $c_2 \neq \frac{2}{3}$, then the above conditions lead to the following tableau*:

0	0	0	0	0	0	0
c_2	c_2	0	0	c_2	$*$	0
1	$\frac{(3c_2-3c_2^2-1)}{c_2(2-3c_2)}$	$\frac{1-c_2}{c_2(2-3c_2)}$	0	$\frac{(3c_2-3c_2^2-1)}{c_2(2-3c_2)}$	$\frac{1-c_2}{c_2(2-3c_2)}$	$*$
1	$\frac{6c_2-3c_2-1}{6c_2}$	$\frac{1}{6c_2(1-c_2)}$	$\frac{2-3c_2}{6(1-c_2)}$	$1-\frac{1}{2c_2}$	$\frac{1}{2c_2}$	$*$

.

1.2.6. *Order barriers.* As shown in the last sections, it is possible to derive explicit methods of order $p = 1, 2, 3$ using, respectively, $s = 1, 2, 3$ stages. These methods are optimal in the sense that $s < p$ is generally not possible and $s > p$ would lead to more computational effort.

In the ODEs framework, such a result is well known; see [3]. In fact, it is also known that it is possible to build explicit order 4 method using 4-stage schemes. A very interesting feature of explicit methods is that to retrieve an order p scheme with p strictly greater than 4, one needs to use $s > p$ stages. This last result is known as "explicit order barriers"; see, for example, Theorem 370B in [3]. Because ODEs are a special case of BSDEs, the same explicit barriers will be encountered for BSDEs.

This leaves open the case $s = p = 4$ for BSDEs. Theorem 1.7 below shows that generally $s > p$ already for $p = 4$ in the BSDEs framework. This means that the explicit barrier is encountered earlier for BSDEs than for ODEs.

Before stating the main result of this section, let us also recall part (iii) of Theorem 1.4, which reveals an implicit order barrier in the BSDEs framework.

PROPOSITION 1.3 (Implicit barrier). *Assume $(\mathbf{H}r)_3$ holds and $f^z \neq 0$, then there is no implicit order 3 two-stage scheme, under the nondegeneracy assumption $(\mathbf{H}o)_3$.*

THEOREM 1.7 (Explicit barrier). *We assume that $f^y = 0$ and $f^z \neq 0$. There is no explicit four stage methods in the class of methods given in Definition 1.1 which is of order 4, provided that $(\mathbf{H}r)_4$, $(\mathbf{H}o)_4$ hold and that the H-coefficients are given by $H_j^i := H_{t_i,j,c_jh_i}^{\psi_j}$ and $H_{j,k}^i := H_{t_i,j,(c_j-c_k)h_i}^{\phi_j}$ with $\psi_j \in \mathcal{B}_{[0,1]}^3$, $\phi_j \in \mathcal{B}_{[0,1]}^2$, $2 \leq j \leq 5$.*

REMARK 1.7. Theorem 1.7 can be extended to the case of $f^y \neq 0$ and $f^z \neq 0$. Indeed, the fact that $f^y \neq 0$ will add more constraints to the problem. Note, however, that $(\mathbf{H}o)_4$ would need to be reformulated in this case.

1.3. *Outline.* The rest of the paper is organized as follows. In Section 2, we present some preliminary results used to study the order of convergence. We also interpret the approximation of Z as the approximation of a proxy for Z in dimension $d = 1$. Sections 3–5 deal then with the proof of the order for scheme with 1, 2 and 3 stages. Section 6 is dedicated to the case of the four-stage methods and the proof of Theorem 1.7. Finally, the Appendix contains the proofs of the results in Section 1.1 and the proofs of the preliminary results.

1.4. *Notation.* In the sequel C is a positive constant whose value may change from line to line depending on T, d, Λ, X_0 but which does not depend on the choice of the partition π. We write C_p if it depends on some extra positive parameters p.

For $t \in \pi$, R a random variable and r a real number, the notation $R = O_t(r)$ means that $|R| \leq \lambda_t^\pi r$ where λ_t^π is a positive random variable satisfying

$$\mathbb{E}\big[|\lambda_t^\pi|^p\big] \leq C_p$$

for all $p > 0$, $t \in \pi$ and all partitions π.

The continuous and adapted process U belongs to $S^2([0, T])$ if

$$\mathbb{E}\Big[\sup_{s \in [0,T]} |U_s|^2\Big] < \infty.$$

Multiple Itô Integrals. For any process U in $S^2([0, T])$, we consider the following iterated Lebesgue–Itô integrals for a multi-index α with length l:

$$I_{t,s}^\alpha[U] := \begin{cases} U_s, & \text{if } l = 0, \\ \int_t^s I_{t,r}^{\alpha-}[U]\,\mathrm{d}r, & \text{if } l \geq 1 \text{ and } \alpha_l = 0, \\ \int_t^s I_{t,r}^{\alpha-}[U]\,\mathrm{d}W_r^j, & \text{if } l \geq 1 \text{ and } \alpha_l = j, 1 \leq j \leq d. \end{cases}$$

One can recursively check that these integrals are well defined and that $I^\alpha[I^\beta[\cdot]] = I^{\beta*\alpha}[\cdot]$. We will denote by $I_{t,r}^\alpha$ the multiple Itô Integrals of the constant process equal to one.

Abbreviation. For $t \in [0, T]$, we denote $v^\alpha(t, X_t)$ by v_t^α and $f^y(Y_t, Z_t)$ by f_t^y, where f^y is the partial derivatives of f with respect to the variable y. Similarly $f_t^z := f^z(Y_t, Z_t)$ where f^z is the partial derivative of f with respect to z.

2. Preliminaries.

2.1. *Itô–Taylor expansions.* The following proposition is Theorem 5.5.1 in [9] adapted to our context.

PROPOSITION 2.1. *Let $\mathcal{A}$ be a hierarchical set and $\mathcal{B}(\mathcal{A})$ the associated remainder set, for a function v belonging to $\mathcal{G}_b^\beta$ for all $\beta \in \mathcal{B}(\mathcal{A})$. Then*

$$v(t+h, X_{t+h}) = \sum_{\alpha \in \mathcal{A}} v_t^\alpha I_{t,t+h}^\alpha + \sum_{\beta \in \mathcal{B}(\mathcal{A})} I_{t,t+h}^\beta [v^\beta].$$

This leads to the following weak expansion formula:

PROPOSITION 2.2. *Let $m \geq 0$. Then for a function $v \in \mathcal{G}_b^{m+1}$,*

$$\mathbb{E}_t[v(t+h, X_{t+h})] = v_t + h v_t^{(0)} + \frac{h^2}{2} v_t^{(0,0)} + \cdots + \frac{h^m}{m!} v_t^{(0)m} + O_t(h^{m+1}).$$

We now state another key expansion for the results below based on Proposition 2.1 and Definition 1.5.

PROPOSITION 2.3. (i) *Let $m \geq 0$, for $\psi = (\psi^\ell)_{1 \leq \ell \leq d}$ with $\psi^\ell \in \mathcal{B}_{[0,1]}^m$, assuming that $v \in \mathcal{G}_b^{m+2}$, then*

$$\mathbb{E}_t[(H_{t,h}^\psi)^\ell v(t+h, X_{t+h})] = v_t^{(\ell)} + h v_t^{(\ell,0)} + \cdots + \frac{h^m}{m!} v_t^{(\ell)*(0)m} + O_t(h^{m+1}).$$

(ii) *For $\psi = (\psi^\ell)_{1 \leq \ell \leq d}$ with $\psi^\ell \in \mathcal{B}_{[0,1]}^0$, assuming that $v \in \mathcal{G}_b^1$, we have*

$$\mathbb{E}_t[(H_{t,h}^\psi)^\ell v(t+h, X_{t+h})] = O_t(1).$$

(iii) *If $L^{(0)} \circ L^{(\ell)} = L^{(\ell)} \circ L^{(0)}$, for $\ell \in \{1, \ldots, d\}$, then the expansion of* (i) *holds true for $\psi = (1, \ldots, 1)$.*

The proof of this proposition is postponed to the Appendix.

REMARK 2.1. (i) The expansion of Proposition 2.3(i) motivates the definition of the H-coefficient. Indeed, we will apply it to the functions u and $u^{(0)}$ and are able to cancel the low order term for a good choice of coefficients (α_{kj}), (β_j); see the computations of the next sections.

(ii) It is worth noticing that in the (very special) case where $L^{(0)} \circ L^{(\ell)} = L^{(\ell)} \circ L^{(0)}$ for $\ell \in \{1, \ldots, d\}$, one only needs to use in the definition of the scheme, H-coefficients built with the function $\psi = (1, \ldots, 1)$.

We conclude this paragraph by giving some examples of function ψ ($d = 1$).

EXAMPLE 2.1. (i) The function $\psi = \mathbf{1}_{[0,1]}$ belongs to $\mathcal{B}_{[0,1]}^0$.

(ii.a) The polynomial function $x \mapsto \psi(x) = 4 - 6x$ belongs to $\mathcal{B}_{[0,1]}^1$.

(ii.b) For $c \in (0, 1)$, the function $\psi = \frac{1}{c(c-1)} \mathbf{1}_{[1-c,1]} + \frac{c-2}{c-1} \mathbf{1}_{[0,1]}$ belongs to $\mathcal{B}_{[0,1]}^1$.

(iii) For $c, c' \in (0, 1)$, $c \neq c'$,

$$\psi = \frac{1 - c'}{c(1 - c)(c' - c)}\mathbf{1}_{[1-c,1]} + \frac{c - 1}{c'(1 - c')(c' - c)}\mathbf{1}_{[1-c',1]}$$

$$+ \left(1 + \frac{1}{(1 - c)} + \frac{1}{(1 - c')}\right)\mathbf{1}_{[0,1]}$$

belongs to $\mathcal{B}^2_{[0,1]}$.

2.2. *A class of proxy for Z.* The solution of the BSDE (1.2) consists in the pair process (Y, Z). Unlike Y, the second component is not "directly available" in (1.2) since it is defined as the integrand in the martingale part. However, we can use (1.2) to construct first a proxy for Z. As we shall see, the sequence of processes $(Z_i)_{i \leq n}$ are discrete-time approximation of this proxy. The results below are based on the expansion given in Proposition 2.3. The discussion in this section assumes $d = 1$.

DEFINITION 2.1.   For $m \geq 0$, let $\psi \in \mathcal{B}^m_{[0,1]}$

$$(2.1) \qquad\qquad Z^\psi_{t,h} := \mathbb{E}_t\left[H^\psi_{t,h}\int_t^{t+h} Z_u\, \mathrm{d}W_u\right].$$

For later use, we denote $H^\psi_{t,h}(u) = \mathbb{E}_u[H^\psi_{t,h}]$, $t \leq u \leq t + h$.

PROPOSITION 2.4.   *Let $m \geq 0$, and assume that $u \in \mathcal{G}^{m+2}_b$. For $\psi \in \mathcal{B}^m_{[0,1]}$, the following holds*:

$$Z_t = Z^\psi_{t,h} + O(h^{m+1}).$$

PROOF.   One observes that

$$Z^\psi_{t,h} = \frac{1}{h}\mathbb{E}_t\left[\int_t^{t+h} \psi\left(\frac{s - t}{h}\right)Z_s\, \mathrm{d}s\right] = \frac{1}{h}\mathbb{E}_t\left[\int_t^{t+h} \psi\left(\frac{s - t}{h}\right)u^{(1)}_s\, \mathrm{d}s\right].$$

Applying the expansion given in Proposition 2.2 to $u^{(1)}$ up to order m and using the assumption on ψ, we obtain

$$Z^\psi_{t,h} = \sum_{k=0}^m u^{(0)_k*(1)}_t \frac{1}{h}\int_t^{t+h} \psi\left(\frac{s - t}{h}\right)\frac{(s - t)^k}{k!}\, \mathrm{d}s + O_t(h^{m+1})$$

$$= \sum_{k=0}^m u^{(0)_k*(1)}_t \frac{h^k}{k!}\int_0^1 \psi(r)r^k\, \mathrm{d}r + O_t(h^{m+1})$$

$$= Z_t + O_t(|h|^{m+1}),$$

recalling that $Z_t = u^{(1)}_t$.   □

REMARK 2.2. Of course one can build other types of proxies for Z based on (2.1), for example, at $t = 0$,

$$\mathbb{E}\left[H_{0,h}^{\psi} \int_0^h Z_s \, \mathrm{d}W_s + \lambda_1 Z_h + \lambda_2 Z_{h/2} \cdots \right].$$

In this case, ψ will be required to satisfy different constraints in order to obtain the desired order of convergence.

It remains to derive the discrete-time approximation (Z_i).

Observe that, using (1.2),

$$
\begin{aligned}
(2.2) \quad Z_{t,h}^{\psi} &:= \mathbb{E}_t\left[H_{t,h}^{\psi} \int_t^{t+h} Z_u \, \mathrm{d}W_u \right] \\
&= \mathbb{E}_t\left[H_{t,h}^{\psi}\left(Y_{t+h} + \int_t^{t+h} f(Y_u, Z_u) \, \mathrm{d}u \right) \right].
\end{aligned}
$$

In [2, 6], the approximation of the Z process is given by

$$\bar{Z}_{t_i,h_i}^{\mathbf{1}} := \mathbb{E}_{t_i}[H_{t_i,h_i}^{\mathbf{1}} Y_{t_{i+1}}].$$

In order to obtain high-order approximation of the process Z, we discretize the integral term in the right-hand side in (2.2), with $t = t_i$. For $\psi \in \mathcal{B}_{[0,1]}^m$, $m \geq 1$, we will approximate this term by the following:

$$(2.3) \qquad h \sum_{j=1}^q \beta_j \mathbb{E}_{t_i}[H_{t_i,(1-c_j)h_i}^{\phi_j} f(Y_{t_{i,j}}, Z_{t_{i,j}})],$$

where the coefficients $\beta_j \in \mathbb{R}$ and the function ϕ_j belongs to $\mathcal{B}_{[0,1]}^{m-1}$, for $1 \leq j \leq q$.

REMARK 2.3. Alternatively, one can approximate directly

$$\mathbb{E}_t\left[H_{t_i,h_i}^{\psi} \int_{t_i}^{t_{i+1}} f(Y_u, Z_u) \, \mathrm{d}u \right] = \mathbb{E}_{t_i}\left[\int_{t_i}^{t_{i+1}} H_{t_i,h_i}^{\psi}(u) f(Y_u, Z_u) \, \mathrm{d}u \right]$$

by

$$h \sum_{j=1}^q \beta_j \mathbb{E}_{t_i}[H_{t_i,h_i}^{\psi}(t_{i,j}) f(Y_{t_{i,j}}, Z_{t_{i,j}})].$$

However, since generally $H_{t_i,h_i}^{\psi}(t_{i,j}) \neq H_{t_i,(1-c_j)h_i}^{\psi}$, one would then require stronger assumptions on the function ψ and the H-coefficient which, in turn, will lead to higher computational complexity.

The approximation given in (2.3) is still theoretical since it uses the true value $Y_{t_{i,j}}$ and $Z_{t_{i,j}}$. We need to introduce several stages to obtain approximations of these intermediate values.

3. One-stage schemes.

3.1. *Proof of Theorem* 1.3(i).   (1) We first compute the error expansion for the Z part of the scheme. By (1.16), we have, for $1 \leq \ell \leq d$,

$$
\begin{aligned}
\hat{Z}_{t_i}^{\ell} &:= \mathbb{E}_{t_i}\big[(H_{t_i,h_i}^{\psi_1})^{\ell} Y_{t_{i+1}} + h_i \beta_1 (H_{t_i,h_i}^{\phi_1})^{\ell} f(Y_{t_{i+1}}, Z_{t_{i+1}})\big] \\
&= \mathbb{E}_{t_i}\big[(H_{t_i,h_i}^{\psi_1})^{\ell} u_{t_{i+1}} - h_i \beta_1 (H_{t_i,h_i}^{\phi_1})^{\ell} u_{t_{i+1}}^{(0)}\big],
\end{aligned}
$$

recalling (1.4).

Using Proposition 2.3, we get

$$
(3.1) \qquad\qquad \hat{Z}_{t_i}^{\ell} = Z_{t_i}^{\ell} + O_{t_i}(|\pi|),
$$

since $u \in \mathcal{G}_b^2$, recalling $(\mathbf{H}r)_1$ and $\psi_1^{\ell}, \phi_1^{\ell} \in \mathcal{B}_{[0,1]}^0$.

This basically means that as soon as $\psi_1^{\ell} \in \mathcal{B}_{[0,1]}^0$, $1 \leq \ell \leq d$, the choice of β_1 is arbitrary. Indeed, by definition of the truncation error for the Z component [see (1.15)–(1.17)], we have

$$
\mathcal{T}_Z(\pi) = O(|\pi|),
$$

which is the order we aim to obtain.

(2a) We now compute the error expansion for the Y-part. First observe that

$$
\begin{aligned}
\hat{Y}_{t_i} &:= \mathbb{E}_{t_i}\big[Y_{t_{i+1}} + h_i b_1 f(Y_{t_{i+1}}, Z_{t_{i+1}}) + h_i b_2 f(\hat{Y}_{t_i}, \hat{Z}_{t_i})\big] \\
&= \mathbb{E}_{t_i}\big[Y_{t_{i+1}} + h_i b_1 f(Y_{t_{i+1}}, Z_{t_{i+1}}) + h_i b_2 f(Y_{t_i}, \hat{Z}_{t_i})\big] + h_i b_2 \delta f_{t_i},
\end{aligned}
$$

where $\delta f_{t_i} = f(\hat{Y}_{t_i}, \hat{Z}_{t_i}) - f(Y_{t_i}, \hat{Z}_{t_i})$. This leads to

$$
\hat{Y}_{t_i} := \mathbb{E}_{t_i}\big[u_{t_{i+1}} - h_i b_1 u_{t_{i+1}}^{(0)} - h_i b_2 u_{t_i}^{(0)}\big] + h_i b_2 \delta f_{t_i}.
$$

Using Proposition 2.2, we compute

$$
\hat{Y}_{t_i} := u_{t_i} + h_i(1 - b_1 - b_2) u_{t_i}^{(0)} + h_i b_2 \delta f_{t_i} + O_{t_i}(|\pi|^2).
$$

Since f is Lipschitz-continuous and $u^{(0)}$ bounded, we obtain for $|\pi|$ small enough that $\hat{Y}_{t_i} = Y_{t_i} + O_{t_i}(|\pi|)$ which implies that $\delta f_{t_i} = O_{t_i}(|\pi|)$ and thus

$$
(3.2) \qquad\qquad \hat{Y}_{t_i} := Y_{t_i} + h_i(1 - b_1 - b_2) u_{t_i}^{(0)} + O_{t_i}(|\pi|^2).
$$

The condition $b_1 + b_2 = 1$ is thus sufficient to retrieve at least an order-1 scheme.

(2b) Under $(\mathbf{H}o)_1$, this condition is also necessary.

Indeed, combining definition (1.15)–(1.17) and (3.2), we compute

$$
\mathcal{T}_Y(\pi) = \sum_{i=0}^{n-1} h_i |1 - b_1 - b_2|^2 \mathbb{E}\big[|u_{t_i}^{(0)}|^2\big] + O(|\pi|^2).
$$

Interpreting the sum in the last equation as a Riemann sum and taking the limit as $|\pi| \to 0$, we obtain

$$\lim_{|\pi|\downarrow 0} \mathcal{T}_Y(\pi) = |1 - b_1 - b_2|^2 \int_0^T \mathbb{E}\big[|u^{(0)}(t, X_t)|^2\big]\, dt.$$

If $(1 - b_1 - b_2)^2 \neq 0$, since the scheme must be of order 1, we must have

$$\int_0^T \mathbb{E}\big[|u^{(0)}(t, X_t)|^2\big]\, dt = 0$$

for solutions u of (1.4) such that $u \in \mathcal{G}_b^2$, recalling Definition 1.3. In particular, at $t = T$, since $t \mapsto \mathbb{E}[|u^{(0)}(t, X_t)|^2]$ is continuous, we get

$$\mathbb{E}\big[|g^{(0)}(X_T)|^2\big] = 0 \qquad \text{for all } g \in \mathcal{G}_b^2.$$

Under $(\mathbf{H}o)_1$, this yields a contradiction.

3.2. *Proof of Theorem* 1.3(ii). (1a) We first compute the expansion for the Z part. By definition [see (1.16)], we have, for $1 \leq \ell \leq d$,

$$\begin{aligned}
\hat{Z}_{t_i}^\ell &:= \mathbb{E}_{t_i}\big[(H_{t_i,h_i}^{\psi_1})^\ell Y_{t_{i+1}} + h_i \beta_1 (H_{t_i,h_i}^{\phi_1})^\ell f(Y_{t_{i+1}}, Z_{t_{i+1}})\big] \\
&= \mathbb{E}_{t_i}\big[(H_{t_i,h_i}^{\psi_1})^\ell u_{t_{i+1}} - h_i \beta_1 (H_{t_i,h_i}^{\phi_1})^\ell u_{t_{i+1}}^{(0)}\big].
\end{aligned}$$

Using Proposition 2.3, we have

$$(3.3) \qquad \hat{Z}_{t_i}^\ell = Z_{t_i}^\ell + h_i(1 - \beta_1) u_{t_i}^{(\ell,0)} + O_{t_i}(|\pi|^2)$$

since $u \in \mathcal{G}_b^3$ and $\psi_1^\ell \in \mathcal{B}_{[0,1]}^1$, $\phi_1^\ell \in \mathcal{B}_{[0,1]}^0$, $1 \leq \ell \leq d$.

Using a first-order Taylor expansion, this leads to

$$(3.4) \qquad f(Y_{t_i}, \hat{Z}_{t_i}) = -u_{t_i}^{(0)} + h_i(1 - \beta_1) \sum_{\ell=1}^d f_{t_i}^{z^\ell} u_{t_i}^{(\ell,0)} + O_{t_i}(|\pi|^2),$$

recalling that $f \in C_b^2$ under $(\mathbf{H}r)_2$.

From (3.3) we deduce that the condition $1 - \beta_1 = 0$ is sufficient to obtain $\mathcal{T}_Z(\pi) = O(|\pi|^2)$, recalling (1.15)–(1.17).

(1b) If we assume that $(\mathbf{H}o)_2$ holds, this condition is also necessary. Indeed, one computes that

$$\frac{\mathcal{T}_Z(\pi)}{|\pi|^2} = \sum_{i=0}^{n-1} h_i(1 - \beta_1)^2 \sum_{\ell=1}^d \mathbb{E}\big[|u_{t_i}^{(\ell,0)}|^2\big] + O(|\pi|^2)$$

for grids with constant mesh size.

Then by interpreting the sum in the last equation as a Riemann sum, we obtain

$$\lim_{|\pi|\downarrow 0} \frac{\mathcal{T}_Z(\pi)}{|\pi|^2} = (1 - \beta_1)^2 \int_0^T \sum_{\ell=1}^d \mathbb{E}\big[|u^{(\ell,0)}(t, X_t)|^2\big]\, dt,$$

where the limit is taken over the grids with constant mesh size. If $(1 - \beta_1)^2 \neq 0$, since we are looking at a scheme of order 2, we must have

$$\int_0^T \sum_{\ell=1}^d \mathbb{E}[|u^{(\ell,0)}(t, X_t)|^2]\, dt = 0$$

for the solution u of (1.4) such that $u \in \mathcal{G}_b^3$, recalling Definition 1.3. In particular, at $t = T$, since $t \mapsto \sum_{\ell=1}^d \mathbb{E}[|u^{(\ell,0)}(t, X_t)|^2]$ is continuous, we get

$$\sum_{\ell=1}^d \mathbb{E}[|g^{(\ell,0)}(X_T)|^2] = 0 \qquad \text{for all } g \in \mathcal{G}_b^3.$$

Under $(\mathbf{H}o)_2$, this yields a contradiction.

We assume now that the condition $\beta_1 = 1$ holds.

(2a) For the Y-part, we have

$$\hat{Y}_{t_i} := \mathbb{E}_{t_i}\big[Y_{t_{i+1}} + h_i b_1 f(Y_{t_{i+1}}, Z_{t_{i+1}}) + h_i b_2 f(\hat{Y}_{t_i}, \hat{Z}_{t_i})\big]$$

$$= \mathbb{E}_{t_i}\big[Y_{t_{i+1}} + h_i b_1 f(Y_{t_{i+1}}, Z_{t_{i+1}}) + h_i b_2 f(Y_{t_i}, \hat{Z}_{t_i})\big] + h_i b_2 \delta f_{t_i},$$

where $\delta f_{t_i} = f(\hat{Y}_{t_i}, \hat{Z}_{t_i}) - f(Y_{t_i}, \hat{Z}_{t_i})$.

Combining the last equality with (3.4) and recalling that $\beta_1 = 1$, we get

$$\hat{Y}_{t_i} := \mathbb{E}_{t_i}\big[u_{t_{i+1}} - h_i b_1 u_{t_{i+1}}^{(0)} - h_i b_2 u_{t_i}^{(0)}\big] + h_i b_2 \delta f_{t_i} + O_{t_i}(|\pi|^3).$$

Since $u \in \mathcal{G}_b^3$, we use Proposition 2.2 to compute

$$
\begin{aligned}
(3.5) \qquad \hat{Y}_{t_i} = Y_{t_i} &+ h_i(1 - b_1 - b_2)u_{t_i}^{(0)} + h_i^2(\tfrac{1}{2} - b_1)u_{t_i}^{(0,0)} \\
&+ h_i b_2 \delta f_{t_i} + O_{t_i}(|\pi|^3).
\end{aligned}
$$

We observe that $\hat{Y}_{t_i} = Y_{t_i} + O(|\pi|)$ which leads to

$$\delta f_{t_i} = O_{t_i}(|\pi|)$$

since f is Lipschitz-continuous.

Combining (3.5) with the last estimate, we obtain

$$\hat{Y}_{t_i} = Y_{t_i} + h_i(1 - b_1 - b_2)u_{t_i}^{(0)} + O_{t_i}(|\pi|^2).$$

The condition

$$(1 - b_1 - b_2) = 0$$

is sufficient to obtain a method at least of order 1.

(2b) Using the same arguments as in step (2b) of the proof of part (i) of Theorem 1.3, we obtain that this condition is necessary if $(\mathbf{H}o)_2$ holds.

(2c) We thus assume from now on that this condition holds, and we get

$$(3.6) \qquad\qquad \hat{Y}_{t_i} = Y_{t_i} + O_{t_i}(|\pi|^2),$$

which leads, since f is Lipschitz-continuous, to $\delta f_{t_i} = O_{t_i}(|\pi|^2)$. Inserting this estimate back into (3.5), we obtain

$$(3.7) \qquad \hat{Y}_{t_i} = Y_{t_i} + h_i^2 \left(\tfrac{1}{2} - b_1\right) u_{t_i}^{(0,0)} + O_{t_i}(|\pi|^3),$$

recalling that $b_1 + b_2 = 1$.

The condition $\frac{1}{2} - b_1 = 0$ is therefore sufficient to obtain a method at least of order 2.

(2d) If we assume that $(\mathbf{H}o)_2$ holds, this condition is also necessary. Indeed, one computes that

$$\frac{\mathcal{T}_Y(\pi)}{|\pi|} = \sum_{i=0}^{n-1} h_i \left(\frac{1}{2} - b_1\right)^2 \mathbb{E}[|u_{t_i}^{(0,0)}|^2] + O(|\pi|^2)$$

for grids π with constant mesh size.

Then, as the limit of a Riemann sum, we obtain that

$$\lim_{|\pi| \downarrow 0} \frac{\mathcal{T}_Y(\pi)}{|\pi|} = \left(\frac{1}{2} - b_1\right)^2 \int_0^T \mathbb{E}[|u^{(0,0)}(t, X_t)|^2]\, dt,$$

where the limit is taken over the grids with constant mesh size. If $\frac{1}{2} - b_1 \neq 0$, since the scheme must be of order 2, we must have

$$\int_0^T \mathbb{E}[|u^{(0,0)}(t, X_t)|^2]\, dt = 0$$

for solution u of (1.4) such that $u \in \mathcal{G}_b^3$, recalling Definition 1.3. In particular, at $t = T$, since $t \mapsto \mathbb{E}[|u^{(0,0)}(t, X_t)|^2]$ is continuous, we get

$$\mathbb{E}[|g^{(0,0)}(X_T)|^2] = 0 \qquad \text{for all } g \in \mathcal{G}_b^3.$$

Under $(\mathbf{H}o)_2$, this yields a contradiction and completes the proof of the theorem.

4. Two-stage schemes.

4.1. *Proof of Theorem* 1.4. (1a) We first compute the error expansion at the intermediary step (step $j = 2$), recalling that $(\mathbf{H}r)_3$ is in force.

For $1 \leq \ell \leq d$, we have that

$$\hat{Z}_{t_i,2}^\ell := \mathbb{E}_{t_i,2}\big[(H_{t_i,2,c_2 h_i}^{\psi_2})^\ell Y_{t_{i+1}} + h_i c_2 (H_{t_i,2,c_2 h_i}^{\phi_2})^\ell f(Y_{t_{i+1}}, Z_{t_{i+1}})\big]$$

$$= \mathbb{E}_{t_i,2}\big[(H_{t_i,2,c_2 h_i}^{\psi_2})^\ell u_{t_{i+1}} - h_i c_2 (H_{t_i,2,c_2 h_i}^{\phi_2})^\ell u_{t_{i+1}}^{(0)}\big].$$

Since $u \in \mathcal{G}_b^4$, we apply Proposition 2.3 and get, for $1 \leq \ell \leq d$,

$$(4.1) \qquad \hat{Z}_{t_i,2}^\ell = \frac{c_2^2}{2} h_i^2 u_{t_i,2}^{(\ell,0,0)} + O_{t_i,2}(|\pi|^3).$$

Using a first order Taylor expansion, we obtain

$$f(Y_{t_{i,2}}, \hat{Z}_{t_{i,2}}) = -u_{t_{i,2}}^{(0)} - \frac{c_2^2}{2} h_i^2 \sum_{\ell=1}^{d} f_{t_{i,2}}^{z^\ell} u_{t_{i,2}}^{(\ell,0,0)} + O_{t_{i,2}}(|\pi|^3),$$

recalling that $f \in C_b^2$.

(1b) For the Y-part, we have, denoting $\delta f_{t_{i,2}} = f(\hat{Y}_{t_{i,2}}, \hat{Z}_{t_{i,2}}) - f(Y_{t_{i,2}}, \hat{Z}_{t_{i,2}})$,

$$\hat{Y}_{t_{i,2}} := \mathbb{E}_{t_{i,2}} \big[Y_{t_{i+1}} + h_i a_{21} f(Y_{t_{i+1}}, Z_{t_{i+1}}) \big] + a_{22} h_i f(Y_{t_{i,2}}, \hat{Z}_{t_{i,2}}) + a_{22} h_i \delta f_{t_{i,2}}$$

$$= \mathbb{E}_{t_{i,2}} \big[u_{t_{i+1}} - h_i \big(a_{21} u_{t_{i+1}}^{(0)} + a_{22} u_{t_{i,2}}^{(0)} \big) \big] + a_{22} h_i \delta f_{t_{i,2}} + O_{t_{i,2}}(|\pi|^3).$$

Using Proposition 2.2, we compute

$$(4.2) \qquad \hat{Y}_{t_{i,2}} = Y_{t_{i,2}} + \Big(\frac{c_2^2}{2} - a_{21} c_2 \Big) h_i^2 u_{t_{i,2}}^{(0,0)} + a_{22} h_i \delta f_{t_{i,2}} + O_{t_{i,2}}(|\pi|^3),$$

recalling that $u \in \mathcal{G}_b^3$.

Since f is Lipschitz continuous, we get that $\delta f_{t_{i,2}} = O_{t_{i,2}}(|\pi|^2)$.
Inserting this estimate back into (4.2), we obtain

$$\hat{Y}_{t_{i,2}} = Y_{t_{i2}} + \Big(\frac{c_2^2}{2} - a_{21} c_2 \Big) h_i^2 u_{t_{i,2}}^{(0,0)} + O_{t_{i,2}}(|\pi|^3).$$

Combining a first-order Taylor expansion with the last equality and (4.1) leads to

$$(4.3) \qquad \begin{aligned} f(\hat{Y}_{t_{i,2}}, \hat{Z}_{t_{i,2}}) &= -u_{t_{i,2}}^{(0)} + \Big(\frac{c_2^2}{2} - a_{21} c_2 \Big) h_i^2 f_{t_{i2}}^y u_{t_{i,2}}^{(0,0)} \\ &\quad - \frac{c_2^2}{2} h_i^2 \sum_{\ell=1}^{d} f_{t_{i2}}^{z^\ell} u_{t_{i,2}}^{(\ell,0,0)} + O_{t_{i,2}}(|\pi|^3). \end{aligned}$$

(2a) We now study the error at the final step for the Z-part.
We compute the following expansion, for $1 \le \ell \le d$:

$$\begin{aligned} \hat{Z}_{t_i}^\ell &:= \mathbb{E}_{t_i} \big[(H_{t_i, h_i}^{\psi_3})^\ell Y_{t_{i+1}} + \beta_1 h_i (H_{t_i, h_i}^{\phi_3})^\ell f(Y_{t_{i+1}}, Z_{t_{i+1}}) \\ &\qquad\qquad + \tilde{\beta}_2 h_i (H_{t_i, (1-c_2)h_i}^{\phi_3})^\ell f(\hat{Y}_{t_{i,2}}, \hat{Z}_{t_{i,2}}) \big] \\ &= \mathbb{E}_{t_i} \big[(H_{t_i, h_i}^{\psi_3})^\ell u_{t_{i+1}} - \beta_1 h_i (H_{t_i, h_i}^{\phi_3})^\ell u_{t_{i+1}}^{(0)} - \tilde{\beta}_2 h_i (H_{t_i, (1-c_2)h_i}^{\phi_3})^\ell u_{t_{i,2}}^{(0)} \big] \\ &\quad + O_{t_i}(|\pi|^3), \end{aligned}$$

where we used (4.3), Proposition 2.3 and $(\mathbf{H}r)_3$, observing that $f^y u^{(0,0)}$ and $f^{z^\ell} u^{(\ell,0,0)}$, $1 \le \ell \le d$, belong to $\mathcal{G}_b^1$.

Using Proposition 2.3 again, we obtain, for $1 \le \ell \le d$,

$$
(4.4) \qquad
\begin{aligned}
\hat{Z}_{t_i}^\ell - Z_{t_i}^\ell &:= (1 - \beta_1 - \tilde{\beta}_2) h_i u_{t_i}^{(\ell,0)} \\
&\quad + \left(\tfrac{1}{2} - \beta_1 - (1 - c_2)\tilde{\beta}_2\right) h_i^2 u_{t_i}^{(\ell,0,0)} + O_{t_i}\left(|\pi|^3\right).
\end{aligned}
$$

(2b) For the local truncation error on the Z-part to be of order 2, recalling (1.17), it is clear, according to (4.4), that the following condition is sufficient:

$$
(4.5) \qquad 1 - \beta_1 - \tilde{\beta}_2 = 0.
$$

Similarly, to retrieve local truncation error on the Z-part to be of order 3, the following conditions are sufficient:

$$
(4.6) \qquad 1 - \beta_1 - \tilde{\beta}_2 = 0,
$$

$$
(4.7) \qquad c_2 \tilde{\beta}_2 - \tfrac{1}{2} = 0.
$$

(2c) We now prove that condition (4.5) is necessary to obtain an order 2 scheme under $(\mathbf{H}o)_2$, recalling that $(\mathbf{H}o)_3$ implies $(\mathbf{H}o)_2$. We compute, for grids with constant mesh size,

$$
\frac{\mathcal{T}_Z(\pi)}{|\pi|^2} = (1 - \beta_1 - \tilde{\beta}_2)^2 h_i \sum_{\ell=1}^d \mathbb{E}\big[|u_{t_i}^{(\ell,0)}|^2\big] + O_{t_i}\left(|\pi|^2\right)
$$

and then (Riemann sum)

$$
\lim_{|\pi| \downarrow 0} \frac{\mathcal{T}_Z(\pi)}{|\pi|^2} = (1 - \beta_1 - \tilde{\beta}_2)^2 \int_0^T \sum_{\ell=1}^d \mathbb{E}\big[|u^{(\ell,0)}(t, X_t)|^2\big]\, dt.
$$

If $(1 - \beta_1 - \tilde{\beta}_2)^2 \ne 0$, since the scheme must be of order 2, we must have

$$
\int_0^T \sum_{\ell=1}^d \mathbb{E}\big[|u^{(\ell,0)}(t, X_t)|^2\big]\, dt = 0.
$$

In particular, at $t = T$, since $t \mapsto \sum_{\ell=1}^d \mathbb{E}[|u^{(\ell,0)}(t, X_t)|^2]$ is continuous, we get

$$
\sum_{\ell=1}^d \mathbb{E}\big[|g^{(\ell,0)}(X_T)|^2\big] = 0 \qquad \text{for all } g \in \mathcal{G}_b^3.
$$

Under $(\mathbf{H}o)_2$, this yields a contradiction.

(2d) Under $(\mathbf{H}o)_3$, it is thus necessary that $(1 - \beta_1 - \tilde{\beta}_2)^2 = 0$ to retrieve an order 2 and a fortiori an order 3 schemes. The expansion error for the Z part reads then

$$
(4.8) \qquad \hat{Z}_{t_i}^\ell - Z_{t_i}^\ell := (c_2 \tilde{\beta}_2 - \tfrac{1}{2}) h_i^2 u_{t_i}^{(\ell,0,0)} + O_{t_i}\left(|\pi|^3\right), \qquad 1 \le \ell \le d.
$$

Using the same techniques as in step (2c), one will get that condition (4.7) is necessary to obtain an order 3 scheme under $(\mathbf{H}o)_3$.

(3) We study the error expansion on the Y part at the final step. We aim to obtain an order 3 scheme. From the definition of the truncation error, it is obviously necessary that the local truncation error for the Z part is of order 3. We work then under this condition [see step (2d)] and then we have

$$(4.9) \qquad f(Y_{t_i}, \hat{Z}_{t_i}) = -u_{t_i}^{(0)} + O_{t_i}(|\pi|^3).$$

For the Y-part, using (4.3) and (4.9), we have

$$\hat{Y}_{t_i} := \mathbb{E}_{t_i}\big[Y_{t_{i+1}} + h_i b_1 f(Y_{t_{i+1}}, Z_{t_{i+1}}) + h_i b_2 f(\hat{Y}_{t_i,2}, \hat{Z}_{t_i,2})\big]$$

$$+ h_i b_3 f(Y_{t_i}, \hat{Z}_{t_i}) + \delta f_{t_i}$$

$$= \mathbb{E}_{t_i}\bigg[u_{t_{i+1}} - h_i b_1 u_{t_{i+1}}^{(0)} - h_i b_1 u_{t_i,2}^{(0)}$$

$$+ \bigg(\frac{c_2^2}{2} - a_{21}c_2\bigg)h_i^3 f^y u_{t_i,2}^{(0,0)} - \frac{c_2^2}{2}h_i^3 \sum_{\ell=1}^{d} f^{z^\ell} u_{t_i,2}^{(\ell,0,0)}\bigg]$$

$$- h_i b_3 u_{t_i}^{(0)} + h_i \delta f_{t_i} + O_{t_i}(|\pi|^4).$$

Using Proposition 2.2 and $(\mathbf{H}r)_3$,

$$\hat{Y}_{t_i} = \mathbb{E}_{t_i}\big[u_{t_{i+1}} - h_i b_1 u_{t_{i+1}}^{(0)} - h_i b_1 u_{t_i,2}^{(0)} - h_i b_3 u_{t_i}^{(0)}\big]$$

$$+ \bigg(\frac{c_2^2}{2} - a_{21}c_2\bigg)h_i^3 f^y u_{t_i}^{(0,0)} - \frac{c_2^2}{2}h_i^3 \sum_{\ell=1}^{d} f^{z^\ell} u_{t_i}^{(\ell,0,0)} + h_i \delta f_{t_i} + O_{t_i}(|\pi|^4).$$

Using Proposition 2.2, we get

$$\hat{Y}_{t_i} - Y_{t_i} = (1 - b_1 - b_2 - b_3)h_i u_{t_i}^{(0)} + \bigg(\frac{1}{2} - b_1 - b_2(1 - c_2)\bigg)h_i^2 u_{t_i}^{(0,0)}$$

$$+ \bigg(\frac{1}{6} - \frac{b_1 + b_2(1 - c_2)^2}{2}\bigg)h_i^3 u_{t_i}^{(0,0,0)}$$

$$(4.10)$$

$$+ \bigg(\frac{c_2^2}{2} - a_{21}c_2\bigg)h_i^3 f^y u_{t_i}^{(0,0)}$$

$$- \frac{c_2^2}{2}h_i^3 \sum_{\ell=1}^{d} f^{z^\ell} u_{t_i}^{(\ell,0,0)} + h_i \delta f_{t_i} + O_{t_i}(|\pi|^4).$$

Using the last equation, we obtain that $\delta f_{t_i} = O_{t_i}(|\pi|)$, which leads to

$$\hat{Y}_{t_i} - Y_{t_i} = (1 - b_1 - b_2 - b_3)h_i u_{t_i}^{(0)} + O_{t_i}(|\pi|^2).$$

Under $(\mathbf{H}o)_3$, it appears then that the following condition is necessary to retrieve an order ≥ 1 scheme:

$$(4.11) \qquad b_1 + b_2 + b_3 = 1.$$

We then assume that this condition holds and obtain

$$\hat{Y}_{t_i} - Y_{t_i} = h_i \delta f_{t_i} + O_{t_i}(|\pi|^2).$$

We thus compute

$$\delta f_{t_i} = h_i f^y \delta f_{t_i} + O_{t_i}(|\pi|^2).$$

And for $|\pi|$ small enough, $\delta f_{t_i} = O_{t_i}(|\pi|^2)$. Inserting this into (4.10) and recalling that (4.11) is in force, we get that

$$\hat{Y}_{t_i} - Y_{t_i} = (\tfrac{1}{2} - b_1 - b_2(1 - c_2))h_i^2 u_{t_i}^{(0,0)} + O_{t_i}(|\pi|^3).$$

Under $(\mathbf{H}o)_3$, the condition $\tfrac{1}{2} = b_1 + b_2(1 - c_2)$ is then necessary to obtain an order 2 scheme, and we thus assume it holds. Arguing as before we now obtain $\delta f_{t_i} = O_{t_i}(|\pi|^3)$ and then

(4.12)
$$\hat{Y}_{t_i} - Y_{t_i} = \left(\frac{b_2(1 - c_2)c_2}{2} - \frac{1}{12}\right)h_i^3 u_{t_i}^{(0,0,0)} + \left(\frac{c_2^2}{2} - a_{21}c_2\right)h_i^3 f^y u_{t_i}^{(0,0)}$$
$$- \frac{c_2^2}{2}h_i^3 \sum_{\ell=1}^d f^{z^\ell} u_{t_i}^{(\ell,0,0)} + O_{t_i}(|\pi|^4).$$

(3b) If $f^{z^\ell} = 0$ for all $\ell \in \{1, \ldots, d\}$, one obtains that

$$b_2 = \frac{1}{6(1 - c_2)c_2} \quad \text{and} \quad a_{21} = \frac{c_2}{2}$$

are sufficient conditions for the methods to be of order 3.

Under $(\mathbf{H}o)_3$, these are also necessary conditions.

This completes the proof of (i) and (ii).

(4) To prove (iii), we use (4.12) again. We observe that under $(\mathbf{H}o)_3$, if $f^{z^\ell} \neq 0$ for some $\ell \in \{1, \ldots, d\}$, since $c_2 > 0$, the methods is at most of order 2.

4.2. *Proof of Theorem* 1.5.

PROOF.   The computation for the explicit case is almost the same—easier, in fact. The main difference comes from the fact that we are only interested in order 2 schemes. We thus need a bit less regularity. Following the step of the last proof, one then gets the following error expansion:

(4.13)
$$\hat{Y}_{t_i} - Y_{t_i} = (1 - b_1 - b_2)h_i u_{t_i}^{(0)}$$
$$+ (\tfrac{1}{2} - b_1 - b_2(1 - c_2))h_i^2 u_{t_i}^{(0,0)} + O_{t_i}(|\pi|^3)$$

and, for $\ell \in \{1, \ldots, d\}$,

(4.14)
$$\hat{Z}_{t_i}^\ell - Z_{t_i}^\ell := (1 - \beta_1 - \tilde{\beta}_2)h_i u_{t_i}^{(\ell,0)} + O_{t_i}(|\pi|^2).$$

Under $(\mathbf{H}r)_2$, the conditions

$$1 - b_1 - b_2 = 0, \qquad b_2 c_2 = \tfrac{1}{2}, \qquad 1 - \beta_1 - \tilde{\beta}_2 = 0$$

are obviously sufficient. Under $(\mathbf{H}o)_2$, using the same techniques as in steps (2c)–(2d) of the proof of Theorem 1.4, one proves that these conditions are necessary, which completes the proof of the theorem. $\square$

5. Three-stage schemes.

5.1. *Proof of Theorem* 1.6. (1a) We compute the error expansion at the intermediary step $j = 2$.

$$\hat{Y}_{t_{i,2}} := \mathbb{E}_{t_{i,2}}\big[Y_{t_{i+1}} + h_i c_2 f(Y_{t_{i+1}}, Z_{t_{i+1}})\big] = \mathbb{E}_{t_{i,2}}\big[u_{t_{i+1}} - h_i c_2 u_{t_{i+1}}^{(0)}\big],$$

$$Z_{t_{i,2}} := \mathbb{E}_{t_{i,2}}\big[H_{t_{i,2},c_2 h_i}^{\psi_2} Y_{t_{i+1}} + h_i c_2 H_{t_{i,2},c_2 h_i}^{\psi_2} f(Y_{t_{i+1}}, Z_{t_{i+1}})\big]$$

$$= \mathbb{E}_{t_{i,2}}\big[H_{t_{i,2},c_2 h_i}^{\psi_2} u_{t_{i+1}} - h_i c_2 H_{t_{i,2},c_2 h_i}^{\psi_2} u_{t_{i+1}}^{(0)}\big].$$

Under $(\mathbf{H}r)_3$, applying Propositions 2.2 and 2.3, we have

$$\hat{Y}_{t_{i,2}} = Y_{t_{i,2}} - \frac{c_2^2}{2} h_i^2 u_{t_{i,2}}^{(0,0)} + O_{t_{i,2}}(|\pi|^3),$$

$$\hat{Z}_{t_{i,2}}^\ell = Z_{t_{i,2}}^\ell - \frac{c_2^2}{2} h_i^2 u_{t_{i,2}}^{(\ell,0,0)} + O_{t_{i,2}}(|\pi|^3), \qquad \ell \in \{1, \ldots, d\},$$

$$f(\hat{Y}_{t_{i,2}}, \hat{Z}_{t_{i,2}}) := -u_{t_{i,2}}^{(0)} - \frac{c_2^2}{2} h_i^2\left(f^y u_{t_{i,2}}^{(0,0)} + \sum_{\ell=1}^d f_{t_{i,2}}^{z^\ell} u_{t_{i,2}}^{(\ell,0,0)}\right) + O_{t_{i,2}}(|\pi|^3).$$

(1b) Error expansion at step 3.

$$\hat{Y}_{t_{i,3}} := \mathbb{E}_{t_{i,3}}\big[Y_{t_{i+1}} + h_i a_{31} f(Y_{t_{i+1}}, Z_{t_{i+1}}) + h_i a_{32} f(Y_{t_{i,2}}, Z_{t_{i,2}})\big],$$

$$\hat{Z}_{t_{i,3}} := \mathbb{E}_{t_{i,3}}\big[H_{t_{i,3},c_3 h_i}^{\psi_3} Y_{t_{i+1}} + h_i \alpha_{31} H_{t_{i,3},c_3 h_i}^{\phi_3} f(Y_{t_{i+1}}, Z_{t_{i+1}})$$

$$+ h_i \tilde{\alpha}_{32} H_{t_{i,3},(c_3 - c_2)h_i}^{\phi_3} f(Y_{t_{i,2}}, Z_{t_{i,2}})\big].$$

With this definition and using step (1a), we compute

$$\hat{Y}_{t_{i,3}} = \mathbb{E}_{t_{i,3}}\big[u_{t_{i+1}} - h_i a_{31} u_{t_{i+1}}^{(0)} - h_i a_{32} u_{t_{i,2}}^{(0)}\big] + O_{t_{i,3}}(|\pi|^3),$$

which leads to, recalling $a_{31} + a_{32} = c_3$,

$$\hat{Y}_{t_{i,3}} = Y_{t_{i,3}} - \left(\frac{c_3^2}{2} - c_2 a_{32}\right) h_i^2 u_{t_{i,3}}^{(0,0)} + O_{t_{i,3}}(|\pi|^3).$$

Equivalently, we get, for $\ell \in \{1, \ldots, d\}$,

$$\hat{Z}_{t_{i,3}}^\ell = Z_{t_{i,3}}^\ell - \left(\frac{c_3^2}{2} - c_2 \tilde{\alpha}_{32}\right) h_i^2 u_{t_{i,3}}^{(\ell,0,0)} + O_{t_{i,3}}(|\pi|^3).$$

And we obtain

$$f(\hat{Y}_{t_{i,3}}, \hat{Z}_{t_{i,3}}) := -u_{t_{i,3}}^{(0)} - \left(\frac{c_3^2}{2} - c_2 a_{32}\right) h_i^2 f_{t_{i,3}}^y u_{t_{i,3}}^{(0,0)}$$

$$- \left(\frac{c_3^2}{2} - c_2 \tilde{\alpha}_{32}\right) h_i^2 \sum_{\ell=1}^d f_{t_{i,3}}^{z^\ell} u_{t_{i,3}}^{(\ell,0,0)}$$

$$+ O_{t_{i,3}}(|\pi|^3).$$

(1c) Error expansion at the final step for Z.

$$\hat{Y}_{t_i} := \mathbb{E}_{t_i}\left[Y_{t_{i+1}} + h_i b_1 f(Y_{t_{i+1}}, Z_{t_{i+1}}) + h_i b_2 f(\hat{Y}_{t_{i,2}}, \hat{Z}_{t_{i,2}}) + h_i b_3 f(\hat{Y}_{t_{i,3}}, \hat{Z}_{t_{i,3}})\right],$$

$$\hat{Z}_{t_i} := \mathbb{E}_{t_i}\left[H_{t_i,h_i}^{\psi_4} Y_{t_{i+1}} + h_i \beta_1 H_{t_i,h_i}^{\phi_4} f(Y_{t_{i+1}}, Z_{t_{i+1}})\right.$$

$$\left. + h_i \beta_2 H_{t_i,(1-c_2)h_i}^{\phi_4} f(\hat{Y}_{t_{i,2}}, \hat{Z}_{t_{i,2}}) + h_i \tilde{\beta}_3 H_{t_i,(1-c_3)h_i}^{\phi_4} f(\hat{Y}_{t_{i,3}}, \hat{Z}_{t_{i,3}})\right].$$

Using the results of step (1), we then compute, for $\ell \in \{1, \ldots, \ell\}$,

$$\hat{Z}_{t_i}^\ell = \mathbb{E}_{t_i}\left[(H_{t_i,h_i}^{\psi_4})^\ell u_{t_{i+1}} - h_i \beta_1 (H_{t_i,h_i}^{\phi_4})^\ell u_{t_{i+1}}^{(0)}\right.$$

$$\left. - h_i \beta_2 (H_{t_i,(1-c_2)h_i}^{\phi_4})^\ell u_{t_{i,2}}^{(0)} - h_i \tilde{\beta}_3 (H_{t_i,(1-c_3)h_i}^{\phi_4})^\ell u_{t_{i,3}}^{(0)}\right]$$

$$- \beta_2 \frac{c_2^2}{2} h_i^3 \mathbb{E}_{t_i}\left[(H_{t_i,(1-c_2)h_i}^{\phi_4})^\ell \left(f_{t_{i,2}}^y u_{t_{i,2}}^{(0,0)} + \sum_{j=1}^d f_{t_{i,2}}^{z^j} u_{t_{i,2}}^{(j,0,0)}\right)\right]$$

$$- \tilde{\beta}_3 \left(\frac{c_3^2}{2} - c_2 a_{32}\right) h_i^3 \mathbb{E}_{t_i}\left[(H_{t_i,(1-c_3)h_i}^{\phi_4})^\ell f_{t_{i,3}}^y u_{t_{i,3}}^{(0,0)}\right]$$

$$- \tilde{\beta}_3 \left(\frac{c_3^2}{2} - c_2 \tilde{\alpha}_{32}\right) h_i^3 \mathbb{E}_{t_i}\left[(H_{t_i,(1-c_3)h_i}^{\phi_4})^\ell \sum_{j=1}^d f_{t_{i,3}}^{z^j} u_{t_{i,3}}^{(j,0,0)}\right]$$

$$+ O_{t_i}(|\pi|^3).$$

Under $(\mathbf{H}r)_3$, since $f^y u^{(0,0)}, f^{z^j} u^{(j,0,0)} \in \mathcal{G}_b^1$, $j \in \{1, \ldots, d\}$, we obtain using Proposition 2.3(ii), for all $\ell \in \{1, \ldots, d\}$,

$$\mathbb{E}_{t_i}\left[(H_{t_i,(1-c_2)h_i}^{\phi_4})^\ell \left(f_{t_{i,2}}^y u_{t_{i,2}}^{(0,0)} + \sum_{j=1}^d f_{t_{i,2}}^{z^j} u_{t_{i,2}}^{(j,0,0)}\right)\right] = O_{t_i}(1),$$

$$\mathbb{E}_{t_i}\left[(H_{t_i,(1-c_3)h_i}^{\phi_4})^\ell f_{t_{i,3}}^y u_{t_{i,3}}^{(0,0)}\right] = O_{t_i}(1)$$

and

$$\mathbb{E}_{t_i}\left[(H_{t_i,(1-c_3)h_i}^{\phi_4})^\ell \sum_{j=1}^d f_{t_{i,3}}^{z^j} u_{t_{i,3}}^{(j,0,0)}\right] = O_{t_i}(1).$$

And then

$$\bar{Z}_{t_i} = \mathbb{E}_{t_i}\big[H^{\psi_4}_{t_i,h_i} u_{t_{i+1}} - h_i\beta_1 H^{\phi_4}_{t_i,h_i} u^{(0)}_{t_{i+1}}$$
$$- h_i\beta_2 H^{\phi_4}_{t_i,(1-c_2)h_i} u^{(0)}_{t_{i,2}} - h_i\tilde{\beta}_3 H^{\phi_4}_{t_i,(1-c_3)h_i} u^{(0)}_{t_{i,3}}\big]$$
$$+ O_{t_i}\big(|\pi|^3\big).$$

Using the expansion of Proposition 2.3, this leads to the following truncation error for the Z part:

$$\mathcal{T}_Z(\pi) := \sum_i (1 - \beta_1 + \beta_2 + \tilde{\beta}_3)^2 h_i^3 \sum_{\ell=1}^d \mathbb{E}\big[|u^{(\ell,0)}_{t_i}|^2\big]$$

(5.1)
$$+ \sum_i \Big(\frac{1}{2} - \beta_1 - \beta_2(1-c_2) - \tilde{\beta}_3(1-c_3)\Big)^2 h_i^5 \sum_{\ell=1}^d \mathbb{E}\big[|u^{(\ell,0,0)}_{t_i}|^2\big]$$

$$+ O\big(|\pi|^6\big).$$

(1d) Error expansion at the final step for Y.

$$\hat{Y}_{t_i} := \mathbb{E}_{t_i}\big[Y_{t_{i+1}} + h_i b_1 f(Y_{t_{i+1}}, Z_{t_{i+1}}) + h_i b_2 f(\hat{Y}_{t_{i,2}}, \hat{Z}_{t_{i,2}}) + h_i b_3 f(\hat{Y}_{t_{i,3}}, \hat{Z}_{t_{i,3}})\big].$$

We compute that

$$\hat{Y}_{t_i} = \mathbb{E}_{t_i}\big[u_{t_{i+1}} - h_i b_1 u^{(0)}_{t_{i+1}} - h_i b_2 u^{(0)}_{t_{i,2}} - h_i b_3 u^{(0)}_{t_{i,3}}\big]$$
$$- b_2 \frac{c_2^2}{2} h_i^3 \mathbb{E}_{t_i}\Bigg[\Bigg(f^y_{t_{i,2}} u^{(0,0)}_{t_{i,2}} + \sum_{\ell=1}^d f^{z^\ell}_{t_{i,2}} u^{(\ell,0,0)}_{t_{i,2}}\Bigg)\Bigg]$$
$$- b_3\Big(\frac{c_3^2}{2} - c_2 a_{32}\Big) h_i^3 \mathbb{E}_{t_i}\big[f^y_{t_{i,3}} u^{(0,0)}_{t_{i,3}}\big]$$
$$- b_3\Big(\frac{c_3^2}{2} - c_2\tilde{\alpha}_{32}\Big) h_i^3 \mathbb{E}_{t_i}\Bigg[\sum_{\ell=1}^d f^{z^\ell}_{t_{i,3}} u^{(\ell,0,0)}_{t_{i,3}}\Bigg]$$
$$+ O_{t_i}\big(|\pi|^4\big),$$

which leads to

$$\hat{Y}_{t_i} = \mathbb{E}_{t_i}\big[u_{t_{i+1}} - h_i b_1 u^{(0)}_{t_{i+1}} - h_i b_2 u^{(0)}_{t_{i,2}} - h_i b_3 u^{(0)}_{t_{i,3}}\big]$$
$$- \Big(b_2\frac{c_2^2}{2} + b_3\frac{c_3^2}{2} - b_3 c_2 a_{32}\Big) h_i^3 f^y u^{(0,0)}_{t_i}$$
$$- \Big(b_2\frac{c_2^2}{2} + b_3\frac{c_3^2}{2} - b_3 c_2\tilde{\alpha}_{32}\Big) h_i^3 \sum_{\ell=1}^d f^{z^\ell}_{t_i} u^{(\ell,0,0)}_{t_i}$$
$$+ O_{t_i}\big(|\pi|^4\big).$$

Using then Proposition 2.2, we obtain the following global truncation error for Y:

$$
\mathcal{T}_Y(\pi) = \sum_i h_i \mathbb{E}\Bigg[\Bigg|(1 - b_1 - b_2 - b_3)u_{t_i}^{(0)}
$$

$$
+ \left(\frac{1}{2} - b_1 - b_2(1 - c_2) - b_3(1 - c_3)\right)h_i u_{t_i}^{(0,0)}
$$

$$
+ \left(\frac{1}{6} - \frac{1}{2}b_1 - \frac{1}{2}b_2(1 - c_2)^2 - \frac{1}{2}b_3(1 - c_3)^2\right)h_i^2 u_{t_i}^{(0,0,0)}
$$

(5.2)

$$
- \left(b_2\frac{c_2^2}{2} + b_3\frac{c_3^2}{2} - b_3 c_2 a_{32}\right)h_i^2 f_{t_i}^y u_{t_i}^{(0,0)}
$$

$$
- \left(b_2\frac{c_2^2}{2} + b_3\frac{c_3^2}{2} - b_3 c_2 \tilde{\alpha}_{32}\right)h_i^2 \sum_{\ell=1}^d f_{t_i}^{z^\ell} u_{t_i}^{(\ell,0,0)}\Bigg|^2\Bigg]
$$

$$
+ O(|\pi|^6).
$$

(2a) If $c_3 \neq c_2$, According to steps (1c) and (1d), the conditions

$$
b_1 + b_2 + b_3 = 1, \qquad b_2 c_2 + b_3 c_3 = \tfrac{1}{2},
$$

$$
b_2 c_2^2 + b_3 c_3^2 = \tfrac{1}{3}, \qquad b_3 a_{32} c_2 = b_3 \tilde{\alpha}_{32} c_2 = \tfrac{1}{6}
$$

and

$$
\beta_1 + \beta_2 + \tilde{\beta}_3 = 1, \qquad \beta_2 c_2 + \tilde{\beta}_3 c_3 = \tfrac{1}{2}
$$

allow us to obtain an order 3 method, recalling that $c_2 \neq 1$.

Observe that the condition on β are weaker than on b and that $a_{32} = \alpha_{32}$. This equality, combined with the other condition on the coefficients, leads to $a_{jk} = \alpha_{jk}$, $1 \leq j, k \leq 3$.

(2b) Under $(\mathbf{H}o)_3$, using the same techniques, as, for example, in the proof of Theorem 1.3, one proves that the above conditions are necessary.

6. Four-stage schemes. This section is dedicated to the proof of Theorem 1.7. We now study the local truncation error for the family of scheme given by

(6.1) $\qquad Y_{i,2} = \mathbb{E}_{t_{i,2}}[Y_{i+1} + h_i c_2 f(Z_{i+1})],$

(6.2) $\qquad Z_{i,2} = \mathbb{E}_{t_{i,2}}[H_{t_{i,2},c_2 h_i}^{\psi_2} Y_{i+1} + h_i c_2 H_{t_{i,2},c_2 h_i}^{\phi_2} f(Z_{i+1})],$

(6.3) $\qquad Y_{i,3} = \mathbb{E}_{t_{i,3}}[Y_{i+1} + h_i a_{31} f(Z_{i+1}) + h_i a_{32} f(Z_{i,2})],$

$$
Z_{i,3} = \mathbb{E}_{t_{i,3}}\big[H_{t_{i,3},c_3 h_i}^{\psi_3} Y_{i+1}
$$

(6.4)

$$
+ h_i\big(\alpha_{31} H_{t_{i,3},c_3 h_i}^{\phi_3} f(Z_{i+1}) + \tilde{\alpha}_{32} H_{t_{i,3},(c_3-c_2)h_i}^{\phi_3} f(Z_{i,2})\big)\big],
$$

$$(6.5) \qquad Y_{i,4} = \mathbb{E}_{t_{i,4}}\big[Y_{i+1} + h_i a_{41} f(Z_{i+1}) + h_i a_{42} f(Z_{i,2}) + h_i a_{43} f(Z_{i,3})\big],$$

$$Z_{i,4} = \mathbb{E}_{t_{i,4}}\big[H^{\psi_4}_{t_{i,4},c_4 h_i} Y_{i+1}$$

$$(6.6) \qquad\qquad + h_i\big(\alpha_{41} H^{\phi_4}_{t_{i,4},c_4 h_i} f(Z_{i+1}) + \tilde\alpha_{42} H^{\phi_4}_{t_{i,4},(c_4-c_2)h_i} f(Z_{i,2})$$

$$+ \tilde\alpha_{43} H^{\phi_4}_{t_{i,4},(c_4-c_3)h_i} f(Z_{i,3})\big)\big].$$

The approximation at step (i) is given by

$$(6.7) \qquad Y_i = \mathbb{E}_{t_i}\big[Y_{i+1} + h_i\big(b_1 f(Z_{i+1}) + b_2 f(Z_{i,2}) + b_3 f(Z_{i,3}) + b_4 f(Z_{i,4})\big)\big],$$

$$Z_i = \mathbb{E}_{t_i}\big[H^{\psi_5}_{t_i,h_i} Y_{i+1}$$

$$(6.8) \qquad\qquad + h_i\big(\beta_1 H^{\phi_5}_{t_i,h_i} f(Z_{i+1}) + \tilde\beta_2 H^{\phi_5}_{t_i,(1-c_2)h_i} f(Z_{i,2})$$

$$+ \tilde\beta_3 H^{\phi_5}_{t_i,(1-c_3)h_i} f(Z_{i,3}) + \tilde\beta_4 H^{\phi_5}_{t_i,(1-c_4)h_i} f(Z_{i,4})\big)\big].$$

We assume that

$$a_{31} + a_{32} = c_3 \quad\text{and}\quad a_{41} + a_{42} + a_{43} = c_4,$$

$$\alpha_{31} + \tilde\alpha_{32} = c_3 \quad\text{and}\quad \alpha_{41} + \tilde\alpha_{42} + \tilde\alpha_{43} = c_4.$$

Moreover, $\psi_2, \psi_3, \psi_4, \psi_5 \in B^3_{[0,1]}$ and $\phi_2, \phi_3, \phi_4, \phi_5 \in B^2_{[0,1]}$.

We first prove that the following set of condition is necessary to retrieve an order 4 method:

LEMMA 6.1. *Assume that $c_2 \neq 1$ and $c_3 \neq 1$.*

(i) *The order 4 conditions for the Y-part are*

$$b_1 + b_2 + b_3 + b_4 = 1, \qquad b_3 \tilde\alpha_{32} c_2 + b_4 \tilde\alpha_{42} c_2 + b_4 \tilde\alpha_{43} c_3 = \tfrac{1}{6},$$

$$b_2 c_2 + b_3 c_3 + b_4 c_4 = \tfrac{1}{2}, \qquad b_3 \tilde\alpha_{32} c_2 c_3 + b_4 \tilde\alpha_{42} c_2 c_3 + b_4 \tilde\alpha_{43} c_3 c_4 = \tfrac{1}{8},$$

$$b_2 c_2^2 + b_3 c_3^2 + b_4 c_4^2 = \tfrac{1}{3}, \qquad b_3 \tilde\alpha_{32} c_2^2 + b_4 \tilde\alpha_{42} c_2^2 + b_4 \tilde\alpha_{43} c_3^2 = \tfrac{1}{12},$$

$$b_2 c_2^3 + b_3 c_3^3 + b_4 c_4^3 = \tfrac{1}{4}, \qquad b_4 \tilde\alpha_{43} \tilde\alpha_{32} c_2 = \tfrac{1}{24}.$$

(ii) *The order 4 conditions for the Z-part are*

$$\beta_1 + \beta_2 + \beta_3 = 1, \qquad \beta_2 c_2^2 + \beta_3 c_3^2 = \tfrac{1}{3},$$

$$\beta_2 c_2 + \beta_3 c_3 = \tfrac{1}{2}, \qquad \beta_3 \alpha_{32} c_3 = \tfrac{1}{6}.$$

REMARK 6.1. (i) If $c_2 = 1$, then $c_3 = c_4 = 1$ and $\beta_1 = 1$, the approximation for Z reads

$$Z_i = \mathbb{E}_{t_i}\big[H^{\psi_5}_{t_i,h_i} Y_{i+1} + h_i H^{\phi_5}_{t_i,h_i} f(Z_{i+1})\big],$$

which leads generally to an order 2 truncation error for Z.

(ii) If $c_2 \neq 1$ and $c_3 = 1$ (then $c_4 = 1$),

$$Z_i = \mathbb{E}_{t_i}\big[H^{\psi_5}_{t_i,h_i}Y_{i+1} + h_i\beta_1 H^{\phi_5}_{t_i,h_i}f(Z_{i+1}) + h_i\beta_2 H^{\phi_5}_{t_i,(1-c_2)h_i}f(Z_{i,2})\big],$$

which leads generally to an order 3 truncation error for Z.

PROOF OF LEMMA 6.1. (1) We first compute the error expansion at the intermediary steps. Observe that since we assume that f does not depends on Y, we only need to consider the approximation of Z for the intermediary stages.

(1a) Error expansion at step 2.

Under $(\mathbf{H}r)_4$, using Proposition 2.3(i), we have for $1 \leq \ell \leq d$,

$$(\hat{Z}_{t_{i,2}})^\ell = \mathbb{E}_{t_{i,2}}\big[(H^{\psi_2}_{t_{i,2},c_2 h_i})^\ell Y_{i+1} + h_i c_2 (H^{\phi_2}_{t_{i,2},c_2 h_i})^\ell f(Z_{t_{i+1}})\big]$$

$$= u^{(\ell)}_{t_{i,2}} - \frac{c_2^2}{2}h_i^2 u^{(\ell,0,0)}_{t_{i,2}} - \frac{c_2^3}{3}h_i^3 u^{(\ell,0,0,0)}_{t_{i,2}} + O_{t_{i,2}}(|\pi|^4),$$

which leads to

$$(6.9) \quad f(\hat{Z}_{t_{i,2}}) = -u^{(0)}_{t_{i,2}} - \frac{c_2^2}{2}h_i^2 \sum_{j=1}^{d} {}^j v_{t_{i,2}} - \frac{c_2^3}{3}h_i^3 \sum_{j=1}^{d} {}^j w_{t_{i,2}} + O_{t_{i,2}}(|\pi|^4),$$

where we set ${}^j v = f^{z^j} u^{(j,0,0)}$ and ${}^j w = f^{z^j} u^{(j,0,0,0)}$, $1 \leq j \leq d$.

(1b) Error expansion at step 3.

Observe that, using (6.9), we have for $1 \leq \ell \leq d$,

$$(\hat{Z}_{t_{i,3}})^\ell = \mathbb{E}_{t_{i,3}}\big[(H^{\psi_3}_{t_{i,3},c_3 h_i})^\ell u_{t_{i+1}} - h_i \alpha_{31}(H^{\phi_3}_{t_{i,3},c_3 h_i})^\ell u^{(0)}_{t_{i+1}}$$

$$- h_i \tilde{\alpha}_{32}(H^{\phi_3}_{t_{i,3},(c_3-c_2)h_i})^\ell u^{(0)}_{t_{i,2}}\big]$$

$$- \mathbb{E}_{t_{i,3}}\left[\tilde{\alpha}_{32}\frac{c_2^2}{2}h_i^3(H^{\phi_3}_{t_{i,3},(c_3-c_2)h_i})^\ell \sum_{j=1}^{d} {}^j v_{t_{i,2}}\right] + O_{t_{i,3}}(|\pi|^4).$$

We also used that $\mathbb{E}_{t_{i,3}}[(H^{\phi_3}_{t_{i,3},(c_3-c_2)h_i})^\ell \sum_{j=1}^{d} {}^j w_{t_{i,2}}] = O_{t_{i,3}}(1)$, recalling Proposition 2.3 and that under $(\mathbf{H}r)_4$, ${}^j w \in \mathcal{G}_b^1$, $1 \leq j \leq d$.

Applying Proposition 2.3, we compute, recalling that $\alpha_{31} + \tilde{\alpha}_{32} = c_3$,

$$(\hat{Z}_{t_{i,3}})^\ell = u^{(\ell)}_{t_{i,3}} - \left(\frac{c_3^2}{2} - \tilde{\alpha}_{32}c_2\right)h_i^2 u^{(\ell,0,0)}_{t_{i3}} - \left(\frac{c_3^3}{3} + \left(\frac{c_2^2}{2} - c_2 c_3\right)\tilde{\alpha}_{32}\right)h_i^3 u^{(\ell,0,0,0)}_{t_{i3}}$$

$$- \tilde{\alpha}_{32}\frac{c_2^2}{2}h_i^3 \mathbb{E}_{t_{i,3}}\left[(H^{\phi_3}_{t_{i,3},(c_3-c_2)h_i})^\ell \sum_{j=1}^{d} {}^j v_{t_{i,2}}\right] + O_{t_{i,3}}(|\pi|^4).$$

Under $(\mathbf{H}r)_4$, ${}^j v \in \mathcal{G}_b^2$, $1 \leq j \leq d$, applying Proposition 2.3(i), we have that

$$\mathbb{E}_{t_{i,3}}\big[(H^{\phi_3}_{t_{i,3},(c_3-c_2)h_i})^{\ell}\,{}^j v_{t_{i,2}}\big] = {}^j v^{(\ell)}_{t_{i,3}} + O_{t_{i,2}}(|\pi|).$$

We straightforwardly deduce that

$$f(\hat{Z}_{t_{i,3}}) = -u^{(0)}_{t_{i,3}} - \left(\frac{c_3^2}{2} - \tilde{\alpha}_{32}c_2\right)h_i^2 \sum_{j=1}^{d}{}^j v_{t_{i3}}$$

(6.10)
$$-\left(\frac{c_3^3}{3} + \left(\frac{c_2^2}{2} - c_2 c_3\right)\tilde{\alpha}_{32}\right)h_i^3 \sum_{j=1}^{d}{}^j w_{t_{i3}}$$

$$-\tilde{\alpha}_{32}\frac{c_2^2}{2}h_i^3 \sum_{\ell=1}^{d}\sum_{j=1}^{d} f^{z^\ell}_{t_{i,3}}\,{}^j v^{(\ell)}_{t_{i,3}} + O_{t_{i,3}}(|\pi|^4).$$

(1c) Error expansion at step 4.

Using (6.9)–(6.10), we obtain for $1 \leq \ell \leq d$,

$$(\hat{Z}_{t_{i,4}})^{\ell} = \mathbb{E}_{t_{i,4}}\big[(H^{\psi_4}_{t_{i,4},c_4 h_i})^{\ell} u_{t_{i+1}}\big] + O_{t_{i,4}}(|\pi|^4)$$

$$- h_i \mathbb{E}_{t_{i,4}}\big[\alpha_{41}(H^{\phi_4}_{t_{i,4},c_4 h_i})^{\ell} u^{(0)}_{t_{i+1}} + \tilde{\alpha}_{42}(H^{\phi_4}_{t_{i,4},(c_4-c_2)h_i})^{\ell} u^{(0)}_{t_{i,2}}$$

$$+ \tilde{\alpha}_{43}(H^{\phi_4}_{t_{i,4},(c_4-c_3)h_i})^{\ell} u^{(0)}_{t_{i,3}}\big]$$

$$- \mathbb{E}_{t_{i,4}}\Bigg[\frac{c_2^2}{2}\tilde{\alpha}_{42}h_i^3(H^{\phi_4}_{t_{i,4},(c_4-c_2)h_i})^{\ell} \sum_{j=1}^{d}{}^j v_{t_{i,2}}$$

$$+ \left(\frac{c_3^2}{2} - \tilde{\alpha}_{32}c_2\right)\tilde{\alpha}_{43}h_i^3(H^{\phi_4}_{t_{i,4},(c_4-c_3)h_i})^{\ell} \sum_{j=1}^{d}{}^j v_{t_{i3}}\Bigg].$$

Using Proposition 2.3, recalling that $\alpha_{41} + \tilde{\alpha}_{42} + \tilde{\alpha}_{43} = c_4$, we compute

$$(\hat{Z}_{t_{i,4}})^{\ell} = u^{(\ell)}_{t_{i,4}} - \left(\frac{c_4^2}{2} - \tilde{\alpha}_{42}c_2 - \tilde{\alpha}_{43}c_3\right)h_i^2 u^{(\ell,0,0)}_{t_{i4}}$$

$$- \left(\frac{c_4^3}{3} + \tilde{\alpha}_{42}\frac{c_2^2}{2} - \tilde{\alpha}_{42}c_2 c_4 + \tilde{\alpha}_{43}\frac{c_3^2}{2} - \tilde{\alpha}_{43}c_3 c_4\right)h_i^3 u^{(\ell,0,0,0)}_{t_{i4}}$$

$$+ \tilde{\alpha}_{42}\frac{c_2^2}{2}h_i^3 \mathbb{E}_{t_{i,4}}\Bigg[(H^{\phi_4}_{t_{i,4},(c_4-c_2)h_i})^{\ell} \sum_{j=1}^{d}{}^j v_{t_{i,2}}\Bigg]$$

$$- \tilde{\alpha}_{43}\left(\frac{c_3^2}{2} - \tilde{\alpha}_{32}c_2\right)h_i^3 \mathbb{E}_{t_{i,4}}\Bigg[(H^{\phi_4}_{t_{i,4},(c_4-c_3)h_i})^{\ell} \sum_{j=1}^{d}{}^j v_{t_{i3}}\Bigg]$$

$$+ O_{t_{i,4}}(|\pi|^4).$$

Applying Proposition 2.3, this leads to, recalling that $(\mathbf{H}r)_4$ is in force,

$$
f(\hat{Z}_{t_{i,4}}) = -u_{t_{i,4}}^{(0)} - \left(\frac{c_4^2}{2} - \tilde{\alpha}_{42}c_2 - \tilde{\alpha}_{43}c_3 \right) h_i^2 \sum_{j=1}^{d} {}^j v_{t_{i4}}
$$

$$
- \left(\frac{c_4^3}{3} + \tilde{\alpha}_{42}\frac{c_2^2}{2} - \tilde{\alpha}_{42}c_2c_4 + \tilde{\alpha}_{43}\frac{c_3^2}{2} - \tilde{\alpha}_{43}c_3c_4 \right) h_i^3 \sum_{j=1}^{d} {}^j w_{t_{i4}}
$$

(6.11)

$$
- \left(\tilde{\alpha}_{42}\frac{c_2^2}{2} + \tilde{\alpha}_{43}\left(\frac{c_3^2}{2} - \tilde{\alpha}_{32}c_2 \right) \right) h_i^3 \sum_{\ell=1}^{d}\sum_{j=1}^{d} f_{t_{i,4}}^{z^\ell\, j}\, v_{t_{i,4}}^{(\ell)}
$$

$$
+ O_{t_{i,4}}(|\pi|^4).
$$

(2a) We now study the error for the Y-part at the final step.
Using (6.9)–(6.11), we obtain

$$
\hat{Y}_{t_i} = \mathbb{E}_{t_i}\left[u_{t_{i+1}} - h_i\left(b_1 u_{t_{i+1}}^{(0)} + b_2 u_{t_{i,2}}^{(0)} + b_3 u_{t_{i,3}}^{(0)} + b_4 u_{t_{i,4}}^{(0)} \right) \right]
$$

$$
- b_2 \mathbb{E}_{t_i}\left[\frac{c_2^2}{2} h_i^3 \sum_{j=1}^{d} {}^j v_{t_{i,2}} + \frac{c_2^3}{3} h_i^4 \sum_{j=1}^{d} {}^j w_{t_{i,2}} \right]
$$

$$
- b_3 \mathbb{E}_{t_i}\left[\left(\frac{c_3^2}{2} - \tilde{\alpha}_{32}c_2 \right) h_i^3 \sum_{j=1}^{d} {}^j v_{t_{i3}} + \left(\frac{c_3^3}{3} + \tilde{\alpha}_{32}\frac{c_2^2}{2} - \tilde{\alpha}_{32}c_2c_3 \right) h_i^4 \sum_{j=1}^{d} {}^j w_{t_{i3}} \right.
$$

$$
\left. + \tilde{\alpha}_{32}\frac{c_2^2}{2} h_i^4 \sum_{\ell=1}^{d}\sum_{j=1}^{d} f_{t_{i,3}}^{z^\ell\, j}\, v_{t_{i,3}}^{(\ell)} \right]
$$

$$
- b_4 \mathbb{E}_{t_i}\left[\left(\frac{c_4^2}{2} - \tilde{\alpha}_{42}c_2 - \tilde{\alpha}_{43}c_3 \right) h_i^3 \sum_{j=1}^{d} {}^j v_{t_{i4}} \right.
$$

$$
\left. + \left(\frac{c_4^3}{3} + \tilde{\alpha}_{42}\frac{c_2^2}{2} - \tilde{\alpha}_{42}c_2c_4 + \tilde{\alpha}_{43}\frac{c_3^2}{2} - \tilde{\alpha}_{43}c_3c_4 \right) h_i^4 \sum_{j=1}^{d} {}^j w_{t_{i4}} \right]
$$

$$
- b_4 \mathbb{E}_{t_i}\left[\left(\tilde{\alpha}_{42}\frac{c_2^2}{2} + \tilde{\alpha}_{43}\left(\frac{c_3^2}{2} - \tilde{\alpha}_{32}c_2 \right) \right) h_i^4 \sum_{\ell=1}^{d}\sum_{j=1}^{d} f_{t_{i,4}}^{z^\ell\, j}\, v_{t_{i,4}}^{(\ell)} \right]
$$

$$
+ O_{t_i}(|\pi|^5).
$$

Under $(\mathbf{H}r)_4$, using Proposition 2.3, we then compute

$$
\hat{Y}_{t_i} = u_{t_i} + h_i(1 - b_1 - b_2 - b_3 - b_4)u_{t_i}^{(0)}
$$

$$
+ h_i^2\left(\frac{1}{2} - b_2(1-c_2) - b_3(1-c_3) - b_4(1-c_4) \right) u_{t_i}^{(0,0)}
$$

$$+ h_i^3 \left(\frac{1}{6} - \frac{b_2(1-c_2)^2 + b_3(1-c_3)^2 + b_4(1-c_4)^2}{2} \right) u_{t_i}^{(0,0,0)}$$

$$+ h_i^4 \left(\frac{1}{24} - \frac{b_2(1-c_2)^3 + b_3(1-c_3)^3 + b_4(1-c_4)^3}{6} \right) u_{t_i}^{(0,0,0,0)}$$

$$- h_i^3 \left(b_2 \frac{c_2^2}{2} + b_3 \left(\frac{c_3^2}{2} - \tilde{\alpha}_{32}c_2 \right) + b_4 \left(\frac{c_4^2}{2} - \tilde{\alpha}_{42}c_2 - \tilde{\alpha}_{43}c_3 \right) \right) \sum_{j=1}^{d} {}^j v_{t_i}$$

$$- h_i^4 \left(b_2 \frac{c_2^2}{2}(1-c_2) + b_3 \left(\frac{c_3^2}{2} - \tilde{\alpha}_{32}c_2 \right)(1-c_3) \right.$$

$$\left. + b_4 \left(\frac{c_4^2}{2} - \tilde{\alpha}_{42}c_2 - \tilde{\alpha}_{43}c_3 \right)(1-c_4) \right) \sum_{j=1}^{d} {}^j v_{t_i}^{(0)}$$

$$- h_i^4 \left(b_2 \frac{c_2^3}{3} + b_3 \left(\frac{c_3^3}{3} + \tilde{\alpha}_{32}\frac{c_2^2}{2} - \tilde{\alpha}_{32}c_2c_3 \right) \right.$$

$$\left. + b_4 \left(\frac{c_4^3}{3} + \tilde{\alpha}_{42}\frac{c_2^2}{2} - \tilde{\alpha}_{42}c_2c_4 + \tilde{\alpha}_{43}\frac{c_3^2}{2} - \tilde{\alpha}_{43}c_3c_4 \right) \right) \sum_{j=1}^{d} {}^j w_{t_i}$$

$$- h_i^4 \left(b_3\tilde{\alpha}_{32}\frac{c_2^2}{2} + b_4 \left(\tilde{\alpha}_{42}\frac{c_2^2}{2} + \tilde{\alpha}_{43} \left(\frac{c_3^2}{2} - \tilde{\alpha}_{32}c_2 \right) \right) \right) \sum_{\ell=1}^{d} \sum_{j=1}^{d} f_{t_i}^{z^\ell j} v_{t_i}^{(\ell)}$$

$$+ O_{t_i}(|\pi|^5).$$

Under $(\mathbf{H}o)_4$, using the same techniques as in the proof of Theorem 1.4, one proves inductively on the order that each factor has to be equal to 0, which leads to the set (i) of conditions of the lemma. It appears that these conditions are the same as in the ODE case. From, for example, Section 322, page 175 in [3], we know that $c_4 = 1$ necessarily.

(3) We now study the error for the Z-part at the final step, taking into account $c_4 = 1$ and $c_2 < c_3 < 1$. We thus have

$$\hat{Z}_{t_i} = \mathbb{E}_{t_i} \big[H_{t_i,h_i}^{\psi_5} Y_{t_{i+1}} + h_i \big(\beta_1 H_{t_i,h_i}^{\phi_5} f(Z_{t_{i+1}}) + \beta_2 H_{t_i,(1-c_2)h_i}^{\phi_5} f(\hat{Z}_{t_i,2})$$

$$+ \tilde{\beta}_3 H_{t_i,(1-c_3)h_i}^{\phi_5} f(Z_{t_i,3}) \big) \big].$$

We are thus considering a 3-stage scheme for the Z part. Using the results of step 1, we obtain the following expansion, for $1 \le \ell \le d$:

$$\hat{Z}_{t_i}^\ell = Z_{t_i}^\ell + (1 - \beta_1 - \beta_2 - \beta_3) u_{t_i}^{(\ell,0)}$$

$$+ \left(\frac{1}{2} - \beta_1 - \beta_2(1-c_2) - \beta_3(1-c_3) \right) h_i u_{t_i}^{(\ell,0,0)}$$

$$+ \left(\frac{1}{6} - \frac{1}{2}\beta_1 - \frac{1}{2}\beta_2(1 - c_2)^2 - \frac{1}{2}\beta_3(1 - c_3)^2 \right) h_i^2 u_{t_i}^{(\ell,0,0,0)}$$

$$- \left(\beta_2 \frac{c_2^2}{2} + \beta_3 \frac{c_3^2}{2} - \beta_3 c_2 \tilde{\alpha}_{32} \right) h_i^2 f_{t_i}^{z^\ell} \sum_{j=1}^{d} u_{t_i}^{(j,0,0)} + O_{t_i}(|\pi|^3).$$

It is then obvious that set (ii) of the condition is sufficient to obtain an order 4 truncation error on Z. Moreover, arguing as, for example, in steps (2b)–(2c) of the proof of Theorem 1.4, by induction on the order required, one proves that these condition are also necessary, provided that $(\mathbf{H}o)_4$ is in force. □

PROOF OF THEOREM 1.7. The set of condition (ii) leads, using case I of Theorem 1.3, with $(b_j) = (\beta_j)$ and $(a_{kj}) = (\alpha_{kj})$, to the only possible value for α_{32} is given by

$$\alpha_{32} = \frac{c_3(c_3 - c_2)}{c_2(2 - 3c_2)}.$$

In our context equations (322b) and (322c) in [3] read

$$b_4 \tilde{\alpha}_{43}(c_3 - c_2)c_3 = \frac{1}{12} - \frac{c_2}{6},$$

$$b_4 \tilde{\alpha}_{43} \tilde{\alpha}_{32} c_2 = \frac{1}{24}.$$

Dividing these two equations, we obtain

$$\frac{(c_3 - c_2)c_3}{\alpha_{32}c_2} = 2 - 4c_2.$$

It follows from the expression of α_{32} that $c_2 = 0$, which is not possible. □

APPENDIX

A.1. Schemes stability.

A.1.1. *Proof of Theorem* 1.1. Using (1.19)–(1.20), we compute, for $1 \geq \eta > 0$ to be fixed later on that

$$|\delta Y_i|^2 \leq \left(1 + \frac{h_i}{\eta} \right) |\mathbb{E}_{t_i}[\delta Y_{i+1}]|^2 + C\frac{\eta}{h_i} |\mathbb{E}_{t_i}[\zeta_i^Y]|^2$$

(A.1)
$$+ Ch_i^2 \left(1 + \frac{\eta}{h_i} \right) \left(\frac{1}{h} B_i + \mathbb{E}_{t_i}[|\delta Y_{i+1}|^2 + |\delta Z_{i+1}|^2] \right),$$

$$|\delta Z_i|^2 \leq C\left(\frac{1}{h_i} B_i + h_i \mathbb{E}_{t_i}[|\delta Y_{i+1}|^2 + |\delta Z_{i+1}|^2] + |\mathbb{E}_{t_i}[\zeta_i^Z]|^2 \right),$$

where $B_i := \mathbb{E}_{t_i}[|\delta Y_{i+1}|^2 - |\mathbb{E}_{t_i}[\delta Y_{i+1}]|^2]$.

Defining for $1 \geq \varepsilon > 0$ to be fixed later on $I_i^\varepsilon := |\delta Y_i|^2 + \varepsilon h_i |\delta Z_i|^2$, we compute

$$I_i^{\varepsilon/2} + \frac{\varepsilon}{2} h_i |\delta Z_i|^2$$

$$\leq \left(1 + \frac{h_i}{\eta} \right) |\mathbb{E}_{t_i}[\delta Y_{i+1}]|^2 + C(\varepsilon + \eta) B_i + C \frac{\eta}{h_i} |\mathbb{E}_{t_i}[\zeta_i^Y]|^2 + C h_i |\mathbb{E}_{t_i}[\zeta_i^Z]|^2$$

$$+ \left(C h_i^2 \left(1 + \frac{\eta}{h_i} \right) + C \varepsilon h_i^2 \right) \mathbb{E}_{t_i} \left[|\delta Y_{i,+1}|^2 + |\delta Z_{i,+1}|^2 \right].$$

Setting $\varepsilon = \eta = \frac{1}{2C}$ and observing that $|\mathbb{E}_{t_i}[\delta Y_{i+1}]|^2 = \mathbb{E}_{t_i}[|\delta Y_{i+1}|^2] - B_i$, we compute that, for h^* small enough

$$(A.2) \qquad I_i^{\varepsilon/2} + \frac{\varepsilon}{2} h_i |\delta Z_i|^2 \leq (1 + C h_i) I_{i+1}^{\varepsilon/2} + C \frac{\eta}{h_i} |\mathbb{E}_{t_i}[\zeta_i^Y]|^2 + C h_i |\mathbb{E}_{t_i}[\zeta_i^Z]|^2.$$

Using the discrete version of Gronwall's lemma, we obtain

$$\max_{0 \leq i \leq n-1} \mathbb{E}[|\delta Y_i|^2] \leq \max_{0 \leq i \leq n-1} I_i^{\varepsilon/2}$$

$$\leq C \left(I_n^{\varepsilon/2} + \sum_{i=0}^{n-1} h_i \mathbb{E} \left[\frac{1}{h_i^2} |\mathbb{E}_{t_i}[\zeta_i^Y]|^2 + |\mathbb{E}_{t_i}[\zeta_i^Z]|^2 \right] \right).$$

The control of $\sum_{i=0}^{n-1} h_i \mathbb{E}[|\delta Z_i|^2]$ is then obtained summing inequality (A.2) over i.

A.1.2. *Proof of Proposition* 1.1.   We simply observe that the solution (Y, Z) of the BSDE is also the solution of a perturbed scheme with $\zeta_i^Y := \hat{Y}_{t_i} - Y_{t_i}$ and $\zeta_i^Z := \hat{Z}_{t_i} - Z_{t_i}$, and with terminal conditions $\tilde{Y}_n := g(X_T)$ and $\tilde{Z}_n := \nabla g^\top(X_T)\sigma(X_T)$. The proof then follows directly from Theorem 1.1.

A.1.3. *Proof of Theorem* 1.2.   Claim (ii) is a direct application of (i) and Proposition 1.2.

We now prove (i).

(1) We define $U_{i,j}$ (resp., $\tilde{U}_{i,j}$) and $V_{i,j}$ (resp., $\tilde{V}_{i,j}$) as $Y_{i,j}$ and $Z_{i,j}$ in Definition 1.1(ii) using U (resp., $\tilde{U}$) instead of Y_{i+1} and V (resp., $\tilde{V}$) instead of Z_{i+1}. Let us also denote

$$F_{i,j} := f(U_{i,j}, V_{i,j}), \qquad \tilde{F}_{i,j} := f(\tilde{U}_{i,j}, \tilde{V}_{i,j}) \quad \text{and} \quad \delta F_{i,j} := F_{i,j} - \tilde{F}_{i,j}.$$

With this notation, we have that

$$\Phi_i^Y(U, V) := \sum_{j=1}^{q+1} b_j f(U_{i,j}, V_{i,j}) \quad \text{and} \quad \Phi_i^Y(\tilde{U}, \tilde{V}) = \sum_{j=1}^{q+1} b_j f(\tilde{U}_{i,j}, \tilde{V}_{i,j}).$$

Since f is Lipschitz-continuous, we compute

$$\mathbb{E}_{t_i}\big[\big|\Phi_i^Y(U,V) - \Phi_i^Y(\tilde{U},\tilde{V})\big|^2\big] \le C\mathbb{E}_{t_i}\big[|\delta U_{i,1}|^2 + |\delta V_{i,1}|^2\big] + \sum_{j=2}^{q+1}\mathbb{E}_{t_i}\big[|\delta F_{i,j}|^2\big].$$

We also have that

$$\mathbb{E}_{t_i}\big[\Phi_i^Z(U,V)\big] = \sum_{j=1}^{q} \beta_j H_{q+1,j}^i f(U_{i,j}, V_{i,j})$$

and

$$\mathbb{E}_{t_i}\big[\Phi_i^Z(\tilde{U},\tilde{V})\big] = \sum_{j=1}^{q} \beta_j H_{q+1,j}^i f(\tilde{U}_{i,j}, \tilde{V}_{i,j}).$$

Combining the Cauchy–Schwarz inequality with property (1.12) and the Lipschitz continuity of f, we compute

$$\mathbb{E}_{t_i}\big[\big|\Phi_i^Z(U,V) - \Phi_i^Z(\tilde{U},\tilde{V})\big|^2\big]$$
$$\le \frac{C}{h_i}\mathbb{E}_{t_i}\big[|\delta U_{i,1}|^2 + |\delta V_{i,1}|^2\big] + \frac{C}{h_i}\sum_{j=2}^{q}\mathbb{E}_{t_i}\big[|\delta F_{i,j}|^2\big].$$

Moreover, we observe, using the Lipschitz-continuity property of f,

$$\mathbb{E}_{t_i}\big[|\delta F_{i,j}|^2\big] \le C\mathbb{E}_{t_i}\big[|\delta U_{i,j}|^2 + |\delta V_{i,j}|^2\big].$$

(2a) For $j = 2$, we compute that

$$\mathbb{E}_{t_i}\big[|\delta U_{i,2}|^2\big] \le C\big(\mathbb{E}_{t_i}\big[|\delta U_{i,1}|^2 + h_i^2|\delta V_{i,1}|^2\big] + h_i^2\mathbb{E}_{t_i}\big[|\delta U_{i,2}|^2 + |\delta V_{i,2}|^2\big]\big),$$

$$\mathbb{E}_{t_i}\big[|\delta V_{i,2}|^2\big] \le C\Big(\frac{1}{h_i}\mathbb{E}_{t_i}\big[|\delta U_{i,1}|^2 - |\mathbb{E}_{t_i}[\delta U_{i,1}]|^2\big] + h_i\mathbb{E}_{t_i}\big[|\delta U_{i,1}|^2 + |\delta V_{i,1}|^2\big]\Big).$$

For $|\pi|$ small enough, we then obtain

$$\mathbb{E}_{t_i}\big[|\delta U_{i,2}|^2 + |\delta V_{i,2}|^2\big]$$
$$\le C\Big(\frac{1}{h_i}\mathbb{E}_{t_i}\big[|\delta U_{i,1}|^2 - |\mathbb{E}_{t_i}[\delta U_{i,1}]|^2\big] + \mathbb{E}_{t_i}\big[|\delta U_{i,1}|^2 + h_i|\delta V_{i,1}|^2\big]\Big),$$

which, since f is Lipschitz, straightforwardly leads to

$$\mathbb{E}_{t_i}\big[|\delta F_{i,2}|^2\big] \le C\Big(\frac{1}{h_i}\mathbb{E}_{t_i}\big[|\delta U_{i,1}|^2 - |\mathbb{E}_{t_i}[\delta U_{i,1}]|^2\big] + \mathbb{E}_{t_i}\big[|\delta U_{i,1}|^2 + h_i|\delta V_{i,1}|^2\big]\Big).$$

(2b) For $2 < j \le q+1$, we have that

$$\mathbb{E}_{t_i}\big[|\delta U_{i,j}|^2\big] \le C\mathbb{E}_{t_i}\Big[|\delta U_{i,1}|^2 + h_i^2\sum_{k=1}^{j-1}|\delta F_{i,j}|^2\Big] + Ch_i^2\mathbb{E}_{t_i}\big[|\delta U_{i,j}|^2 + |\delta V_{i,j}|^2\big],$$

$$\mathbb{E}_{t_i}\big[|\delta V_{i,j}|^2\big] \le C\Big(\frac{1}{h_i}\mathbb{E}_{t_i}\big[|\delta U_{i,1}|^2 - |\mathbb{E}_{t_i}[\delta U_{i,1}]|^2\big] + h_i\sum_{k=1}^{j-1}\mathbb{E}_{t_i}\big[|\delta F_{i,j}|^2\big]\Big)$$

and for $|\pi|$ small enough,

$$\mathbb{E}_{t_i}\big[|\delta F_{i,j}|^2\big]$$
$$\leq C\left(\frac{1}{h_i}\mathbb{E}_{t_i}\big[|\delta U_{i,1}|^2 - |\mathbb{E}_{t_i}[\delta U_{i,1}]|^2\big] + \mathbb{E}_{t_i}\big[|\delta U_{i,1}|^2\big] + h_i\sum_{k=1}^{j-1}\mathbb{E}_{t_i}\big[|\delta F_{i,j}|^2\big]\right).$$

An easy mathematical induction completes the proof.

A.2. Itô–Taylor expansions.

A.2.1. *Proof of Proposition* 2.2.   Using Proposition 2.1 (Theorem 5.5.1 in [9]), we compute

$$v(t+h, X_{t+h}) = \sum_{\alpha\in\mathcal{A}_m} v_t^\alpha I_{t,t+h}^\alpha + \sum_{\beta\in\mathcal{A}_{m+1}\setminus\mathcal{A}_m} I_{t,t+h}^\beta[v^\beta],$$

recalling that $\mathcal{B}(\mathcal{A}_m) = \mathcal{A}_{m+1}\setminus\mathcal{A}_m$.

Taking the conditional expectation on both sides and using Lemma 5.7.1 in [9], we obtain

$$\left|\mathbb{E}_t\big[v(t+h, X_{t+h})\big] - \sum_{k=0}^m v_t^{(0)_k}\frac{h^k}{k!}\right| = \big|\mathbb{E}_t\big[I_{t,t+h}^{(0)_{m+1}}\big[v_\cdot^{(0)_{m+1}}\big]\big]\big|.$$

Since $v\in\mathcal{G}_b^\beta$ for all $\beta\in\mathcal{A}_{m+1}$, in particular $v\in\mathcal{G}_b^{(0)_{m+1}}$, we obtain

$$\big|\mathbb{E}_t\big[I_{t,t+h}^{(0)_{m+1}}\big[v^{(0)_{m+1}}\big]\big]\big| = O_t(h^{m+1}),$$

which completes the proof.

A.2.2. *Proof of Proposition* 2.3.   (i) (1) Using Proposition 2.1 (Theorem 5.5.1 in [9]), we compute

$$\big(H_{t,t+h}^\psi\big)^\ell v(t+h, X_{t+h}) - \sum_{\alpha\in\mathcal{A}_{m+1}} v_t^\alpha\big(H_{t,t+h}^\psi\big)^\ell I_{t,t+h}^\alpha$$
$$= \sum_{\beta\in\mathcal{A}_{m+2}\setminus\mathcal{A}_{m+1}}\big(H_{t,t+h}^\psi\big)^\ell I_{t,t+h}^\beta[v_\cdot^\beta],$$

recalling that $\mathcal{B}(\mathcal{A}_{m+1}) = \mathcal{A}_{m+2}\setminus\mathcal{A}_{m+1}$.

(2) We now compute $\mathbb{E}_t[(H_{t,t+h}^\psi)^\ell I_{t,t+h}^\alpha]$ for $\alpha\in\mathcal{A}_{m+1}$, recalling that $(H_{t,t+h}^\psi)^\ell := \frac{1}{h}I_{t,t+h}^{(\ell)}[\psi_{t,h}^\ell]$; see Definition 1.5(ii).

If $\alpha^+ \neq (\ell)$, we observe that $\mathbb{E}_t[(H_{t,t+h}^\psi)^\ell I_{t,t+h}^\alpha] = 0$; see, for example, Lemma 5.7.2 in [9].

Now, let α be such that $\ell(\alpha) = q$, $1 \leq q \leq m+1$ and $\alpha^+ = (\ell)$. Then there exits $1 \leq l \leq q$, such that $\alpha = (0)_{l-1} * (\ell) * (0)_{q-l}$, and we have

$$\mathbb{E}_t\left[(H_{t,h}^{\psi})^{\ell} I_{t,t+h}^{\alpha}\right]$$

$$= \frac{1}{h} I_{t,t+h}^{(0)_{q-l}} \left[\mathbb{E}_t\left[I_{t,\cdot}^{(\ell)}\left[\psi^{\ell}\left(\frac{\cdot - t}{h}\right)\right] I_{t,\cdot}^{(\ell)}\left[I_{t,\cdot}^{(0)_{l-1}}\right]\right]\right]$$

(A.3)
$$= \frac{1}{h(l-1)!} I_{t,t+h}^{(0)_{q-l}} \left[I_{t,\cdot}^{(0)}\left[\psi^{\ell}\left(\frac{\cdot - t}{h}\right)(\cdot - t)^{l-1}\right]\right]$$

$$= \frac{1}{h(l-1)!(q-l)!} \int_t^{t+h} (t+h-u)^{q-l}(u-t)^{l-1}\psi^{\ell}\left(\frac{u-t}{h}\right) du$$

$$= \frac{h^{q-1}}{(l-1)!(q-l)!} \int_0^1 (1-r)^{q-l} r^{l-1} \psi^{\ell}(r) \, dr.$$

Since $\psi^{\ell} \in \mathcal{B}_{[0,1]}^m$,

$$\mathbb{E}_t\left[(H_{t,h}^{\psi})^{\ell} I_{t,t+h}^{\alpha}\right] = \frac{h^{q-1}}{(q-1)!} \mathbf{1}_{\{\alpha_1 = \ell\}}.$$

(3) Using Lemma 5.7.2 in [9], for $\beta \in \mathcal{A}_{m+2} \setminus \mathcal{A}_{m+1}$ and $1 \leq j \leq d$, we have

$$\mathbb{E}_t\left[(H_{t,t+h}^{\psi})^j I_{t,t+h}^{\beta}[v^{\beta}]\right] := \frac{1}{h}\mathbb{E}_t\left[I_{t,t+h}^{(j)}[\psi_{t,h}^j] I_{t,t+h}^{\beta}[v^{\beta}]\right] = 0 \qquad \text{if } \beta^+ \neq (j).$$

We are now considering $\beta \in \mathcal{A}_{m+2} \setminus \mathcal{A}_{m+1}$ such that $\beta^+ = (j)$, that is, β with at most one nonzero component. According to the notation of Lemma 5.7.2 in [9] (see the beginning of Section 5.7 in [9]), we then compute that

$$k_0(\beta) + k_1(\beta) = m+1 \quad \text{and} \quad k_0((j)) = k_1((j)) = 0.$$

Since $\ell((j)^+) = 1$, we obtain $\omega((j), \beta) = m+2$ and using again Lemma 5.7.2, we obtain

$$\left|\mathbb{E}_t\left[\sum_{\beta \in \mathcal{A}_{m+2}} (H_{t,t+h}^{\psi})^j I_{t,t+h}^{\beta}[v^{\beta}]\right]\right| = O_t(h^{m+1}),$$

recalling that $v \in \mathcal{G}_b^{\beta}$, for $\beta \in \mathcal{A}_{m+2}$.

(ii) This is a straightforward consequence of Itô's formula applied to v and the fact that $v^{(0)}$ and $v^{(\ell)}$ are bounded under $\mathcal{G}_b^1$.

(iii) We follow the arguments of (i). In particular, since $\psi = (1, \dots, 1)$ in (A.3), using the basic properties of the Beta function, one obtains

$$\mathbb{E}\left[(H_{t,h}^{\psi})^j I_{t,t+h}^{\alpha}\right] = \frac{h^{q-1}}{q!}$$

for $\ell(\alpha) = q$ and $\alpha^+ = (j)$, $1 \leq q \leq m+1$, $1 \leq j \leq d$. The proof is completed observing that $v^{(\alpha)} = v^{(j)*(0)_{q-1}}$ for such α under the assumption $L^{(0)} \circ L^{(j)} = L^{(j)} \circ L^{(0)}$.

REFERENCES

[1] BOUCHARD, B. and ELIE, R. (2008). Discrete-time approximation of decoupled forward-backward SDE with jumps. *Stochastic Process*. *Appl*. **118** 53–75. MR2376252

[2] BOUCHARD, B. and TOUZI, N. (2004). Discrete-time approximation and Monte-Carlo simulation of backward stochastic differential equations. *Stochastic Process*. *Appl*. **111** 175–206. MR2056536

[3] BUTCHER, J. C. (2008). *Numerical Methods for Ordinary Differential Equations*, 2nd ed. Wiley, Chichester. MR2401398

[4] CRISAN, D. and MANOLARAKIS, K. (2012). Solving backward stochastic differential equations using the cubature method: Application to nonlinear pricing. *SIAM J. Financial Math*. **3** 534–571. MR2968045

[5] GEISS, C., GEISS, S. and GOBET, E. (2012). Generalized fractional smoothness and L_p-variation of BSDEs with non-Lipschitz terminal condition. *Stochastic Process*. *Appl*. **122** 2078–2116. MR2921973

[6] GOBET, E. and LABART, C. (2007). Error expansion for the discretization of backward stochastic differential equations. *Stochastic Process*. *Appl*. **117** 803–829. MR2330720

[7] GOBET, E., LEMOR, J.-P. and WARIN, X. (2005). A regression-based Monte Carlo method to solve backward stochastic differential equations. *Ann*. *Appl*. *Probab*. **15** 2172–2202. MR2152657

[8] GOBET, E. and MAKHLOUF, A. (2010). L_2-time regularity of BSDEs with irregular terminal functions. *Stochastic Process*. *Appl*. **120** 1105–1132. MR2639740

[9] KLOEDEN, P. E. and PLATEN, E. (1992). *Numerical Solution of Stochastic Differential Equations*. *Applications of Mathematics* (*New York*) **23**. Springer, Berlin. MR1214374

[10] KUSUOKA, S. (2001). Approximation of expectation of diffusion process and mathematical finance. In *Taniguchi Conference on Mathematics Nara '98*. *Adv*. *Stud*. *Pure Math*. **31** 147–165. Math. Soc. Japan, Tokyo. MR1865091

[11] LYONS, T. and VICTOIR, N. (2004). Cubature on Wiener space. *Proc*. *R*. *Soc*. *Lond*. *Ser. A Math*. *Phys*. *Eng*. *Sci*. **460** 169–198. MR2052260

[12] NINOMIYA, S. (2003). A new simulation scheme of diffusion processes: Application of the Kusuoka approximation to finance problems. *Math*. *Comput*. *Simulation* **62** 479–486. 3rd IMACS Seminar on Monte Carlo Methods—MCM 2001 (Salzburg). MR1988392

[13] NINOMIYA, S. (2003). A partial sampling method applied to the Kusuoka approximation. *Monte Carlo Methods Appl*. **9** 27–38. MR1987415

[14] NINOMIYA, S. and VICTOIR, N. (2008). Weak approximation of stochastic differential equations and application to derivative pricing. *Appl*. *Math*. *Finance* **15** 107–121. MR2409419

[15] PARDOUX, É. and PENG, S. (1992). Backward stochastic differential equations and quasilinear parabolic partial differential equations. In *Stochastic Partial Differential Equations and Their Applications* (*Charlotte*, *NC*, 1991). *Lecture Notes in Control and Inform*. *Sci*. **176** 200–217. Springer, Berlin. MR1176785

[16] PARDOUX, É. and PENG, S. G. (1990). Adapted solution of a backward stochastic differential equation. *Systems Control Lett*. **14** 55–61. MR1037747

[17] ZHANG, J. (2004). A numerical scheme for BSDEs. *Ann*. *Appl*. *Probab*. **14** 459–488. MR2023027

DEPARTMENT OF MATHEMATICS
IMPERIAL COLLEGE LONDON
SOUTH KENSINGTON CAMPUS
LONDON SW7 2AZ
UNITED KINGDOM
E-MAIL: j.chassagneux@imperial.ac.uk
        d.crisan@imperial.ac.uk