The One Step Malliavin scheme: new discretization of BSDEs implemented with deep learning regressions

Négyesi, Bálint; Andersson, Kristoffer ; Oosterlee, Cornelis W.

DOI
10.1093/imanum/drad092

Publication date
2024

Document Version
Final published version

Published in
IMA Journal of Numerical Analysis

IMA Journal of Numerical Analysis (2024) **00**, 1–53
https://doi.org/10.1093/imanum/drad092

The One Step Malliavin scheme: new discretization of BSDEs implemented with deep learning regressions

Balint Negyesi*
Delft Institute of Applied Mathematics (DIAM), Delft University of Technology, PO Box 5031, 2600 GA Delft, The Netherlands
*Corresponding author: B.Negyesi@tudelft.nl

Kristoffer Andersson
Research Group of Scientific Computing, Centrum Wiskunde & Informatica, PO Box 94079, 1090 GB Amsterdam, The Netherlands
Kristoffer.Andersson@cwi.nl

and

Cornelis W. Oosterlee
Mathematical Institute, Utrecht University, Postbus 80010, 3508 TA Utrecht, The Netherlands
C.W.Oosterlee@uu.nl

[Received on 22 June 2022; revised on 2 June 2023]

A novel discretization is presented for decoupled forward–backward stochastic differential equations (FBSDE) with differentiable coefficients, simultaneously solving the BSDE and its Malliavin sensitivity problem. The control process is estimated by the corresponding linear BSDE driving the trajectories of the Malliavin derivatives of the solution pair, which implies the need to provide accurate Γ estimates. The approximation is based on a merged formulation given by the Feynman–Kac formulae and the Malliavin chain rule. The continuous time dynamics is discretized with a theta-scheme. In order to allow for an efficient numerical solution of the arising semidiscrete conditional expectations in possibly high dimensions, it is fundamental that the chosen approach admits to differentiable estimates. Two fully-implementable schemes are considered: the BCOS method as a reference in the one-dimensional framework and neural network Monte Carlo regressions in case of high-dimensional problems, similarly to the recently emerging class of Deep BSDE methods (Han *et al.* (2018 Solving high-dimensional partial differential equations using deep learning. *Proc. Natl. Acad. Sci.*, **115**, 8505–8510); Huré *et al.* (2020 Deep backward schemes for high-dimensional nonlinear PDEs. *Math. Comp.*, **89**, 1547–1579)). An error analysis is carried out to show $\mathbb{L}^2$ convergence of order $1/2$, under standard Lipschitz assumptions and additive noise in the forward diffusion. Numerical experiments are provided for a range of different semilinear equations up to 50 dimensions, demonstrating that the proposed scheme yields a significant improvement in the control estimations.

Keywords: backward stochastic differential equations; Malliavin calculus; deep BSDE; neural networks; BCOS; gamma estimates.

1. Introduction

In this paper, we are concerned with the numerical solution of a system of forward–backward stochastic differential equations (FBSDE) where the randomness in the backward equation (BSDE) is driven by a

forward stochastic differential equation (SDE). These systems are written in the general form

$$X_t = x_0 + \int_0^t \mu(s, X_s)\,\mathrm{d}s + \int_0^t \sigma(s, X_s)\,\mathrm{d}W_s, \tag{1.1a}$$

$$Y_t = g(X_T) + \int_t^T f(s, X_s, Y_s, Z_s)\,\mathrm{d}s - \int_t^T \left(Z_s\,\mathrm{d}W_s\right)^T, \tag{1.1b}$$

where $\{W_t\}_{0\le t\le T}$ is a d-dimensional Brownian motion and $\mu : [0,T]\times\mathbb{R}^{d\times 1} \to \mathbb{R}^{d\times 1}$, $\sigma : [0,T]\times\mathbb{R}^{d\times 1} \to \mathbb{R}^{d\times d}$, $g : \mathbb{R}^{d\times 1} \to \mathbb{R}^{q\times 1}$ and $f : [0,T]\times\mathbb{R}^{d\times 1}\times\mathbb{R}^{q\times 1}\times\mathbb{R}^{q\times d} \to \mathbb{R}^{q\times 1}$ are all deterministic mappings of time and space, with some fixed $T > 0$. Adhering to the stochastic control terminology, we often refer to Z as the *control process*. We shall work under the standard well-posedness assumptions of Pardoux & Peng (1992), which require Lipschitz continuity of the corresponding coefficients in order to ensure the existence of a unique solution pair $\{(Y_t, Z_t)\}_{0\le t\le T}$ adapted to the augmented natural filtration. The main motivation to study FBSDE systems lies in their connection with parabolic, second-order partial differential equations (PDE), generalizing the well-known Feynman–Kac relations to nonlinear settings. Indeed, considering the semilinear, parabolic terminal problem

$$\begin{aligned} \partial_t u(t,x) + \frac{1}{2}\mathrm{Tr}\{\sigma\sigma^T(t,x)\mathrm{Hess}_x u(t,x)\} + \nabla_x u(t,x)\mu(t,x) + f(t,x,u,\nabla_x u(t,x)\sigma(t,x)) = 0 \\ u(T,x) = g(x), \end{aligned} \tag{1.2}$$

the Markov solution to (1.1) coincides with the solution of (1.2) in an almost sure sense, provided by the *nonlinear Feynman–Kac* relations

$$Y_t = u(t, X_t), \qquad Z_t = \nabla_x u(t, X_t)\sigma(t, X_t). \tag{1.3}$$

Consequently, the BSDE formulation provides a stochastic representation to the simultaneous solution of a parabolic problem and its gradient, which is an advantageous feature for several applications in stochastic control and finance, where sensitivities play a fundamental role. These relations can be extended to *viscosity solutions* in case (1.2) does not admit to a classical solution—see Pardoux & Peng (1992). Moreover, it is known—see Pardoux & Peng (1992); El Karoui *et al.* (1997); Hu *et al.* (2011); Mastrolia *et al.* (2017)—that under suitable regularity assumptions the solution pair of the backward equation is differentiable in the Malliavin sense Nualart (2006), and the Malliavin derivatives $\{(D_s Y_t, D_s Z_t)\}_{0\le s,t\le T}$ satisfy a linear BSDE themselves, where the Z process admits to a continuous modification provided by $Z_t = D_t Y_t$.

From a numerical standpoint, the main challenge in solving BSDEs stems from the approximation of conditional expectations. Indeed, a discretization of the backward equation in (1.1b) yields a sequence of recursively nested conditional expectations at each point in the discretized time window. Over the years, several methods have been proposed to tackle the solution of the FBSDE system using: PDE methods in Ma *et al.* (1994); forward Picard iterations in Bender & Denk (2007); quantization techniques in Bally & Pagès (2003); chaos expansion formulas in Briand & Labart (2014); Fourier cosine expansions in Ruijter & Oosterlee (2015, 2016) and regression Monte Carlo approaches in Bouchard & Touzi (2004); Gobet *et al.* (2005); Bender & Steiner (2012). These methods have shown great results in low-dimensional

settings; however, the majority of them suffers from the curse of dimensionality, meaning that their computational complexity scales exponentially in the number of dimensions. Although, regression Monte Carlo methods have been successfully proven to overcome this burden, they are difficult to apply beyond $d = 10$ dimensions due to the necessity of a finite regression basis. The primary challenge in the numerical solution of BSDEs is related to the approximation of the Z process. In particular, the standard backward Euler discretization results in a conditional expectation estimate of Z, which scales inverse proportionally with the step size of the time discretization—see Bouchard & Touzi (2004). This phenomenon poses a significant amount of difficulty in least-squares Monte Carlo frameworks, as the corresponding regression targets have diverging conditional variances in the continuous limit.

Recently, the field has received renewed attention due to the pioneering paper of Han *et al.* (2018), in which they reformulate the backward discretization in a forward fashion, parametrize the control process of the solution by deep neural networks and train the resulting sequence of networks in a global optimization given by the terminal condition of (1.1b). Their method has enjoyed various modifications and extensions, see, e.g., Beck *et al.* (2019); Fujii *et al.* (2019). In particular, Huré *et al.* (2020) proposed an alternative where the optimization of the sequence of neural networks is done in a backward recursive manner, similarly to classical regression Monte Carlo approaches. We refer to the class of these deep learning based formulations as *Deep BSDE* methods, which have shown remarkable empirical results in solving high-dimensional problems. Note, however, that the approach of Han *et al.* (2018) solely captures the deterministic mapping connecting the forward diffusion in (1.1) to the solution pair of the BSDE at $t = 0$. Even though the extension of Huré *et al.* (2020) gives such approximations at future time steps, the accuracy of both methods degrades significantly in the Z part of the solution. The total approximation errors of such Deep BSDE methods have been investigated in Han & Long (2020); Huré *et al.* (2020); Germain *et al.* (2021). The results in Han & Long (2020) provide a *posteriori estimate* driven by the error in the terminal condition, whereas the analyses in Huré *et al.* (2020); Germain *et al.* (2021) show that due to the universal approximation theorem (UAT) of deep neural networks, the total approximation error of neural network parametrizations is consistent with the discretization in terms of regression biases.

The main motivation behind the present paper roots in the observations above. In order to provide more accurate solutions for the Z process, we exploit the aforementioned relation between the Malliavin derivative of Y and the control process by solving the linear BSDE driving the trajectories of DY. Hence, we are faced with the solution of one scalar-valued BSDE and one d-dimensional BSDE at each point in time. This raises the need for a new discrete scheme, which we call the *One Step Malliavin (OSM)* scheme. The discretization of the linear BSDE of the Malliavin derivatives is based on a merged formulation of the Feynman–Kac formulae in (1.3) and the chain rule formula of Malliavin calculus Nualart (2006). As we shall see, the resulting discrete time approximation of the Z process possesses the same order of conditional variance as the ones of the Y process, making the scheme significantly more attractive in a regression Monte Carlo framework compared to classical Euler discretizations. On the other hand, our formulation carries an extra layer of difficulty, in that we are forced to approximate the *'the Z of the Z, i.e. Γ processes'* (Gobet & Turkedjiev, 2017, Pg.1184) in the Malliavin BSDE, which are, in light of (1.3), related to the Hessian matrix of the solution of the corresponding parabolic problem (1.2). In this regard, our setting shares similarities with *second-order backward SDEs (2BSDEs)* Cheridito *et al.* (2007) and fully nonlinear problems Fahim *et al.* (2011). We analyze the discrete time approximation errors and show that under certain assumptions the new scheme has the same $\mathbb{L}^2$ convergence rate of order $1/2$ as the backward Euler scheme of BSDEs (Bouchard & Touzi, 2004).

Two fully-implementable approaches are investigated to solve the resulting discretization. First, we provide an extension to the BSDE-COS (BCOS) method (Ruijter & Oosterlee, 2015) and approximate solutions to one-dimensional problems by Fourier cosine expansions. Ultimately, the presence of Γ

estimates induces d^2 many additional conditional expectations to be approximated at each point in time, which makes the OSM scheme less tractable for classical Monte Carlo parametrizations when d is large. Thereafter, inspired by the encouraging results of Deep BSDE methods in case of high-dimensional equations, we propose a neural network least-squares Monte Carlo approach similar to the one of Huré *et al.* (2020), where the Y, Z and Γ processes are parametrized by fully-connected, feedforward deep neural networks. Subsequently, parameters of these networks are optimized in a recursive fashion, backwards over time, where at each time step two distinct gradient descent optimizations are performed, minimizing losses corresponding to the aforementioned discretization. Motivated by the UAT property of neural networks in Sobolev spaces, similarly to Huré *et al.* (2020), we consider two variants of the latter approach: one in which the Γ process is parametrized by a matrix-valued deep neural network; and one in which the Γ process is approximated as the Jacobian of the parametrization of the Z process, inspired by (1.3). The total approximation error is investigated similarly to Huré *et al.* (2020); Germain *et al.* (2021) and shown to be consistent with the discretization under the assumption of perfectly converging gradient descent iterations. We demonstrate the accuracy and robustness of our problem formulation with numerical experiments. In particular, using BCOS as a benchmark method for one-dimensional problems, we empirically assess the regression errors induced by gradient descent. We provide examples up to $d = 50$ dimensions.

The rest of the paper is organized as follows. In Section 2, we provide the necessary theoretical foundations, followed by Section 3 where the new discrete scheme is formulated. In Section 4, a discrete time approximation error analysis is given, bounding the total discretization error of the proposed scheme. Section 5 is concerned with the implementation of the discretization scheme, giving two fully-implementable approaches for the arising conditional expectations. First, the BCOS method (Ruijter & Oosterlee, 2015) is extended in case of one-dimensional problems, then a Deep BSDE (Han *et al.*, 2018; Huré *et al.*, 2020) approach is formulated for high-dimensional equations. A complete regression error analysis is provided, building on the universal approximation properties of neural networks. Our analysis is concluded by numerical experiments presented in Section 6, which confirm the theoretical results and showcase great accuracy over a wide range of different problems.

2. Backward stochastic differential equations and Malliavin calculus

In the following section, we introduce the notions of BSDEs and Malliavin calculus used throughout the paper.

2.1 *Preliminaries*

Let us fix $0 \leq T < \infty$ and $d, q, n, k \in \mathbb{N}^+$. We are concerned with a filtered probability space $\left(\Omega, \mathcal{F}, \mathbb{P}, \{\mathcal{F}\}_{0\leq t\leq T}\right)$, where $\mathcal{F} = \mathcal{F}_T$ and $\{\mathcal{F}\}_{0\leq t\leq T}$ is the natural filtration generated by a d-dimensional Brownian motion $\{W_t\}_{0\leq t\leq T}$ augmented by $\mathbb{P}$-null sets of Ω. In what follows, all equalities concerning $\mathcal{F}_t$-measurable random variables are meant in the $\mathbb{P}$-a.s. sense and all expectations—unless otherwise stated—are meant under $\mathbb{P}$. Throughout the whole paper, we rely on the following notations:

- $|x| := [\mathrm{Tr}\, x^T x]^{1/2}$ for the Frobenius norm of any $x \in \mathbb{R}^{q\times d}$. In case of scalar and vector inputs this coincides with the standard Euclidean norm. Additionally, we put $\langle x|y\rangle$ for the Euclidean inner product of $x, y \in \mathbb{R}^d$.
- $\mathbb{S}^p(\mathbb{R}^{q\times d})$ for the space of continuous and progressively measurable stochastic processes $Y : \Omega \times [0, T] \to \mathbb{R}^{q\times d}$ such that $\mathbb{E}\left[\sup_{0\leq t\leq T} |Y|^p\right] < \infty$.

- $\mathbb{H}^p(\mathbb{R}^{q\times d})$ for the space of progressively measurable stochastic processes $Z : \Omega \times [0,T] \to \mathbb{R}^{q\times d}$ such that $\mathbb{E}\left[\left(\int_0^T |Z_t|^2 \,\mathrm{d}t\right)^{p/2}\right] < \infty$.
- $\mathbb{L}^p_{\mathcal{F}_t}(\mathbb{R}^{q\times d})$ for the space of $\mathcal{F}_t$-measurable random variables $\xi : \Omega \to \mathbb{R}^{q\times d}$ such that $\mathbb{E}\left[|\xi|^p\right] < \infty$.
- $L^2([0,T];\mathbb{R}^q)$ for the Hilbert space of deterministic functions $h : [0,T] \to \mathbb{R}^q$ such that $\int_0^T |h(t)|^2 \,\mathrm{d}t < \infty$. Additionally, we denote its inner product by $\langle h|g\rangle_{L^2} := \int_0^T \langle h(t)|g(t)\rangle \,\mathrm{d}t$.
- $\nabla_x f := \left(\frac{\partial f}{\partial x_1}, \ldots, \frac{\partial f}{\partial x_d}\right)$ for the gradient of a scalar-valued, multivariate function $(t,x,y,z) \mapsto f(t,x,y,z)$ with respect to $x \in \mathbb{R}^d$, defined as a row vector, and analogously for $\nabla_y f, \nabla_z f$. Similarly, we denote the Jacobian matrix of a vector-valued function $\psi : \mathbb{R}^d \to \mathbb{R}^q$ by $\nabla_x \psi \in \mathbb{R}^{q\times d}$. For notational convenience, we set the Jacobian matrix of row and column vector-valued functions in the same fashion.
- $C^k_b(\mathbb{R}^d;\mathbb{R}^q)$, $C^k_p(\mathbb{R}^d;\mathbb{R}^q)$ for the set of k-times continuously differentiable functions $\varphi : \mathbb{R}^d \to \mathbb{R}^q$ such that all partial derivatives up to order k are bounded or have polynomial growth, respectively.
- $\mathbb{E}_n[\Phi] := \mathbb{E}\left[\Phi|\mathcal{F}_{t_n}\right]$ for conditional expectations with respect to the natural filtration, given a time partition $0 = t_0 < t_1 < \cdots < t_N = T$. We occasionally use the notation $\mathbb{E}^x_n[\Phi] := \mathbb{E}\left[\Phi|X_{t_n} = x\right]$ when the filtration is generated by a Markov process X.
- $\mathbf{1}_{q,d}, \mathbf{0}_{q,d}$ for $\mathbb{R}^{q\times d}$ matrices full of ones and zeros, respectively.

By slight abuse of notation, we put $\mathbb{S}^p(\mathbb{R}) := \mathbb{S}^p(\mathbb{R}^{1\times 1})$, $\mathbb{H}^p(\mathbb{R}^d) := \mathbb{H}^p(\mathbb{R}^{1\times d})$, $\mathbf{1}_d := \mathbf{1}_{1,d}$ and $\mathbf{0}_d := \mathbf{0}_{1\times d}$.

We recall the most important notions of Malliavin differentiability and refer to Nualart (2006) for a more detailed account on the subject. Consider the space of random processes $W(h) := \int_0^T h(t)\,\mathrm{d}W_t$ with $h \in L^2([0,T];\mathbb{R}^{1\times d})$. Let us now define the subspace $\mathcal{R} \subseteq \mathbb{L}^2_{\mathcal{F}_T}$ of smooth, scalar-valued random variables, which are of the form $\Phi = \varphi(W(h_1),\ldots,W(h_n))$ with some $\varphi \in C^\infty_p(\mathbb{R}^n;\mathbb{R})$. The Malliavin derivative of Φ is then defined as the $\mathbb{R}^{1\times d}$-valued stochastic process $D_s\Phi := \sum_{i=1}^n \partial_i\varphi(W(h_1),\ldots,W(h_n))h_i(s)$. The derivative operator can be extended to the closure of $\mathcal{R}$ with respect to the norm

$$\|\Phi\|_{\mathbb{D}^{1,p}} := \left(\mathbb{E}\left[|\Phi|^p + \left(\int_0^T \left|D_s\Phi\right|^2 \,\mathrm{d}s\right)^{p/2}\right]\right)^{1/p},$$

see Nualart (2006, Prop.1.2.1). We denote this closure as the space of Malliavin differentiable, $\mathbb{R}$-valued random variables by $D^{1,p}(\mathbb{R})$. For the space of vector-valued $\Phi = (\Phi_1,\ldots,\Phi_q)$ Malliavin differentiable random variables, we put $\Phi \in \mathbb{D}^{1,p}(\mathbb{R}^q)$ when $\Phi_i \in \mathbb{D}^{1,p}(\mathbb{R})$ for each $i = 1,\ldots,q$. The Malliavin derivative $D_s\Phi \in \mathbb{R}^{q\times d}$ is then the matrix-valued stochastic process whose ith row is $D_s\Phi_i$. The final result that extends the chain rule of elementary calculus to the Malliavin differentiation operator is fundamental for the present paper, essentially enabling the formulation of the upcoming discrete scheme.

LEMMA 2.1 (Malliavin chain rule lemma). Let $\psi \in C_b^1(\mathbb{R}^d;\mathbb{R}^q)$ and fix $p \geq 1$. Consider $F \in \mathbb{D}^{1,p}(\mathbb{R}^d)$. Then $\psi(F) \in \mathbb{D}^{1,p}(\mathbb{R}^q)$, furthermore for each $0 \leq s \leq T$

$$D_s\psi(F) = \nabla_x\psi(F)D_sF. \tag{2.1}$$

The lemma can be relaxed to the case where ψ is only Lipschitz continuous—see Nualart (2006, Prop.1.2.4).

2.2 *Backward stochastic differential equations*

We first provide the necessary theoretical foundations for the well-posedness of the underlying FBSDE system in (1.1) guaranteeing the existence of a unique solution triple. Given the stronger assumptions later required for their Malliavin differentiability, we restrict the presentation to standard Lipschitz assumptions. For a more general exposure, we refer to Chassagneux & Richou (2016) and the references therein.

It is well-known—see, e.g., Karatzas & Shreve (1998)—that the SDE in (1.1a) admits to a unique strong solution $\{X_t\}_{0\leq t\leq T} \in \mathbb{S}^p(\mathbb{R}^{d\times 1})$ whenever $x_0 \in \mathbb{L}^p_{\mathcal{F}_0}(\mathbb{R}^{d\times 1})$ and μ, σ are Lipschitz continuous in the spatial variable, i.e.,

$$\left|\mu(t,x_1) - \mu(t,x_2)\right| + \left|\sigma(t,x_1) - \sigma(t,x_2)\right| \leq L_{\mu,\sigma}\left|x_1 - x_2\right| \tag{2.2}$$

for all $t \in [0,T]$, $x_1, x_2 \in \mathbb{R}^{d\times 1}$, with some $L_{\mu,\sigma} > 0$. Additionally, the solution $\{X_t\}_{0\leq t\leq T}$ satisfies the following estimates for all $p \geq 1$

$$\mathbb{E}\left[\sup_{0\leq t\leq T}\left|X_t\right|^p\right] \leq C_p, \qquad \mathbb{E}\left[\left|X_t - X_s\right|^p\right] \leq C_p\left|t-s\right|^{p/2}, \tag{2.3}$$

with constant C_p only depending on p, T, d. In case of the Arithmetic Brownian Motion (ABM) with constant μ and σ, (1.1a) admits to the unique solution $X_t = x_0 + \mu t + \sigma W_t$. In particular, the Malliavin chain rule formula in Lemma 2.1 implies that $D_sX_t = \mathbb{1}_{s\leq t}\sigma$.

The well-posedness of the backward equation in (1.1b) is guaranteed—see, e.g., El Karoui *et al.* (1997)—by the Lipschitz continuity of the driver, on top of the polynomial growth of the terminal condition

$$\left|f(t,x,y_1,z_1) - f(t,x,y_2,z_2)\right| \leq L_{f,g}\left(\left|y_1 - y_2\right| + \left|z_1 - z_2\right|\right), \qquad \left|f(t,x,y,z)\right| + \left|g(x)\right| \leq L_{f,g}\left(1 + \left|x\right|^p\right), \tag{2.4}$$

for any $t \in [0,T]$, $y_1, y_2 \in \mathbb{R}^q$, $z_1, z_2 \in \mathbb{R}^{q\times d}$, with some $L_{f,g} > 0$ and $p \geq 2$. These conditions, combined with the ones for the SDEs above, imply the existence of a unique solution pair $Y \in \mathbb{S}^p(\mathbb{R}^q)$, $Z \in \mathbb{H}^p(\mathbb{R}^{q\times d})$ satisfying (1.1b). Let us now fix $q = 1$ and restrict the further analysis to scalar-valued backward equations. Thereafter, under the aforementioned conditions, the FBSDE system in (1.1) admits to a unique solution triple $\{(X_t, Y_t, Z_t)\}_{0\leq t\leq T} \in \mathbb{S}^p(\mathbb{R}^{d\times 1}) \times \mathbb{S}^p(\mathbb{R}) \times \mathbb{H}^p(\mathbb{R}^{1\times d})$.

2.3 *Malliavin differentiable FBSDE systems*

This paper is focused on a special class of FBSDE systems such that the solution triple $\{(X_t, Y_t, Z_t)\}_{0\le t\le T}$ is differentiable in the Malliavin sense. The Malliavin differentiability of the forward equation is guaranteed by the following theorem due to Nualart in Nualart (2006, Thm.2.2.1).

LEMMA 2.2 (Malliavin differentiability of SDEs, Nualart, 2006). Let $x_0 \in \mathbb{L}^p_{\mathcal{F}_0}(\mathbb{R}^{d\times 1})$, $\mu \in C_b^{0,1}([0,T]\times \mathbb{R}^{d\times 1};\mathbb{R}^{d\times 1})$, $\sigma \in C_b^{0,1}([0,T]\times\mathbb{R}^{d\times 1};\mathbb{R}^{d\times d})$ and $\mu(t,0), \sigma(t,0)$ be uniformly bounded for all $0\le t\le T$. Put $\{X_t\}_{0\le t\le T}$ for the unique solution of (1.1a). Then for all $t\in[0,T]$, $X_t \in \mathbb{D}^{1,p}(\mathbb{R}^{d\times 1})$ and there exists a continuous modification of its Malliavin derivative $\{D_sX_t\}_{0\le s,t\le T} \in \mathbb{S}^p(\mathbb{R}^{d\times d})$, which satisfies the linear SDE

$$D_sX_t = \mathbb{1}_{s\le t}\left\{\sigma(s,X_s) + \int_s^t \nabla_x\mu(r,X_r)D_sX_r\,\mathrm{d}r + \int_s^t \nabla_x\sigma(r,X_r)D_sX_r\,\mathrm{d}W_r.\right\}, \tag{2.5}$$

where $\nabla_x\sigma$ denotes a $\mathbb{R}^{d\times d\times d}$-valued tensor with $\left[\nabla_x\sigma\right]_{ijk} = \partial_k[\sigma]_{ij}$. Furthermore, there exists a constant C_p, only depending on p, T, d, such that

$$\sup_{s\in[0,T]}\mathbb{E}\left[\sup_{t\in[s,T]}\left|D_sX_t\right|^p\right] \le C_p, \qquad \mathbb{E}\left[\left|D_sX_r - D_sX_t\right|^p\right] \le C_p\left|r-t\right|^{p/2}, \quad \forall r,t\ge s. \tag{2.6}$$

The main implication of the proposition above is that under relatively mild assumptions on the bounded continuous differentiability of the coefficients in (1.1a), the Malliavin derivative of the solution satisfies a linear SDE, where the random coefficients depend on the solution of the SDE itself. Intriguingly, a similar assertion can be made about the solution pair of the backward equation in (1.1b), which—on top of establishing their Malliavin differentiability—also creates a connection between the Malliavin derivative DY and the control process. This is stated by the following theorem originally from Pardoux & Peng (1992), which we state under the loosened conditions of El Karoui *et al.* (1997, Prop.5.9).

THEOREM 2.3 (Malliavin differentiability of BSDEs, El Karoui *et al.*, 1997). Let the coefficients of (1.1a) satisfy the conditions of Lemma 2.2 and assume $f \in C_b^{0,1,1,1}([0,T]\times\mathbb{R}^{d\times 1},\mathbb{R},\mathbb{R}^{1\times d};\mathbb{R})$, $g\in C_b^1(\mathbb{R}^{d\times 1};\mathbb{R})$. Fix $p\ge 2$. Put $\{(Y_t,Z_t)\}_{0\le t\le T}$ for the unique solution pair of (1.1b). Then for all $t\in[0,T]$ $Y_t\in\mathbb{D}^{1,2}(\mathbb{R})$, $Z_t\in\mathbb{D}^{1,2}(\mathbb{R}^{1\times d})$ and there exist modifications of their Malliavin derivatives $\{D_sY_t\}_{0\le s,t\le T}\in\mathbb{S}^p(\mathbb{R}^{1\times d})$, $\{D_sZ_t\}_{0\le,s,t\le T}\in\mathbb{H}^p(\mathbb{R}^{d\times d})$, which satisfy the following linear BSDE:

$$\begin{aligned} D_sY_t &= \nabla_x g(X_T)D_sX_T \\ &\quad + \int_t^T \nabla_x f(r,X_r,Y_r,Z_r)D_sX_r + \nabla_y f(r,X_r,Y_r,Z_r)D_sY_r + \nabla_z f(r,X_r,Y_r,Z_r)D_sZ_r\,\mathrm{d}r \\ &\quad - \int_t^T \left((D_sZ_r)^T\mathrm{d}W_r\right)^T, \quad 0\le s\le t\le T, \\ D_sY_t &= \mathbf{0}_d, \quad D_sZ_t = \mathbf{0}_{d,d}, \quad 0\le t<s\le T. \end{aligned} \tag{2.7}$$

Furthermore, there exists a continuous modification of the control process such that $Z_t = D_t Y_t$ almost surely for all $0 \leq t \leq T$.

We emphasize the linearity of (2.7) and remark that the corresponding random coefficients of the linear equation depend on the solution of (1.1). Henceforth, in light of Lemma 2.2 and Theorem 2.3, we define $\{D_s X_t\}_{0\leq s,t\leq T}$ and $\{D_s Y_t\}_{0\leq s,t\leq T}, \{D_s Z_t\}_{0\leq s,t\leq T}$ as the versions of the corresponding Malliavin derivatives satisfying (2.5) and (2.7), respectively. For the rest of the paper, in order to ease the presentation, we introduce the notations $\mathbf{X}_t := (X_t, Y_t, Z_t)$, $\mathbf{D}_s\mathbf{X}_t := (D_s X_t, D_s Y_t, D_s Z_t)$ and $f^D(t, \mathbf{X}_t, \mathbf{D}_s\mathbf{X}_t) := \nabla_x f(t, \mathbf{X}_t) D_s X_t + \nabla_y f(t, \mathbf{X}_t) D_s Y_t + \nabla_z f(t, \mathbf{X}_t) D_s Z_t$ for all $0 \leq s, t \leq T$.

Path regularity and Hölder continuity. For $\{X_t\}_{0\leq t\leq T} \in \mathbb{S}^p(\mathbb{R}^{d\times 1})$, we have that the solution of the forward SDE is a continuous $\mathbb{R}^{d\times 1}$-valued random process, which is bounded in the supremum norm. Similar statements can be made about its Malliavin derivative $\{D_s X_t\}_{0\leq s,t\leq T}$. In particular, the Hölder regularity estimates in (2.3) and (2.6) ensure that the corresponding processes are not just continuous, but also have a modification admitting to α-Hölder continuous trajectories of order $\alpha \in (0, 1/2)$ provided by the Kolmogorov–Chentsov theorem—see, e.g., Karatzas & Shreve (1998). Since the $1/2$-Hölder regularity of (Y, Z) plays a crucial role in the convergence analysis of the discrete scheme—see Theorem 4.3 in particular—we elaborate on the conditions under which the continuous parts of the solutions to (1.1b) and (2.7) admit to similar estimates. Indeed, one can show that if the solutions $(Y, Z) \in \mathbb{S}^p(\mathbb{R}) \times \mathbb{H}^p(\mathbb{R}^{d\times 1})$ of (1.1b) satisfy the condition $\sup_{0\leq t\leq T} \mathbb{E}\left[|Z_t|^p\right] < \infty$ then there exists a constant C_p such that

$$\mathbb{E}\left[|Y_t - Y_s|^p\right] \leq C_p\, |t - s|^{p/2}\,, \tag{2.8}$$

see Hu *et al.* (2011, Corollary 2.7). In particular, the Y process admits to an α-Hölder continuous modification of order $\alpha \in (0, 1/2 - 1/p)$. Under the conditions of Theorem 2.3, this is naturally guaranteed, and for $p = 2$ it implies the *mean-squared continuity* of the Y process. Moreover, the Z process admits to a continuous modification solving (2.7), which guarantees $Z \in \mathbb{S}^p(\mathbb{R}^{1\times d})$ and, in particular, boundedness in the supremum norm. Under stronger assumptions, one can also establish a similar path regularity result of the control process. Imkeller & Dos Reis (2010, Thm.5.5) show that with additional conditions, essentially requiring second-order bounded differentiability of the corresponding coefficients μ, σ, f and g, the following also holds for all $p \geq 2$

$$\mathbb{E}\left[|Z_t - Z_s|^p\right] \leq C_p\, |t - s|^{p/2}\,. \tag{2.9}$$

Hu *et al.* prove a similar result in (Hu *et al.*, 2011, Thm.2.6) under slightly different assumptions in the general non-Markovian framework. We omit the explicit presentation of the necessary conditions for (2.9) to hold, nevertheless emphasize that Assumption 4.1 of the convergence analysis in Section 4 ensures the path regularity of the Z process and in particular implies mean-squared continuous trajectories.

3. The discrete scheme

In the following section, the proposed discretization scheme is introduced. The objective of the discretization is to simultaneously solve the pair of FBSDE systems given by (1.1) and the FBSDE system of its Malliavin derivatives provided by Lemma 2.2 and Theorem 2.3. Therefore, we are concerned with

the solution to the following pair of FBSDE systems:

$$X_t = x_0 + \int_0^t \mu(r, X_r)\,\mathrm{d}r + \int_0^t \sigma(r, X_r)\,\mathrm{d}W_r, \tag{3.1a}$$

$$Y_t = g(X_T) + \int_t^T f(r, \mathbf{X}_r)\,\mathrm{d}r - \int_t^T Z_r\,\mathrm{d}W_r, \tag{3.1b}$$

$$D_s X_t = \mathbb{1}_{s\le t}\left[\sigma(s, X_s) + \int_s^t \nabla_x \mu(r, X_r) D_s X_r\,\mathrm{d}r + \int_s^t \nabla_x \sigma(r, X_r) D_s X_r\,\mathrm{d}W_r\right], \tag{3.1c}$$

$$D_s Y_t = \mathbb{1}_{s\le t}\left[\nabla_x g(X_T) D_s X_T + \int_t^T f^D(r, \mathbf{X}_r, \mathbf{D}_s\mathbf{X}_r)\,\mathrm{d}r - \int_t^T \left((D_s Z_r)^T\,\mathrm{d}W_r\right)^T\right]. \tag{3.1d}$$

The solution is a pair of triples of stochastic processes $\{(X_t, Y_t, Z_t)\}_{0\le t\le T}$ and $\{(D_s X_t, D_s Y_t, D_s Z_t)\}_{0\le s,t\le T}$ such that (3.1) holds $\mathbb{P}$ almost surely. Consider a discrete time partition $\pi^N := \{t_0, \ldots, t_N\}$ with $0 = t_0 < t_1 < \cdots < t_N = T$ and set $\Delta W_n := W_{t_{n+1}} - W_{t_n}$, $\Delta t_n := t_{n+1} - t_n$, $|\pi| := \max_{0\le n\le N-1} t_{n+1} - t_n$. We denote the discrete time approximations by $\mathbf{X}_n^\pi := (X_n^\pi, Y_n^\pi, Z_n^\pi)$ and $\mathbf{D}_n\mathbf{X}_m^\pi := (D_n X_m^\pi, D_n Y_m^\pi, D_n Z_m^\pi)$ for each $0 \le n, m \le N$.

The forward component in (3.1a) is approximated by the classical Euler–Maruyama scheme, i.e.,

$$X_0^\pi := x_0, \quad X_{n+1}^\pi := X_n^\pi + \mu\left(t_n, X_n^\pi\right)\Delta t_n + \sigma\left(t_n, X_n^\pi\right)\Delta W_n^\pi, \tag{3.2}$$

for each $n = 0, \ldots, N-1$. It is well-known—see, e.g., Kloeden & Platen (1992)—that under standard Lipschitz assumptions on the drift and diffusion coefficients, these estimates admit to

$$\limsup_{|\pi|\to 0} \frac{1}{|\pi|}\mathbb{E}\left[\left|X_{t_n} - X_n^\pi\right|^2\right] < \infty. \tag{3.3}$$

Classically, the backward component in (3.1b) is approximated in two steps. In order to meet the necessary adaptivity requirements of the solution pair (Y, Z), one takes appropriate conditional expectations of (3.1b) and the same equation multiplied with the Brownian increment ΔW_n^T. Using standard properties of stochastic integrals, Itô's isometry and a *theta-discretization* of the remaining time integrals with parameters $\vartheta_y, \vartheta_z > 0$ subsequently give—see, e.g., Ruijter & Oosterlee (2015)

$$Y_N^\pi = g\left(X_N^\pi\right), \quad Z_N^\pi = \nabla_x g\left(X_N^\pi\right)\sigma(t_N, X_N^\pi), \tag{3.4a}$$

$$Z_n^\pi = -\frac{1-\vartheta_z}{\vartheta_z}\mathbb{E}_n\left[Z_{n+1}^\pi\right] + \frac{1}{\Delta t_n \vartheta_z}\mathbb{E}_n\left[\Delta W_n^T Y_{n+1}^\pi\right] + \frac{1-\vartheta_z}{\vartheta_z}\mathbb{E}_n\left[\Delta W_n^T f(t_{n+1}, \mathbf{X}_{n+1}^\pi)\right], \tag{3.4b}$$

$$Y_n^\pi = \Delta t_n \vartheta_y f\left(t_n, X_n^\pi, Y_n^\pi, Z_n^\pi\right) + \mathbb{E}_n\left[Y_{n+1}^\pi\right] + \Delta t_n(1-\vartheta_y)\mathbb{E}_n\left[f\left(t_{n+1}, \mathbf{X}_{n+1}^\pi\right)\right]. \tag{3.4c}$$

In case $\vartheta_y = \vartheta_z = 1$, this scheme is called the standard *Euler scheme for BSDEs*.

3.1 *The OSM scheme*

The novelty of the hereby proposed discretization is that on top of solving (3.1b), we also solve the linear BSDE in (3.1d) driving the Malliavin derivatives of the solution pair. Exploiting the relation between DY and Z established by Theorem 2.3, we set the control estimates according to the discrete time approximations of the Malliavin BSDE. As in the case of the forward component itself, the Malliavin derivative in (3.1c) is approximated by an Euler–Maruyama discretization, giving estimates

$$D_n X^\pi_m := \begin{cases} \mathbb{1}_{m=n}\sigma\left(t_n, X^\pi_n\right), & 0 \le m \le n \le N, \\ D_n X^\pi_{m-1} + \nabla_x \mu\left(t_{m-1}, X^\pi_{m-1}\right) D_n X^\pi_{m-1} \Delta t_{m-1} & \\ \qquad + \nabla_x \sigma\left(t_{m-1}, X^\pi_{m-1}\right) D_n X^\pi_{m-1} \Delta W_{m-1}, & 0 \le n < m \le N. \end{cases} \tag{3.5}$$

Unlike in the case of X^π_n, the convergence of these approximations is not straightforward due to the fact that the initial condition $D_n X^\pi_n = \sigma\left(t_n, X^\pi_n\right)$ already depends on the discrete approximation X^π_n provided by (3.2). Nonetheless, as we shall soon see, our discretization of the linear BSDE in (3.1d) only relies on the approximations $D_n X^\pi_{n+1}$ for each $n = 0, \dots, N-1$. This is a significant relaxation of the convergence criterion, as it can be shown that under relatively mild assumptions on the coefficients in (3.1a), $D_n X^\pi_{n+1}$ defined by (3.5) inherits the convergence rate of (3.3)—see Appendix A for details.

The discretization of the backward component in (3.1d) is done as follows. For any $n = 0, \dots, N-1$

$$D_{t_n} Y_{t_n} = D_{t_n} Y_{t_{n+1}} + \int_{t_n}^{t_{n+1}} f^D\left(r, \mathbf{X}_r, \mathbf{D}_{t_n}\mathbf{X}_r\right) \mathrm{d}r - \int_{t_n}^{t_{n+1}} \left(\left(D_{t_n} Z_r\right)^T \mathrm{d}W_r\right)^T, \tag{3.6}$$

subject to the terminal condition. Multiplying this equation with ΔW_n from the left, Itô's isometry implies

$$\begin{aligned} \mathbb{E}_n\left[\int_{t_n}^{t_{n+1}} D_{t_n} Z_r \,\mathrm{d}r\right] &= \mathbb{E}_n\left[\Delta W_n \left(D_{t_n} Y_{t_{n+1}} + \int_{t_n}^{t_{n+1}} f^D\left(r, \mathbf{X}_r, \mathbf{D}_{t_n}\mathbf{X}_r\right) \mathrm{d}r\right)\right], \\ D_{t_n} Y_{t_n} &= \mathbb{E}_n\left[D_{t_n} Y_{t_{n+1}} + \int_{t_n}^{t_{n+1}} f^D\left(r, \mathbf{X}_r, \mathbf{D}_{t_n}\mathbf{X}_r\right) \mathrm{d}r\right]. \end{aligned} \tag{3.7}$$

In order to avoid implicitness on Y, we approximate the continuous time integrals with the left and right rectangle rules, respectively, and obtain discrete time approximations

$$D_n Z^\pi_n = \frac{1}{\Delta t_n}\mathbb{E}_n\left[\Delta W_n\left(D_n Y^\pi_{n+1} + \Delta t_n f^D\left(t_{n+1}, \mathbf{X}^\pi_{n+1}, \mathbf{D}_n \mathbf{X}^\pi_{n+1}\right)\right)\right], \tag{3.8}$$

$$D_n Y^\pi_n = \mathbb{E}_n\left[D_n Y^\pi_{n+1} + \Delta t_n f^D\left(t_{n+1}, \mathbf{X}^\pi_{n+1}, \mathbf{D}_n \mathbf{X}^\pi_{n+1}\right)\right]. \tag{3.9}$$

At this point, to make the scheme viable, one relies on estimates $D_n Y^\pi_m, D_n Z^\pi_m$ on top of the Euler–Maruyama approximations of DX given by (3.5). This is done by a merged formulation of the Feynman–Kac formulae in (1.3) and the Malliavin chain rule in Lemma 2.1. Indeed, given the Markov nature of the FBSDE system, the solutions of (3.1b) can be written as $Y_t = y(t, X_t), Z_t = z(t, X_t)$ for some sufficiently

smooth deterministic functions $y : [0,T] \times \mathbb{R}^{d\times 1} \to \mathbb{R}$, $z : [0,T] \times \mathbb{R}^{d\times 1} \to \mathbb{R}^{1\times d}$. Moreover, the Malliavin chain rule implies that

$$D_{t_n} Y_r = \nabla_x y(r, X_r) D_{t_n} X_r, \qquad D_{t_n} Z_r = \nabla_x z(r, X_r) D_{t_n} X_r =: \gamma(r, X_r) D_{t_n} X_r, \tag{3.10}$$

for some deterministic functions $y : [0,T]\times\mathbb{R}^{d\times 1} \to \mathbb{R}$ and $z : [0,T]\times\mathbb{R}^{d\times 1} \to \mathbb{R}^{1\times d}$, where we defined $\gamma : [0,T] \times \mathbb{R}^{d\times 1} \to \mathbb{R}^{d\times d}$ as the Jacobian matrix of $z(r, X_r)$, and similarly $\Gamma_t := \gamma(t, X_t)$. Furthermore, due to the Feynman–Kac relations, we also have $z(r, X_r) = \nabla_x y(r, X_r)\sigma(r, X_r)$ and therefore

$$D_{t_n} Y_r = z(r, X_r)\sigma^{-1}(r, X_r) D_{t_n} X_r, \qquad D_{t_n} Z_r = \gamma(r, X_r) D_{t_n} X_r. \tag{3.11}$$

Motivated by these relations, we approximate the discretized Malliavin derivatives in (3.8) according to

$$D_n Y_m^\pi := Z_m^\pi \sigma^{-1}\left(t_m, X_m^\pi\right) D_n X_m^\pi, \qquad D_n Z_m^\pi := \Gamma_m^\pi D_n X_m^\pi, \qquad 0 \le n, m \le N. \tag{3.12}$$

Henceforth, the discrete approximations of the Y process driven by (3.1b) are given in an identical fashion to (3.4c) with $\vartheta_y \in [0,1]$ as a free parameter of the discretization. Moreover, in order to be able to control the $\mathbb{L}^2$ projection error of $D_n Z_m^\pi$ with discrete Grönwall estimates—see Step 1 of Theorem 4.3 in particular—we make the $\nabla_z f$ part of f^D implicit in $D_n Z_n^\pi$, and introduce the notation $\mathbf{D}_n \mathbf{X}_{n+1,n}^\pi := \left(D_n X_{n+1}^\pi, D_n Y_{n+1}^\pi, D_n Z_n^\pi\right)$. Subject to the terminal conditions in (3.1b) and (3.1d), on top of the Malliavin chain rule estimates in (3.12), this leads to the following discrete scheme, which we shall call the *One Step Malliavin* (OSM) scheme

$$Y_N^\pi = g\left(X_N^\pi\right), \quad Z_N^\pi = \nabla_x g\left(X_N^\pi\right)\sigma\left(t_N, X_N^\pi\right), \quad \Gamma_N^\pi = [\nabla_x(\nabla_x g\sigma)]\left(t_N, X_N^\pi\right), \tag{3.13a}$$

$$\Gamma_n^\pi \sigma\left(t_n, X_n^\pi\right) = D_n Z_n^\pi = \frac{1}{\Delta t_n}\mathbb{E}_n\left[\Delta W_n\left(D_n Y_{n+1}^\pi + \Delta t_n f^D\left(t_{n+1}, \mathbf{X}_{n+1}^\pi, \mathbf{D}_n \mathbf{X}_{n+1,n}^\pi\right)\right)\right], \tag{3.13b}$$

$$Z_n^\pi = \mathbb{E}_n\left[D_n Y_{n+1}^\pi + \Delta t_n f^D\left(t_{n+1}, \mathbf{X}_{n+1}^\pi, \mathbf{D}_n \mathbf{X}_{n+1,n}^\pi\right)\right], \tag{3.13c}$$

$$Y_n^\pi = \vartheta_y \Delta t_n f\left(t_n, X_n^\pi, Y_n^\pi, Z_n^\pi\right) + \mathbb{E}_n\left[Y_{n+1}^\pi + (1-\vartheta_y)\Delta t_n f\left(t_{n+1}, \mathbf{X}_{n+1}^\pi\right)\right]. \tag{3.13d}$$

The scheme is made fully implementable by an appropriate parametrization to approximate the arising conditional expectations.

REMARK 3.1 (Comparison of discretizations). There are two key differences between the standard Euler discretization in (3.4) and the OSM scheme in (3.13). First, unlike in the former, the OSM scheme's solution is a triple of discrete random processes, including an additional layer of Γ estimates. Moreover, it can be seen that the estimate in (3.13c) exhibits a better conditional variance than that of (3.4b). In case of the standard Euler discretization, the Z process is approximated through Itô's isometry and the corresponding discrete time approximations include a $1/\Delta t_n$ factor—second term in (3.4b)—which leads to a quadratically exploding conditional variance of the resulting estimates. Several variance reduction techniques have been proposed to mitigate this problem—we mention Alanko & Avellaneda (2013); Gobet & Turkedjiev (2017). On the other hand, within the OSM scheme, the Z process is approximated by the continuous solution of the Malliavin BSDE in (3.1d) and therefore it carries the same conditional

variance behavior as the Y estimate. In case of a fully-implementable regression Monte Carlo setting, this explains why the OSM scheme may provide more accurate control approximations.

Alternative formulations. (3.13) is not the first approach to the BSDE problem building on Theorem 2.3. Turkedjiev (2015) proposed a discrete time approximation scheme, where the Z process is estimated by an integration by parts formula stemming from Malliavin calculus and discovered in Ma & Zhang (2002, Thm.3.1). Hu *et al.* (2011) proposed an explicit scheme in the case of non-Markovian BSDEs, where the control process is estimated using a representation formula implied by the linearity of the Malliavin BSDE (3.1d)—see El Karoui *et al.* (1997, Prop.5.5). Briand & Labart (2014) offer a different approach to BSDEs, where building on chaos expansion formulas, the Z process is taken as the Malliavin derivative of Y given by Theorem 2.3. The difference between these formulations and (3.13) is mostly twofold. The OSM scheme is concerned with solving the entire pair of FBSDE systems (3.1) and not just the backward component in (3.1b). This means that, unlike in Hu *et al.* (2011); Briand & Labart (2014); Turkedjiev (2015), discrete time approximations give Γ estimates as well. Additionally, one important difference in the OSM scheme compared to the approaches (Hu *et al.*, 2011; Turkedjiev, 2015) is that the conditional expectations in (3.13) project $\mathcal{F}_{t_{n+1}}$-measurable random variables onto $\mathcal{F}_{t_n}$, whereas in the case of those works the arguments of the conditional expectations are $\mathcal{F}_T$-measurable. An important implication of this difference is that—unlike Hu *et al.* (2011); Turkedjiev (2015)—in order to simulate the arguments of the arising conditional expectations in (3.13), one does not rely on discrete time approximations of the Malliavin derivatives $D_n X^\pi_m$ over the whole time window ($n \leq m \leq N$), but only in between adjacent time steps $D_n X^\pi_{n+1}$. This is an advantage from the convergence analysis perspective whenever one does not have analytical access to the trajectories of $\{D_s X_t\}_{0\leq s,t\leq T}$. In fact, ensuring the convergence of the Euler–Maruyama scheme for the Malliavin derivative in (3.5) for any $n \leq m \leq N$ is known to be nontrivial, see Hu *et al.* (2011, Remark 5.1). On the other hand, as shown in Appendix A, under suitable regularity assumptions, $D_n X^\pi_{n+1}$ converges in the $\mathbb{L}^2$-sense with a rate of $1/2$, which renders the convergence of the discrete time approximations of the OSM scheme possible.

4. Discretization error analysis

Having introduced the discrete scheme simultaneously solving the FBSDE system itself and the FBSDE system of its solutions' Malliavin derivatives, we investigate the errors induced by the discretization of continuous processes in (3.13). It is known—see Bouchard & Touzi (2004)—that the $\mathbb{L}^2$ discretization errors of the backward Euler scheme in (3.4) admit to

$$\max_{0\leq n\leq N} \mathbb{E}\big[\big|Y_{t_n} - Y^\pi_n\big|^2\big] + \mathbb{E}\left[\sum_{n=0}^{N-1} \int_{t_n}^{t_{n+1}} \big|Z_r - Z^\pi_n\big|^2 \,\mathrm{d}r\right] \leq C\Big(\mathbb{E}\big[\big|g(X_T) - g\left(X^\pi_N\right)\big|^2\big] + \varepsilon^Z(|\pi|) + |\pi|\Big), \tag{4.1}$$

where $\varepsilon^Z(|\pi|) := \mathbb{E}\big[\sum_{n=0}^{N-1} \int_{t_n}^{t_{n+1}} \big|Z_r - \bar{Z}^{n+1}_n\big|^2 \,\mathrm{d}r\big]$ with $\bar{Z}^{n+1}_n := 1/\Delta t_n \mathbb{E}_n\big[\int_{t_n}^{t_{n+1}} Z_r \,\mathrm{d}r\big]$ according to Zhang (2004). The purpose of the following section is to show a similar result for the proposed OSM scheme and prove that it is *consistent* in the $\mathbb{L}^2$-sense, i.e., the discrete time approximations errors converge to zero as the mesh size of the time partition $|\pi|$ vanishes. In particular, we shall see that under standard Lipschitz assumptions on the driver f of the BSDE (3.1b) and the driver f^D of the linear Malliavin BSDE (3.1d), and additive noise in the forward diffusion, the convergence is of order $\mathcal{O}(|\pi|^{1/2})$.

ASSUMPTION 4.1 The following assumptions are in place.

($\mathbf{A}^{\mu,\sigma}$) SDE

- ($\mathbf{A}_1^{\mu,\sigma}$) the forward equation has constant drift and diffusion coefficients (Arithmetic Brownian motion);
- ($\mathbf{A}_2^{\mu,\sigma}$) the forward SDE has a uniformly elliptic diffusion coefficient, i.e., for any $\zeta \in \mathbb{R}^{1\times d}$ there exists a $\beta > 0$ such that $\zeta\sigma\sigma^T\zeta^T > \beta\,|\zeta|^2$ [1];

($\mathbf{A}^{f,g}$) BSDE

- ($\mathbf{A}_1^{f,g}$) $g \in C_b^{2+\alpha}(\mathbb{R})$ with some $\alpha > 0$, furthermore g is also bounded;
- ($\mathbf{A}_2^{f,g}$) $f \in C_b^{0,2,2,2}(\mathbb{R})$;
- ($\mathbf{A}_3^{f,g}$) f and its partial derivatives $\nabla_x f, \nabla_y f, \nabla_z f$ are all 1/2-Hölder continuous in time.

The conditions above are not minimal—see also Subsection 4.2. Nevertheless, for the sake of the present analysis they are sufficient. In particular, since bounded continuous differentiability implies Lipschitz continuity due to the mean-value theorem, by Theorem 2.3 we have that under Assumption 4.1 the FBSDE (3.1a)–(3.1b) is Malliavin differentiable, and the Malliavin derivatives of its solutions satisfy the FBSDE (3.1c)–(3.1d). Additionally, due to Delarue & Menozzi (2006, Thm. 2.1), we can also exploit the following useful result from the theory of parabolic PDEs.

LEMMA 4.2 (Delarue & Menozzi, 2006). Under Assumption 4.1, the parabolic PDE in (1.2) admits to a unique solution $u \in C_b^{1,2}(\mathbb{R})$.

Thereafter, provided by Lemma 4.2, one can use the merged formulation of the Malliavin chain rule lemma Lemma 2.1 and the nonlinear Feynman–Kac relations given by (3.11), in order to get the explicit formulas for the solutions of (3.1d) depending only on time and the state variable. We remark that in our setting $\sigma \in \mathbb{R}^{d\times d}$, the existence of the inverse is guaranteed by the uniform ellipticity condition set on σ in Assumption 4.1. In case the Brownian motion and the forward diffusion have different dimensions, similar statements can be made about right inverses—see Turkedjiev (2015). Another important implication of the estimate above is that Assumption 4.1, through Lemma 4.2, also implies that the driver of the Malliavin BSDE f^D is Lipschitz continuous in its spatial arguments within the bounded domain. Indeed, the mean-value theorem for $f \in C_b^{0,2,2,2}(\mathbb{R})$ implies that f and all its first-order derivatives in (x, y, z) are Lipschitz continuous, consequently for any uniformly bounded argument (DX, DY, DZ) the following holds:

$$\begin{aligned} |f(t_1,\mathbf{x}_1) - f(t_2,\mathbf{x}_2)| &\le L_f\big(|t_1-t_2|^{1/2} + |x_1-x_2| + |y_1-y_2| + |z_1-z_2|\big), \\ |\xi_1|, |\eta_1|, |\zeta_1| \le L_{f^D} : |f^D(t_1,\mathbf{x}_1,\boldsymbol{\xi}_1) - f^D(t_2,\mathbf{x}_2,\boldsymbol{\xi}_2)| &\le L_{f^D}\big(|t_1-t_2|^{1/2} + |x_1-x_2| + |y_1-y_2| + |z_1-z_2| \\ &\quad + |\xi_1-\xi_2| + |\eta_1-\eta_2| + |\zeta_1-\zeta_2|\big), \end{aligned} \tag{4.2}$$

[1] We remark that this condition is equivalent to $A = \sigma\sigma^T$ being a positive definite matrix.

with $\mathbf{x}_i = (x_i, y_i, z_i)$, $\boldsymbol{\xi}_i := (\xi_i, \eta_i, \zeta_i)$, $i = 1, 2$; for all $t_i \in [0, T]$, $x_i \in \mathbb{R}^{d\times 1}$, $y_i \in \mathbb{R}$, $z_i, \eta_i \in \mathbb{R}^{1\times d}$ and $\xi_i, \zeta_i \in \mathbb{R}^{d\times d}$, where $L_f, L_{f^D} > 0$. Here we also used the assumption of Hölder continuity established by $(\mathbf{A}_3^{f,g})$.

Given the usual time partition, it is clear that the discrete approximations (3.13) are deterministic functions of X_n^π and thereupon we put $Y_n^\pi =: y^\pi\left(t_n, X_n^\pi\right) =: y_n^\pi\left(X_n^\pi\right), Z_n^\pi =: z^\pi\left(t_n, X_n^\pi\right) =: z_n^\pi\left(X_n^\pi\right)$, $\Gamma_n^\pi =: \gamma^\pi\left(t_n, X_n^\pi\right) =: \gamma_n^\pi\left(X_n^\pi\right)$. In light of (3.12), we use the approximations

$$D_n Y_{n+1}^\pi = Z_{n+1}^\pi \sigma^{-1}\left(t_{n+1}, X_{n+1}^\pi\right) D_n X_{n+1}^\pi, \qquad D_n Z_n^\pi = \Gamma_n^\pi D_n X_n^\pi. \tag{4.3}$$

We introduce the short-hand notations $\Delta X_n^\pi := X_{t_n} - X_n^\pi$, $\Delta Y_n^\pi = Y_{t_n} - Y_n^\pi, \Delta Z_n^\pi = Z_{t_n} - Z_n^\pi$, $\Delta D_n X_{n+1}^\pi := D_{t_n} X_{t_{n+1}} - D_n X_{n+1}^\pi$, $\Delta D_n Y_{n+1}^\pi := D_{t_n} Y_{t_{n+1}} - D_n Y_{n+1}^\pi$ and $\Delta\Gamma_n^\pi := \Gamma_{t_n} - \Gamma_n^\pi$. Under the conditions of Assumption 4.1, provided by Lemma 2.2 and Theorem 2.3, we have that the processes (X, Y, Z, DX, DY) are all mean-squared continuous in time, i.e., there exists a general constant C such that for all $s, t, r \in [0, T]$

$$\begin{aligned} &\mathbb{E}\left[\left|X_t - X_r\right|^2\right] \le C\left|t - r\right|, \quad \mathbb{E}\left[\left|Y_t - Y_r\right|^2\right] \le C\left|t - r\right|, \quad \mathbb{E}\left[\left|Z_t - Z_r\right|^2\right] \le C\left|t - r\right|, \\ &\mathbb{E}\left[\left|D_s Y_t - D_s Y_r\right|^2\right] \le C\left|t - r\right|, \quad \mathbb{E}\left[\left|D_s X_t - D_s X_r\right|^2\right] \le C\left|t - r\right|, \quad \forall r, t \ge s. \end{aligned} \tag{4.4}$$

Finally, we use

$$\overline{DZ}_n^{n+1} := \frac{1}{\Delta t_n}\mathbb{E}_n\left[\int_{t_n}^{t_{n+1}} D_{t_n} Z_r \,\mathrm{d}r\right] \tag{4.5}$$

for the $\mathbb{L}^2$-projection of the corresponding Malliavin derivative with respect to the $\mathcal{F}_{t_n}$ σ-algebra, with which we can define the $\mathbb{L}^2(\mathbb{R}^{d\times d})$-regularity of DZ as follows:

$$\varepsilon^{DZ}(|\pi|) := \sum_{n=0}^{N-1}\mathbb{E}\left[\int_{t_n}^{t_{n+1}}\left|D_{t_n} Z_r - \overline{DZ}_n^{n+1}\right|^2 \mathrm{d}r\right]. \tag{4.6}$$

Under the condition of constant diffusion coefficients in Assumption 4.1, we have that $D_{t_n} Z_r = D_{t_m} Z_r = \Gamma_r \sigma$ for any $t_n, t_m < r$. Thereafter, exploiting the fact that due to Assumption 4.1 the terminal condition of the Malliavin BSDE (3.1d) is also Lipschitz continuous, one can apply Zhang (2004, Thm.3.1) and get

$$\limsup_{|\pi|\to 0}\frac{1}{|\pi|}\varepsilon^{DZ}(|\pi|) < \infty. \tag{4.7}$$

4.1 *Discrete-time approximation error*

The main goal of this section is to give an upper bound for the discrete time approximation errors defined by

$$\mathcal{E}^{\pi}(|\pi|) := \max_{0\le n\le N}\mathbb{E}\big[\big|\Delta Y_n^{\pi}\big|^2\big] + \max_{0\le n\le N}\mathbb{E}\big[\big|\Delta Z_n^{\pi}\big|^2\big] + \mathbb{E}\left[\sum_{n=0}^{N-1}\int_{t_n}^{t_{n+1}}\big|\big(\Gamma_r - \Gamma_n^{\pi}\big)\,\sigma\big|^2\,\mathrm{d}r\right] \le C\,|\pi|. \tag{4.8}$$

This is established by the following theorem.

THEOREM 4.3 (Consistency of the OSM scheme). Under Assumption 4.1, the scheme defined by (3.13) for any $\vartheta_y \in [0,1]$ has $\mathbb{L}^2$-convergence of order $1/2$, i.e.,

$$\limsup_{|\pi|\to 0}\frac{1}{|\pi|}\mathcal{E}^{\pi}(|\pi|) < \infty. \tag{4.9}$$

Proof. Throughout the proof, C denotes a constant independent of the time partition, whose value may vary from line to line. We proceed in steps and prove estimates for each component of the discretization error.

Step 1. *Estimate for DZ.* First, we establish an estimate for the corresponding discretization error of the DZ-component with respect to the $\mathbb{L}^2$-projection $\overline{DZ}_n^{n+1}$. Let us fix $n = 0,\dots,N-1$. Comparing (3.7) with (4.5), we find

$$\Delta t_n\overline{DZ}_n^{n+1} = \mathbb{E}_n\big[\Delta W_n D_{t_n}Y_{t_{n+1}}\big] + \mathbb{E}_n\left[\Delta W_n\int_{t_n}^{t_{n+1}} f^D\left(r,\mathbf{X}_r,\mathbf{D}_{t_n}\mathbf{X}_r\right)\mathrm{d}r\right]. \tag{4.10}$$

Combining this with the definition of the discrete scheme ((3.13b)) gives

$$\begin{aligned}\Delta t_n\big(\overline{DZ}_n^{n+1} - D_nZ_n^{\pi}\big) &= \mathbb{E}_n\left[\Delta W_n\left(\Delta D_nY_{n+1}^{\pi} - \mathbb{E}_n\left[\Delta D_nY_{n+1}^{\pi}\right]\right)\right]\\ &\quad + \mathbb{E}_n\left[\Delta W_n\left(\int_{t_n}^{t_{n+1}} f^D\left(r,\mathbf{X}_r,\mathbf{D}_{t_n}\mathbf{X}_r\right) - f^D\left(t_{n+1},\mathbf{X}_{n+1}^{\pi},\mathbf{D}_n\mathbf{X}_{n+1,n}^{\pi}\right)\mathrm{d}r\right)\right],\end{aligned} \tag{4.11}$$

using the tower property of conditional expectations. In Frobenius norm, the conditional $\mathbb{L}^2(\mathbb{R}^d)$ Cauchy–Schwarz inequality subsequently implies

$$\begin{aligned}\Delta t_n\big|\overline{DZ}_n^{n+1} - D_nZ_n^{\pi}\big| &\le (d\Delta t_n)^{1/2}\left(\mathbb{E}_n\left[\left|\Delta D_nY_{n+1}^{\pi} - \mathbb{E}_n\left[\Delta D_nY_{n+1}^{\pi}\right]\right|^2\right]\right)^{1/2}\\ &\quad + (d\Delta t_n)^{1/2}\left(\mathbb{E}_n\left[\left|\int_{t_n}^{t_{n+1}} f^D\big(r,\mathbf{X}_r,\mathbf{D}_{t_n}\mathbf{X}_r\big) - f^D\big(t_{n+1},\mathbf{X}_{n+1}^{\pi},\mathbf{D}_n\mathbf{X}_{n+1,n}^{\pi}\big)\,\mathrm{d}r\right|^2\right]\right)^{1/2},\end{aligned} \tag{4.12}$$

by the independence of Brownian increments. Hence, due to the $L^2([0,T];\mathbb{R}^d)$ Cauchy–Schwarz inequality, we gather

$$\begin{aligned}\Delta t_n\big|\overline{DZ}_n^{n+1}-D_nZ_n^\pi\big| \le (d\Delta t_n)^{1/2}\big(\mathbb{E}_n\big[\,\big|\Delta D_nY_{n+1}^\pi-\mathbb{E}_n\left[\Delta D_nY_{n+1}^\pi\right]\big|^2\,\big]\big)^{1/2}\\ +d^{1/2}\Delta t_n\left(\mathbb{E}_n\left[\int_{t_n}^{t_{n+1}}\left|f^D\left(r,\mathbf{X}_r,\mathbf{D}_{t_n}\mathbf{X}_r\right)-f^D\left(t_{n+1},\mathbf{X}_{n+1}^\pi,\mathbf{D}_n\mathbf{X}_{n+1,n}^\pi\right)\right|^2\mathrm{d}r\right]\right)^{1/2}.\end{aligned} \tag{4.13}$$

Using the inequality $a,b\in\mathbb{R}:(a+b)^2\le 2(a^2+b^2)$, we collect the following $\mathbb{L}^2(\mathbb{R}^{d\times d})$ upper bound:

$$\begin{aligned}\Delta t_n\mathbb{E}\big[\big|\overline{DZ}_n^{n+1}-D_nZ_n^\pi\big|^2\big] \le 2d\big(\mathbb{E}\big[\big|\Delta D_nY_{n+1}^\pi\big|^2\big]-\mathbb{E}\big[\big|\mathbb{E}_n\left[\Delta D_nY_{n+1}^\pi\right]\big|^2\big]\big)\\ +2d\Delta t_n\mathbb{E}\left[\int_{t_n}^{t_{n+1}}\left|f^D\left(r,\mathbf{X}_r,\mathbf{D}_{t_n}\mathbf{X}_r\right)-f^D\left(t_{n+1},\mathbf{X}_{n+1}^\pi,\mathbf{D}_n\mathbf{X}_{n+1,n}^\pi\right)\right|^2\mathrm{d}r\right].\end{aligned} \tag{4.14}$$

According to (4.2), the uniform boundedness of $\mathbf{D}_{t_n}\mathbf{X}_r$ implies that f^D is Lipschitz continuous in all its spatial arguments and $1/2$-Hölder continuous in time, with a universal constant L_{f^D}. This, combined with the mean-squared continuities of the $X, Y, Z, D_{t_n}X$ and $D_{t_n}Y$ in (4.4), implies

$$\begin{aligned}\Delta t_n\mathbb{E}\big[\big|\overline{DZ}_n^{n+1}-D_nZ_n^\pi\big|^2\big] \le 2d\big(\mathbb{E}\big[\big|\Delta D_nY_{n+1}^\pi\big|^2\big]-\mathbb{E}\big[\big|\mathbb{E}_n\left[\Delta D_nY_{n+1}^\pi\right]\big|^2\big]\big)\\ +14dL_{f^D}^2\Delta t_n\Big\{C\Delta t_n^2+2\Delta t_n\big(\mathbb{E}\big[\big|\Delta X_{n+1}^\pi\big|^2\big]+\mathbb{E}\big[\big|\Delta Y_{n+1}^\pi\big|^2\big]+\mathbb{E}\big[\big|\Delta Z_{n+1}^\pi\big|^2\big]\big)\\ +2\Delta t_n\big(\mathbb{E}\big[\big|\Delta D_nX_{n+1}^\pi\big|^2\big]+\mathbb{E}\big[\big|\Delta D_nY_{n+1}^\pi\big|^2\big]\big)\\ +\mathbb{E}\left[\int_{t_n}^{t_{n+1}}\big|D_{t_n}Z_r-D_nZ_n^\pi\big|^2\,\mathrm{d}r\right]\Big\},\end{aligned} \tag{4.15}$$

where we again used $(a+b)^2\le 2(a^2+b^2)$ for $a,b\in\mathbb{R}$. By the definition of $\overline{DZ}_n^{n+1}$ in (4.5), the last term can be split as follows:

$$\mathbb{E}\left[\int_{t_n}^{t_{n+1}}\big|D_{t_n}Z_r-D_nZ_n^\pi\big|^2\,\mathrm{d}r\right]=\mathbb{E}\left[\int_{t_n}^{t_{n+1}}\big|D_{t_n}Z_r-\overline{DZ}_n^{n+1}\big|^2\,\mathrm{d}r\right]+\Delta t_n\mathbb{E}\big[\big|\overline{DZ}_n^{n+1}-D_nZ_n^\pi\big|^2\big]. \tag{4.16}$$

Plugging this back in (4.15) yields

$$\begin{aligned}\Delta t_n \mathbb{E}\big[\big|\overline{DZ}_n^{n+1} - D_n Z_n^\pi\big|^2\big] \le 2d\big(\mathbb{E}\big[\big|\Delta D_n Y_{n+1}^\pi\big|^2\big] - \mathbb{E}\big[\big|\mathbb{E}_n\big[\Delta D_n Y_{n+1}^\pi\big]\big|^2\big]\big) \\ + 14dL_{f^D}^2 \Delta t_n \Big\{ C\Delta t_n^2 + 2\Delta t_n \mathbb{E}\big[\big|\Delta X_{n+1}^\pi\big|^2\big] \\ + 2\Delta t_n\big(\mathbb{E}\big[\big|\Delta Y_{n+1}^\pi\big|^2\big] + \mathbb{E}\big[\big|\Delta Z_{n+1}^\pi\big|^2\big]\big) \\ + 2\Delta t_n\big(\mathbb{E}\big[\big|\Delta D_n X_{n+1}^\pi\big|^2\big] + \mathbb{E}\big[\big|\Delta D_n Y_{n+1}^\pi\big|^2\big]\big) \\ + \mathbb{E}\left[\int_{t_n}^{t_{n+1}} \big|D_{t_n} Z_r - \overline{DZ}_n^{n+1}\big|^2 \,\mathrm{d}r\right] \\ + \Delta t_n \mathbb{E}\big[\big|\overline{DZ}_n^{n+1} - D_n Z_n^\pi\big|^2\big]\Big\}. \end{aligned} \tag{4.17}$$

For sufficiently small time steps satisfying $14dL_{f^D}^2 \Delta t_n \le 1/2$, we can therefore gather the estimate

$$\begin{aligned}\Delta t_n \mathbb{E}\big[\big|\overline{DZ}_n^{n+1} - D_n Z_n^\pi\big|^2\big] \le 4d \left\{\mathbb{E}\big[\big|\Delta D_n Y_{n+1}^\pi\big|^2\big] - \mathbb{E}\big[\big|\mathbb{E}_n\big[\Delta D_n Y_{n+1}^\pi\big]\big|^2\big]\right\} \\ + 28dL_{f^D}^2 \Delta t_n \Big\{ C\Delta t_n^2 + 2\Delta t_n \mathbb{E}\big[\big|\Delta X_{n+1}^\pi\big|^2\big] \\ + 2\Delta t_n\big(\mathbb{E}\big[\big|\Delta Y_{n+1}^\pi\big|^2\big] + \mathbb{E}\big[\big|\Delta Z_{n+1}^\pi\big|^2\big]\big) \\ + 2\Delta t_n\big(\mathbb{E}\big[\big|\Delta D_n X_{n+1}^\pi\big|^2\big] + \mathbb{E}\big[\big|\Delta D_n Y_{n+1}^\pi\big|^2\big]\big) \\ + \mathbb{E}\left[\int_{t_n}^{t_{n+1}} \big|D_{t_n} Z_r - \overline{DZ}_n^{n+1}\big|^2 \,\mathrm{d}r\right]\Big\}. \end{aligned} \tag{4.18}$$

Step 2. *Estimate for Z*. With the above result in hand, we give an estimate for the control process. Under Assumption 4.1, provided by Theorem 2.3, we identify the control process Z by its continuous modification given by DY and establish pointwise estimates. Indeed, from (3.7) and the definition of the discrete scheme in (3.13c), it follows

$$\Delta Z_n^\pi = \mathbb{E}_n\left[\Delta D_n Y_{n+1}^\pi\right] + \mathbb{E}_n\left[\int_{t_n}^{t_{n+1}} f^D\left(r, \mathbf{X}_r, \mathbf{D}_{t_n}\mathbf{X}_r\right) - f^D\left(t_{n+1}, \mathbf{X}_{n+1}^\pi, \mathbf{D}_n \mathbf{X}_{n+1,n}^\pi\right) \mathrm{d}r\right]. \tag{4.19}$$

Applying the Young inequality of the form $(a+b)^2 \le (1+\rho\Delta t_n)a^2 + (1+\frac{1}{\rho\Delta t_n})b^2$ with any $\rho > 0$; using the Jensen and $L^2([0,T];\mathbb{R}^d)$ Cauchy–Schwarz inequalities gives

$$\begin{aligned}\mathbb{E}\big[\big|\Delta Z_n^\pi\big|^2\big] \le (1+\rho\Delta t_n)\mathbb{E}\big[\big|\mathbb{E}_n\left[\Delta D_n Y_{n+1}^\pi\right]\big|^2\big] \\ + \frac{1}{\rho}(1+\rho\Delta t_n)\mathbb{E}\left[\int_{t_n}^{t_{n+1}} \big|f^D\left(r, \mathbf{X}_r, \mathbf{D}_{t_n}\mathbf{X}_r\right) - f^D\left(t_{n+1}, \mathbf{X}_{n+1}^\pi, \mathbf{D}_n \mathbf{X}_{n+1,n}^\pi\right)\big|^2 \mathrm{d}r\right]. \end{aligned} \tag{4.20}$$

Exploiting the Lipschitz and Hölder continuity of f^D in (4.2) and using the mean-squared continuities of $X, Y, Z, D_{t_n}X$ and $D_{t_n}Y$ in (4.4), we subsequently gather

$$\begin{aligned}\mathbb{E}\big[\big|\Delta Z_n^\pi\big|^2\big] \le (1+\rho\Delta t_n)\mathbb{E}\big[\big|\mathbb{E}_n\left[\Delta D_n Y_{n+1}^\pi\right]\big|^2\big] \\ + \frac{7L_{f^D}^2}{\rho}(1+\rho\Delta t_n)\Big\{C\Delta t_n^2 + 2\Delta t_n\big(\mathbb{E}\big[\big|\Delta X_{n+1}^\pi\big|^2\big] + \mathbb{E}\big[\big|\Delta Y_{n+1}^\pi\big|^2\big] + \mathbb{E}\big[\big|\Delta Z_{n+1}^\pi\big|^2\big]\big) \\ + 2\Delta t_n\big(\mathbb{E}\big[\big|\Delta D_n X_{n+1}^\pi\big|^2\big] + \mathbb{E}\big[\big|\Delta D_n Y_{n+1}^\pi\big|^2\big]\big) \\ + \mathbb{E}\left[\int_{t_n}^{t_{n+1}} \big|D_{t_n}Z_r - D_n Z_n^\pi\big|^2\,\mathrm{d}r\right]\Big\}.\end{aligned} \tag{4.21}$$

Splitting the last term according to (4.16), substituting the upper bound (4.18) and choosing $\rho^* := 28dL_{f^D}^2$ then yields

$$\begin{aligned}\mathbb{E}\big[\big|\Delta Z_n^\pi\big|^2\big] \le (1+\rho^*\Delta t_n)\mathbb{E}\big[\big|\Delta D_n Y_{n+1}^\pi\big|^2\big] \\ + \frac{1+\rho^*\Delta t_n}{2}\Big\{C\Delta t_n^2 + \big(1+28dL_{f^D}^2\Delta t_n\big)\Delta t_n\big(\mathbb{E}\big[\big|\Delta X_{n+1}^\pi\big|^2\big] + \mathbb{E}\big[\big|\Delta Y_{n+1}^\pi\big|^2\big] + \mathbb{E}\big[\big|\Delta Z_{n+1}^\pi\big|^2\big]\big) \\ + \big(1+28dL_{f^D}^2\Delta t_n\big)\Delta t_n\big(\mathbb{E}\big[\big|\Delta D_n X_{n+1}^\pi\big|^2\big] + \mathbb{E}\big[\big|\Delta D_n Y_{n+1}^\pi\big|^2\big]\big) \\ + \frac{1+28dL_{f^D}^2\Delta t_n}{2}\mathbb{E}\left[\int_{t_n}^{t_{n+1}} \big|D_{t_n}Z_r - \overline{DZ}_n^{n+1}\big|^2\,\mathrm{d}r\right]\Big\},\end{aligned} \tag{4.22}$$

for any sufficiently small $\Delta t_n < 1$. At this point, we can make use of the fact that due to $(\mathbf{A}_1^{\mu,\sigma})$ in Assumption 4.1 $X_n^\pi = \sigma W_{t_n} = X_{t_n}$ and $D_n X_{n+1}^\pi = \sigma \equiv D_{t_n}X_{t_{n+1}}$, which in particular implies $X_{t_n} - X_n^\pi \equiv 0, D_{t_n}X_{t_{n+1}} - D_n X_{n+1}^\pi \equiv 0$ and

$$\Delta D_n Y_{n+1}^\pi = \Delta Z_{n+1}^\pi, \qquad D_{t_n}Z_{t_n} - D_n Z_n^\pi = \Delta\Gamma_n^\pi\sigma, \tag{4.23}$$

in light of (4.3). Plugging these estimates back in (4.22) subsequently gives

$$\begin{aligned}\mathbb{E}\big[\big|\Delta Z_n^\pi\big|^2\big] \le (1+C_z\Delta t_n)\mathbb{E}\big[\big|\Delta Z_{n+1}^\pi\big|^2\big] \\ + C_z\left\{\Delta t_n^2 + \Delta t_n\mathbb{E}\big[\big|\Delta Y_{n+1}^\pi\big|^2\big] + \mathbb{E}\left[\int_{t_n}^{t_{n+1}} \big|D_{t_n}Z_r - \overline{DZ}_n^{n+1}\big|^2\,\mathrm{d}r\right]\right\}.\end{aligned} \tag{4.24}$$

Step 3. *Estimate for Y*. Given f's Lipschitz continuity in (x, y, z) and 1/2-Hölder continuity in t by (4.2), the mean-squared continuities of X, Y and Z in (4.4); through subsequent applications of the Young–, Jensen– and Cauchy–Schwarz inequalities analogously to the previous steps, we derive the following

inequality from the dynamics of Y in (3.1b) and the discrete scheme in (3.13d):

$$\begin{aligned}\mathbb{E}\big[\big|\Delta Y_n^\pi\big|^2\big] \le (1+\beta\Delta t_n)\mathbb{E}\big[\big|\Delta Y_{n+1}^\pi\big|^2\big] \\ + \frac{8L_f^2}{\beta}(1+\beta\Delta t_n)\Big\{C\Delta t_n^2 + 2\vartheta_y^2\Delta t_n\big(\mathbb{E}\big[\big|\Delta Y_n^\pi\big|^2\big] + \mathbb{E}\big[\big|\Delta Z_n^\pi\big|^2\big]\big) \\ + 2(1-\vartheta_y)^2\Delta t_n\big(\mathbb{E}\big[\big|\Delta Y_{n+1}^\pi\big|^2\big] + \mathbb{E}\big[\big|\Delta Z_{n+1}^\pi\big|^2\big]\big)\Big\}, \end{aligned} \tag{4.25}$$

with any $\beta > 0$.

Step 4. *Combined estimate for Y and Z*. Combining the estimates in (4.24) and (4.25) gives

$$\begin{aligned}\left(1 - \frac{16L_f^2(1+\beta)\vartheta_y^2}{\beta}\Delta t_n\right)\big(\mathbb{E}\big[\big|\Delta Y_n^\pi\big|^2\big] + \mathbb{E}\big[\big|\Delta Z_n^\pi\big|^2\big]\big) \le (1+C_y\Delta t_n)\big(\mathbb{E}\big[\big|\Delta Y_{n+1}^\pi\big|^2\big] + \mathbb{E}\big[\big|\Delta Z_{n+1}^\pi\big|^2\big]\big) \\ + C\left\{\Delta t_n^2 + \mathbb{E}\left[\int_{t_n}^{t_{n+1}}\big|D_{t_n}Z_r - \overline{DZ}_n^{n+1}\big|^2\,\mathrm{d}r\right]\right\}, \end{aligned} \tag{4.26}$$

with $C_y = \beta + \frac{16L_f^2(1+\beta)}{\beta}(1-\vartheta_y)^2 + C_z$. Then, for any given $\beta > 0$ and sufficiently small time step admitting to $\frac{16L_f^2(1+\beta)\vartheta_y^2}{\beta}\Delta t_n < 1$, we derive

$$\begin{aligned}\mathbb{E}\big[\big|\Delta Y_n^\pi\big|^2\big] + \mathbb{E}\big[\big|\Delta Z_n^\pi\big|^2\big] \le (1+C\Delta t_n)\big(\mathbb{E}\big[\big|\Delta Y_{n+1}^\pi\big|^2\big] + \mathbb{E}\big[\big|\Delta Z_{n+1}^\pi\big|^2\big]\big) \\ + C\left\{\Delta t_n^2 + \mathbb{E}\left[\int_{t_n}^{t_{n+1}}\big|D_{t_n}Z_r - \overline{DZ}_n^{n+1}\big|^2\,\mathrm{d}r\right]\right\}. \end{aligned} \tag{4.27}$$

Thereupon, the discrete Grönwall lemma implies that

$$\begin{aligned}\max_{0\le n\le N}\mathbb{E}\big[\big|\Delta Y_n^\pi\big|^2\big] + \max_{0\le n\le N}\mathbb{E}\big[\big|\Delta Z_n^\pi\big|^2\big] \le C\Big\{\mathbb{E}\big[\big|g(X_T) - g\big(X_N^\pi\big)\big|^2\big] \\ + \mathbb{E}\big[\big|\nabla_x g(X_T)\sigma(t_N,X_T) - \nabla_x g\big(X_{t_N}^\pi\big)\sigma(t_N,X_N^\pi)\big|^2\big] \\ + \varepsilon^{DZ}(|\pi|) + |\pi|\Big\}, \end{aligned} \tag{4.28}$$

where we also used the definition in (4.6). The proclaimed estimate for the (Y,Z) part then follows from the observation that under Assumption 4.1 the terminal conditions of both the BSDE in (3.1b) and the Malliavin BSDE in (3.1d) are analytically observed; and the fact that, according to (4.7), $\varepsilon^{DZ}(|\pi|)$ is also $\mathcal{O}(|\pi|)$.

Step 5. *Final estimate for* Γ. It remains to show the consistency of the Γ estimate. From (4.14) and (4.16), we get

$$\begin{aligned}
&\mathbb{E}\Big[\int_{t_n}^{t_{n+1}} \big|D_{t_n} Z_r - D_n Z_n^\pi\big|^2 \,\mathrm{d}r\Big] \\
&\qquad \le \mathbb{E}\left[\int_{t_n}^{t_{n+1}} \big|D_{t_n} Z_r - \overline{DZ}_n^{n+1}\big|^2 \,\mathrm{d}r\right] + 2d\left\{\mathbb{E}\big[\big|\Delta D_n Y_{n+1}^\pi\big|^2\big] - \mathbb{E}\big[\big|\mathbb{E}_n\left[\Delta D_n Y_{n+1}^\pi\right]\big|^2\big]\right\} \\
&\qquad\quad + 2d\Delta t_n \mathbb{E}\left[\int_{t_n}^{t_{n+1}} \left|f^D\left(X_{t_N}^\pi\right)\left(r, \mathbf{X}_r, \mathbf{D}_{t_n}\mathbf{X}_r\right) - f^D\left(X_{t_N}^\pi\right)\left(t_{n+1}, \mathbf{X}_{n+1}^\pi, \mathbf{D}_n \mathbf{X}_{n+1,n}^\pi\right)\right|^2 \mathrm{d}r\right].
\end{aligned} \tag{4.29}$$

Summation from $n = 0, \ldots, N-1$ thus gives

$$\begin{aligned}
&\mathbb{E}\Big[\sum_{n=0}^{N-1}\int_{t_n}^{t_{n+1}} \big|D_{t_n} Z_r - D_n Z_n^\pi\big|^2 \,\mathrm{d}r\Big] \\
&\qquad \le \mathbb{E}\left[\sum_{n=0}^{N-1}\int_{t_n}^{t_{n+1}} \big|D_{t_n} Z_r - \overline{DZ}_n^{n+1}\big|^2 \,\mathrm{d}r\right] + 2d\mathbb{E}\big[\big|\Delta D_{N-1} Y_N^\pi\big|^2\big] \\
&\qquad\quad + 2d\sum_{n=1}^{N-1}\left\{\mathbb{E}\big[\big|\Delta D_{n-1} Y_n^\pi\big|^2\big] - \mathbb{E}\big[\big|\mathbb{E}_n\left[\Delta D_n Y_{n+1}^\pi\right]\big|^2\big]\right\} \\
&\qquad\quad + 2d\sum_{n=0}^{N-1}\Delta t_n \mathbb{E}\left[\int_{t_n}^{t_{n+1}} \left|f^D\left(r, \mathbf{X}_r, \mathbf{D}_{t_n}\mathbf{X}_r\right) - f^D\left(t_{n+1}, \mathbf{X}_{n+1}^\pi, \mathbf{D}_n \mathbf{X}_{n+1,n}^\pi\right)\right|^2 \mathrm{d}r\right],
\end{aligned} \tag{4.30}$$

where we changed the summation index for the first part of the third term. Using the relations in (4.23) implied by Assumption 4.1, we can upper bound the summation term by the estimate (4.20)

$$\begin{aligned}
&\mathbb{E}\left[\sum_{n=0}^{N-1}\int_{t_n}^{t_{n+1}} \big|D_{t_n} Z_r - D_n Z_n^\pi\big|^2 \,\mathrm{d}r\right] \\
&\quad \le \mathbb{E}\left[\sum_{n=0}^{N-1}\int_{t_n}^{t_{n+1}} \big|D_{t_n} Z_r - \overline{DZ}_n^{n+1}\big|^2 \,\mathrm{d}r\right] + 2d\mathbb{E}\big[\big|\Delta D_{N-1} Y_N^\pi\big|^2\big] \\
&\qquad + 2d\varrho\sum_{n=1}^{N-1}\Delta t_n \mathbb{E}\big[\big|\mathbb{E}_n\left[\Delta D_n Y_{n+1}^\pi\right]\big|^2\big] \\
&\qquad + 2d\sum_{n=0}^{N-1}(1/\varrho + 2\Delta t_n)\mathbb{E}\left[\int_{t_n}^{t_{n+1}} \left|f^D(r, \mathbf{X}_r, \mathbf{D}_{t_n}\mathbf{X}_r) - f^D(t_{n+1}, \mathbf{X}_{n+1}^\pi, \mathbf{D}_n \mathbf{X}_{n+1,n}^\pi)\right|^2 \mathrm{d}r\right],
\end{aligned} \tag{4.31}$$

for any $\varrho > 0$. Similar steps as in (4.21) subsequently give

$$\begin{aligned}
&\mathbb{E}\Big[\sum_{n=0}^{N-1}\int_{t_n}^{t_{n+1}} \big|D_{t_n}Z_r - D_nZ_n^\pi\big|^2\,\mathrm{d}r\Big] \\
&\quad\le \mathbb{E}\left[\sum_{n=0}^{N-1}\int_{t_n}^{t_{n+1}} \big|D_{t_n}Z_r - \overline{DZ}_n^{n+1}\big|^2\,\mathrm{d}r\right] + 2d\mathbb{E}\big[\big|\Delta D_{N-1}Y_N^\pi\big|^2\big] \\
&\qquad + 2d\varrho\sum_{n=1}^{N-1}\Delta t_n\mathbb{E}\big[\big|\mathbb{E}_n\left[\Delta D_nY_{n+1}^\pi\right]\big|^2\big] \\
&\qquad + 14L_{f^D}^2 d\sum_{n=0}^{N-1}(1/\varrho + 2\Delta t_n)\Big\{C\Delta t_n^2 + 2\Delta t_n\mathbb{E}\big[\big|\Delta X_{n+1}^\pi\big|^2\big] + 2\Delta t_n\mathbb{E}\big[\big|\Delta Y_{n+1}^\pi\big|^2\big] \\
&\qquad\qquad + 2\Delta t_n\mathbb{E}\big[\big|\Delta Z_{n+1}^\pi\big|^2\big] + 2\Delta t_n\mathbb{E}\big[\big|\Delta D_nX_{n+1}^\pi\big|^2\big] \\
&\qquad\qquad + 2\Delta t_n\mathbb{E}\big[\big|\Delta D_nY_{n+1}^\pi\big|^2\big] + \mathbb{E}\left[\int_{t_n}^{t_{n+1}} \big|D_{t_n}Z_r - D_nZ_n^\pi\big|^2\,\mathrm{d}r\right]\Big\}.
\end{aligned} \tag{4.32}$$

By choosing $\varrho^* = 56L_{f^D}^2 d$, we have that for any sufficiently small $|\pi|$ satisfying $28L_{f^D}^2 d\,|\pi| < 1/4$

$$\begin{aligned}
&\mathbb{E}\Big[\sum_{n=0}^{N-1}\int_{t_n}^{t_{n+1}} \big|D_{t_n}Z_r - D_nZ_n^\pi\big|^2\,\mathrm{d}r\Big] \\
&\quad\le 2\mathbb{E}\left[\sum_{n=0}^{N-1}\int_{t_n}^{t_{n+1}} \big|D_{t_n}Z_r - \overline{DZ}_n^{n+1}\big|^2\,\mathrm{d}r\right] + 4d\mathbb{E}\big[\big|\Delta D_{N-1}Y_N^\pi\big|^2\big] \\
&\qquad + 4d\varrho^*\sum_{n=1}^{N-1}\Delta t_n\mathbb{E}\big[\big|\mathbb{E}_n\left[\Delta D_nY_{n+1}^\pi\right]\big|^2\big] \\
&\qquad + \sum_{n=0}^{N-1}\left(1/2 + 56L_{f^D}^2 d\Delta t_n\right)\Big\{C\Delta t_n^2 + 2\Delta t_n\mathbb{E}\big[\big|\Delta X_{n+1}^\pi\big|^2\big] + 2\Delta t_n\mathbb{E}\big[\big|\Delta Y_{n+1}^\pi\big|^2\big] \\
&\qquad\qquad + 2\Delta t_n\big(\mathbb{E}\big[\big|\Delta Z_{n+1}^\pi\big|^2\big] + \mathbb{E}\big[\big|\Delta D_nX_{n+1}^\pi\big|^2\big] + \mathbb{E}\big[\big|\Delta D_nY_{n+1}^\pi\big|^2\big]\big)\Big\}.
\end{aligned} \tag{4.33}$$

Once again applying the relations in (4.23), Jensen's inequality, the convergence of the $\mathbb{L}^2$-regularity of DZ in (4.7) and the estimate (4.28) proven in the previous step now shows the proclaimed convergence of the Γ estimates.

This concludes the proof. □

The final result in (4.9) expresses that the $\mathbb{L}^2$ convergence rate of the discrete time approximations induced by (3.13) is of order $\mathcal{O}(|\pi|^{1/2})$ under the conditions imposed in Assumption 4.1. Comparing the convergence bound of Theorem 4.3 to that of the classical backward Euler discretization in (4.1),

three observations need to be made. First, in contrast to the backward Euler discretization, the OSM scheme admits to a bound where the Z process is controlled by the maximum error over the discrete time steps—see (4.8). This is due to the fact that under the OSM formulation, Theorem 2.3 guarantees a continuous version of the control process bounded in the supremum norm, and thus allows for pointwise estimates. Additionally, we see that even though the hereby proposed discretization solves a *larger problem* by incorporating Γ estimates, it exhibits the same, optimal rate of convergence well-known for the classical backward Euler discretization of BSDEs in (4.1). At last, unlike in the aforementioned case, our final estimate does not include the strong discretization errors of the terminal conditions of the BSDEs (3.1b) and (3.1d). This is merely due to the fact that under Assumption 4.1 we assumed constant diffusion coefficients, which led to the corresponding terms canceling in (4.28). Similarly, we exploited that under our conditions the Malliavin BSDE's terminal condition is Lipschitz continuous, leading to an $\mathcal{O}(|\pi|^{1/2})$ convergence of the $\mathbb{L}^2$-regularity of DZ according to (4.7). In case of irregular terminal conditions and nonanalytical forward diffusions, it is expected that the corresponding terms would also contribute to the final estimate.

4.2 *Assumptions revisited*

In order to conclude the discussion on the discrete time approximation errors, we elaborate on the conditions set in Assumption 4.1. Key aspects of their relevance are highlighted and potential ways to generalize the results are pointed out in order to encourage further research.

Not surprisingly, compared to classical discretizations excluding the Malliavin components, necessarily stricter conditions need to be posed in order to ensure Malliavin differentiability of the original FBSDE system in (3.1a)–(3.1b). The differentiability requirements on the coefficients f and g in $(\mathbf{A}_1^{f,g})$–$(\mathbf{A}_2^{f,g})$ are inherently linked to the Malliavin differentiability of the FBSDE in (3.1). However, the Malliavin differentiability of the solution pair holds under significantly milder assumptions. We refer to Mastrolia *et al.* (2017) for a recent account on the subject, where it is shown that first-order continuous differentiability, with not necessarily bounded $\nabla_x g, \nabla_x f$ is sufficient.

The reason why we nonetheless decided to restrict the assumptions to second-order bounded differentiability is mostly related to Lemma 4.2 and the Lipschitz continuity of f^D in (4.2). Although the Lipschitz continuity of $\nabla_x f, \nabla_y f, \nabla_z f$ are all guaranteed by the $C_b^{0,2,2,2}$ assumption, the same cannot be said about the Malliavin derivative arguments $\mathbf{D}_s\mathbf{X}_t$ of f^D. More precisely, in order to have Lipschitz continuity in all spatial arguments, one—on top of the boundedness of the partial derivatives of f—also needs to have the uniform boundedness of all the Malliavin derivatives (DX, DY, DZ). Due to the Malliavin chain rule estimates in (3.11), under the assumption of constant diffusion coefficients in $(\mathbf{A}_1^{\mu,\sigma})$, the uniform boundedness of the Malliavin derivatives is implied by the twice bounded differentiability of the solution of the parabolic problem in (1.2). This is guaranteed by Lemma 4.2, requiring the conditions in $(\mathbf{A}_1^{f,g})$–$(\mathbf{A}_2^{f,g})$ to be satisfied. In case the uniform boundedness of (DY, DZ) is not readily available, one can truncate the corresponding arguments of f^D similarly to Chassagneux & Richou (2016), and discretize the truncated Malliavin problem accordingly. Thereafter, the total discrete time approximation error can be decomposed into a truncation and discretization component, which guarantee convergence for an appropriately chosen, adaptive truncation range. A detailed presentation of this argument will be part of our future research.

Throughout the analysis, we also often relied on the assumption that the underlying forward diffusion admits to constant drift and diffusion coefficients due to $(\mathbf{A}_1^{\mu,\sigma})$. In particular, this assumption allowed us to neglect the contribution of error terms such as $\mathbb{E}\big[\big|X_{t_n} - X_n^\pi\big|^2\big]$ and $\mathbb{E}\big[\big|D_{t_n}X_{t_{n+1}} - D_n X_{n+1}^\pi\big|^2\big]$—see,

e.g., (4.23). However, it is well-known that the strong convergence of Euler–Maruyama approximations is of order $1/2$—see (3.3)—carrying the same order of convergence as the rest of the terms in our estimates. The convergence of the Malliavin derivative $D_n X^\pi_{n+1}$ with respect to an Euler–Maruyama discretization in (3.5) is more troublesome. In fact, as highlighted by related works in the literature—see Hu *et al.* (2011, Remark 5.1)—it is difficult to guarantee the convergence of $D_s X^\pi$ over the whole time horizon. It is important to highlight that the OSM scheme in (3.13) does not require approximations of the corresponding Malliavin derivative over the whole time window, but only in between adjacent time steps $D_n X^\pi_{n+1}$. This is a major relieve in terms of convergence as one can easily show that within this one time stepping (OSM) scheme, $D_n X^\pi_{n+1}$ inherits the convergence properties of the forward diffusion under mild assumptions—see Appendix A.

The main difficulty with respect to general forward diffusions is related to the Malliavin chain rule approximations given by (3.11). In fact, when $D_n X^\pi_{n+1} \neq D_{t_n} X_{t_{n+1}}$, one needs to deal with product terms such as

$$\begin{aligned} D_{t_n} Y_{t_{n+1}} - D_n Y^\pi_{n+1} = &\left[Z_{t_{n+1}} \sigma^{-1}\left(t_{n+1}, X_{t_{n+1}}\right) - Z^\pi_{n+1} \sigma^{-1}\left(t_{n+1}, X^\pi_{n+1}\right)\right] D_{t_n} X_{t_{n+1}} \\ &+ Z^\pi_{n+1} \sigma^{-1}\left(t_{n+1}, X^\pi_{n+1}\right)\left[D_{t_n} X_{t_{n+1}} - D_n X^\pi_{n+1}\right]. \end{aligned} \tag{4.34}$$

These pose a significant amount of difficulty when one—unlike in the case of $(\mathbf{A}_1^{\mu,\sigma})$—does not have the uniform boundedness of σ^{-1} and $\{D_s X_t\}_{0\leq s,t\leq T}$. Additionally, in order to ensure the boundedness of the discrete estimates Z^π_{n+1}, a certain truncation procedure would be required, further complicating the analysis. Therefore, we decided to restrict the assumptions to constant diffusion coefficients and to leave the general case for future research.

Remark 4.4 (Nonconstant drift and Girsanov's theorem).We remark that the assumption of a constant drift coefficient is mostly a matter convenience. Indeed, with a straightforward change of measure argument via the Girsanov theorem, one can merge the corresponding nonconstant drift contribution onto the driver of the BSDE and—as long as the drift itself satisfies the continuously bounded differentiable assumptions posed on $\nabla_x f$—the same analysis holds.

5. Fully implementable schemes with differentiable function approximators and neural networks

Having established a convergence result for the discrete time approximation's error induced by (3.13), we now turn to fully-implementable schemes where the appearing conditional expectations are numerically approximated by a certain machinery. In other words, we are concerned with the following modification of the discrete scheme in (3.13):

$$\widehat{Y}^\pi_N = g\left(X^\pi_N\right), \quad \widehat{Z}^\pi_N = \nabla_x g\left(X^\pi_N\right) \sigma\left(t_N, X^\pi_N\right), \quad \widehat{\Gamma}^\pi_N = [\nabla_x(\nabla_x g \sigma)](t_N, X^\pi_N), \tag{5.1a}$$

$$\check{\Gamma}^\pi_n \sigma\left(t_n, X^\pi_n\right) = D_n \check{Z}^\pi_n = \frac{1}{\Delta t_n} \mathbb{E}_n\left[\Delta W_n \left(D_n \widehat{Y}^\pi_{n+1} + \Delta t_n f^D\left(t_{n+1}, \widehat{\mathbf{X}}^\pi_{n+1}, \mathbf{D}_n \check{\mathbf{X}}^\pi_{n+1,n}\right)\right)\right], \quad \widehat{\Gamma}^\pi_n \leftarrow \mathcal{P}(\check{\Gamma}^\pi_n), \tag{5.1b}$$

$$\check{Z}_n^\pi = \mathbb{E}_n\big[D_n\widehat{Y}_{n+1}^\pi + \Delta t_n f^D(t_{n+1}, \widehat{\mathbf{X}}_{n+1}^\pi, \mathbf{D}_n\widehat{\mathbf{X}}_{n+1,n}^\pi)\big], \qquad \widehat{Z}_n^\pi \leftarrow \mathcal{P}(\check{Z}_n^\pi), \tag{5.1c}$$

$$\check{Y}_n^\pi = \vartheta_y \Delta t_n f\left(t_n, X_n^\pi, \check{Y}_n^\pi, \widehat{Z}_n^\pi\right) + \mathbb{E}_n\big[\widehat{Y}_{n+1}^\pi + (1-\vartheta_y)\Delta t_n f\big(t_{n+1}, \widehat{\mathbf{X}}_{n+1}^\pi\big)\big], \quad \widehat{Y}_n^\pi \leftarrow \mathcal{P}(\check{Y}_n^\pi), \tag{5.1d}$$

with the notations $\widehat{\mathbf{X}}_{n+1}^\pi := \big(\widehat{X}_{n+1}^\pi, \widehat{Y}_{n+1}^\pi, \widehat{Z}_{n+1}^\pi\big)$, $\mathbf{D}_n\check{\mathbf{X}}_{n+1,n}^\pi := (D_nX_{n+1}^\pi, D_n\widehat{Y}_{n+1}^\pi, D_n\check{Z}_n^\pi)$ and $\mathbf{D}_n\widehat{\mathbf{X}}_{n+1,n}^\pi := \big(D_nX_{n+1}^\pi, D_n\widehat{Y}_{n+1}^\pi, D_n\widehat{Z}_n^\pi\big)$, where $D_n\widehat{Y}_{n+1} := \widehat{Z}_{n+1}^\pi\sigma^{-1}(t_{n+1}, X_{n+1}^\pi)D_nX_{n+1}^\pi$ and $D_n\widehat{Z}_n^\pi := \widehat{\Gamma}_n^\pi D_nX_n^\pi$—similarly as in (3.12). The final approximations are denoted by $(\widehat{Y}_n^\pi, \widehat{Z}_n^\pi, \widehat{\Gamma}_n^\pi)$ and $\mathcal{P}$ denotes a *machinery*, which, given approximations at future time steps, estimates the *true* conditional expectations $(\check{Y}_n^\pi, \check{Z}_n^\pi, \check{\Gamma}_n^\pi)$. It is worth to notice that (5.1c) is explicit, whereas (5.1b) and (5.1d) are both implicit when $\vartheta_y > 0$. Due to the Markov feature of the corresponding problem, we can write all estimates as deterministic functions of the state process $\check{Y}_n^\pi =: \check{y}_n^\pi\left(X_n^\pi\right)$, $\check{Z}_n^\pi =: \check{z}_n^\pi\left(X_n^\pi\right)$, $\check{\Gamma}_n^\pi =: \check{\gamma}_n^\pi\left(X_n^\pi\right)$ and $\widehat{Y}_n^\pi =: \widehat{y}_n^\pi\left(X_n^\pi\right), \widehat{Z}_n^\pi =: \widehat{z}_n^\pi\left(X_n^\pi\right), \widehat{\Gamma}_n^\pi =: \widehat{\gamma}_n^\pi\left(X_n^\pi\right)$ at each time instance.

In the literature there exist several techniques to numerically approximate conditional expectations, see, e.g., Bally & Pagès (2003); Bouchard & Touzi (2004); Briand & Labart (2014). In what follows, we investigate two specific approaches in the context of the OSM scheme. We first give an extension to the BCOS method (Ruijter & Oosterlee, 2015), which shall later be used as a benchmark method for one-dimensional problems. Our main approximation tool is based on a least-squares Monte Carlo formulation similar to those of the Deep BSDE methods (Han *et al.*, 2018; Huré *et al.*, 2020), where the functions parametrizing the solution triple are fully-connected, feedforward neural networks. Due to the universal approximation properties of neural networks in Sobolev spaces, this will allow us to distinguish between two variants. In the first one, the Γ process is parametrized by a matrix-valued neural network whose parameters are optimized in a stochastic gradient descent iteration. In the second, this parametrization is circumvented and, in light of (1.3), the Γ estimates are directly calculated as the Jacobian of the Z process. However, such directly linked estimates induce an additional source of error, which shall be addressed in Theorem 5.2, where we give an error bound for the complete approximation error of the fully-implementable OSM scheme, given the cumulative regression errors of neural network regressions, similarly to the ones proven in Han & Long (2020); Huré *et al.* (2020).

5.1 *The BCOS method*

We recall the most fundamental notions of the BCOS method (Ruijter & Oosterlee, 2015). In order to keep the presentation concise, for the sake of this section, we restrict ourselves to the one-dimensional case. BCOS is an extension of the COS method (Fang & Oosterlee, 2009) to the setting of FBSDE systems, whose main idea is to recover the probability densities of certain random variables given that their characteristic function is available. The key idea of the BCOS method can be summarized as follows. In general, for a Markov problem, conditional expectations are of the form

$$I(x) := \mathbb{E}\left[v\left(t_{n+1}, X_{n+1}^\pi\right)|X_n^\pi = x\right] = \int_{\mathbb{R}} v(t_{n+1}, \rho)p(\rho|x)\,\mathrm{d}\rho, \tag{5.2}$$

where $p(\rho|x)$ is the conditional transition density function from state (t, x) to state (t_{n+1}, ρ). Assuming that the integrand above decays in the infinite limit, one can truncate the integration range to a sufficiently wide finite domain $[a, b]$. Thereafter, the Fourier cosine expansion of the deterministic

mapping $v(t_{n+1},\cdot):[a,b]\to\mathbb{R}$ reads as[2]

$$v(t_{n+1},\rho)=\sum_{k=0}^{\infty}{}'\,\mathcal{V}_k(t_{n+1})\cos\left(k\pi\frac{\rho-a}{b-a}\right), \tag{5.3}$$

where the series coefficients are given by $\mathcal{V}_k(t_{n+1}):=\frac{2}{b-a}\int_a^b v(t_{n+1},\rho)\cos(k\pi\frac{\rho-a}{b-a})\,\mathrm{d}\rho$. Plugging these estimates back in the conditional expectation, with an additional truncation of the Fourier expansion to a finite number of K coefficients, gives the approximation (Fang & Oosterlee, 2009)

$$I(x)\approx\widehat{I}(x):=\sum_{k=0}^{K-1}{}'\,\mathcal{V}_k(t_{n+1})\Re\{\Phi(k|x)\}, \tag{5.4}$$

where $\Phi(k|x):=\phi(\frac{k\pi}{b-a}|x)e^{ik\pi\frac{x-a}{b-a}}$ and $\phi(u|x)$ is the conditional characteristic function of the Markov transition. In case the underlying Markov process is an Euler–Maruyama approximation of the solution to a forward SDE, the conditional characteristic function is given by $\phi(u|x)=\exp(iu\mu(t_n,x)\Delta t_n-\frac{1}{2}u^2\sigma^2(t_n,x)\Delta t_n)$. Using an integration by parts argument—see Ruijter & Oosterlee (2015, Appendix A.1) and Appendix B—similar results can be constructed for conditional expectations of the forms

$$J(x):=\mathbb{E}_n^x\left[v\left(t_{n+1},X_{n+1}^{\pi}\right)\Delta W_n\right]\approx\widehat{J}(x):=\Delta t_n\sigma(t_n,x)\sum_{k=0}^{K-1}{}'-\frac{k\pi}{b-a}\mathcal{V}_k(t_{n+1})\Im\{\Phi(k|x)\}, \tag{5.5}$$

$$\begin{aligned}K(x):=\mathbb{E}_n^x\left[v\left(t_{n+1},X_{n+1}^{\pi}\right)(\Delta W_n)^2\right]&\approx\widehat{K}(x):=\Delta t_n\sum_{k=0}^{K-1}{}'\,\mathcal{V}_k(t_{n+1})\Re\{\Phi(k|x)\}\\&-\Delta t_n^2\sigma^2(t_n,x)\sum_{k=0}^{K-1}\left(\frac{k\pi}{b-a}\right)^2\mathcal{V}_k(t_{n+1})\Re\{\Phi(k|x)\}.\end{aligned} \tag{5.6}$$

Built on these approximations, the BCOS method goes as follows. One first needs to recover the coefficients of the terminal conditions either analytically or via Discrete Cosine Transforms (DCT). These coefficients are plugged into conditional expectations of the form (5.4), (5.5) and (5.6), providing estimates for the solutions at t_{N-1}. In order to make the scheme fully-implementable, one also relies on a machinery that recovers these coefficients while going to time step n, from time step $n+1$ in a backward recursive algorithm. This step can either be done by Fast Fourier Transforms (FFT) (Ruijter & Oosterlee, 2015) when the coefficients of the SDE are constant, or with DCT when they are not (Ruijter & Oosterlee, 2016). When one is faced with an implicit conditional expectation ($\vartheta_y>0$) Picard iterations are performed, which—under Lipschitz assumptions and sufficiently small time steps—converge exponentially fast to the unique fixed point solution.

[2] We adhere to the standard notation where $\sum_{k=0}^{\prime K-1}a_k:=a_0/2+\sum_{k=1}^{K-1}a_k$, i.e., the first element is multiplied by $1/2$.

In particular, the BCOS approximations for (5.1) read as follows—for a more detailed derivation, see Appendix C

$$\widehat{y}_N^{\pi}(x) = g(x), \quad \widehat{z}_N^{\pi}(x) = \partial_x g(x)\sigma(T,x), \quad \widehat{\gamma}_N^{\pi}(x) = \partial_x\left(\partial_x g\sigma\right)(T,x), \tag{5.7a}$$

$$\widehat{\gamma}_n^{\pi}(x)\sigma(t_n,x) = {\sum_{k=0}^{K-1}}' \widehat{\mathcal{DZ}}_k(t_{n+1})\cos\left(k\pi\frac{x-a}{b-a}\right), \tag{5.7b}$$

$$\begin{aligned}\widehat{z}_n^{\pi}(x) = {}& \sigma(t_n,x)(1+\partial_x\mu(t_n,x)\Delta t_n){\sum_{k=0}^{K-1}}' \mathcal{W}_k(t_{n+1})\Re\{\Phi(k|x)\} \\ & - \sigma^2(t_n,x)\partial_x\sigma(t_n,x)\Delta t_n {\sum_{k=0}^{K-1}}' \frac{k\pi}{b-a}\mathcal{W}_k(t_{n+1})\Im\{\Phi(k|x)\} \\ & + \Delta t_n\widehat{\gamma}_n^{\pi}(x)\sigma(t_n,x){\sum_{k=0}^{K-1}}' \mathcal{F}_k^z(t_{n+1})\Re\{\Phi(k|x)\},\end{aligned} \tag{5.7c}$$

$$\widehat{y}_n^{\pi}(x) = {\sum_{k=0}^{K-1}}' \widehat{\mathcal{Y}}_k(t_n)\cos\left(k\pi\frac{x-a}{b-a}\right), \tag{5.7d}$$

where we defined

$$\begin{aligned} h_{n+1}^{\pi}\left(X_{n+1}^{\pi}\right) &:= \widehat{y}_{n+1}^{\pi}\left(X_{n+1}^{\pi}\right) + (1-\vartheta_y)\Delta t_n f\left(t_{n+1}, X_{n+1}^{\pi}, \widehat{y}_{n+1}^{\pi}\left(X_{n+1}^{\pi}\right), \widehat{z}_{n+1}^{\pi}\left(X_{n+1}^{\pi}\right)\right), \\ w_{n+1}^{\pi}\left(X_{n+1}^{\pi}\right) &:= \left(1+\partial_y f\left(t_{n+1}, \widehat{\mathbf{X}}_{n+1}^{\pi}\right)\right)\widehat{z}_{n+1}^{\pi}\left(X_{n+1}^{\pi}\right)\sigma^{-1}\left(t_{n+1}, X_{n+1}^{\pi}\right) + \Delta t_n\partial_x f\left(t_n, \widehat{\mathbf{X}}_{n+1}^{\pi}\right)\end{aligned} \tag{5.8}$$

for the explicit parts of the discrete approximations (5.1d) and (5.1c), respectively. The coefficients

$$\begin{aligned} \mathcal{W}_k(t_{n+1}) &:= \frac{2}{b-a}\int_a^b w_{n+1}^{\pi}(\rho)\cos\left(k\pi\frac{\rho-a}{b-a}\right)\mathrm{d}\rho, \quad \mathcal{H}_k(t_{n+1}) := \frac{2}{b-a}\int_a^b h_{n+1}^{\pi}(\rho)\cos\left(k\pi\frac{\rho-a}{b-a}\right)\mathrm{d}\rho, \\ \mathcal{F}_k^z(t_{n+1}) &:= \frac{2}{b-a}\int_a^b \partial_z f(t_{n+1},\rho)\cos\left(k\pi\frac{\rho-a}{b-a}\right)\mathrm{d}\rho \end{aligned} \tag{5.9}$$

are approximated by their DCT counterparts $\widehat{\mathcal{W}}_k(t_{n+1}), \widehat{\mathcal{H}}_k(t_{n+1}), \widehat{\mathcal{F}}_k^z(t_{n+1})$, respectively. $\widehat{\mathcal{DZ}}_k(t_{n+1})$ is recovered with DCT on the approximations $\mathbb{E}_n^x\left[\Delta t_n^{-1}\Delta W_n w_{n+1}^{\pi}\left(X_{n+1}^{\pi}\right)D_n X_{n+1}^{\pi}\right]/\left(1-\mathbb{E}_n^x\left[\Delta W_n\partial_z f(t_{n+1}, \widehat{\mathbf{X}}_{n+1}^{\pi})\right]\right)$. Thereafter, the BCOS formulas in (5.4), (5.5) and (5.6), together with the Euler–Maruyama estimates (3.5), imply the estimates for Γ and Z. The Z estimates are plugged into the approximation of the Y process in (5.1d). The coefficients $\widehat{\mathcal{Y}}_k(t_n)$ are recovered from the estimates $y_n^{P,\pi}(x) = \vartheta_y\Delta t_n f(t_n, x, y_n^{P-1,\pi}(x), \widehat{z}_n^{\pi}(x)) + \mathbb{E}_n^x\left[h_{n+1}^{\pi}\right]$ after a sufficient number of P Picard iterations are taken. This completes the BCOS algorithm for the OSM scheme.

For a detailed account on the contributions of the corresponding truncation and approximation errors of the BCOS method, we refer to Fang & Oosterlee (2009); Ruijter & Oosterlee (2015, 2016) and the references therein. Although the method can be extended to higher-dimensional diffusion processes, it suffers from the curse of dimensionality through the inevitable spatial discretization required in the Fourier frequency domain.

5.2 *Neural networks*

In recent years, neural networks have shown excellent empirical results when deployed in a regression Monte Carlo framework for BSDEs (Han *et al.*, 2018; Fujii *et al.*, 2019; Huré *et al.*, 2020). In what follows, we are concerned with the class of feedforward, fully-connected deep neural networks, particularly in the context of approximating high-dimensional conditional expectations. This family of functions $\Psi(\cdot|\Theta) : \mathbb{R}^{d\times 1} \to \mathbb{R}^{q\times d}$ can be described as a hierarchical sequence of compositions

$$\Psi(x|\Theta) := a_{\text{out}} \circ A_{L+1}(\cdot|\theta_{L+1}) \circ a \circ A_L(\cdot|\theta_L) \circ a \circ \cdots \circ a \circ A_1(\cdot|\theta_1) \circ x. \tag{5.10}$$

The affine transformations $A_l, l = 1, \ldots, L$ are called *hidden layers* and are of the form $A_l(y|\theta_l := (W_{l-1}^l, b_l)) := W_{l-1}^l y + b_l$, with $W_{l-1}^l \in \mathbb{R}^{S_l \times S_{l-1}}$ being a matrix of *weights* and $b_l \in \mathbb{R}^{S_l \times 1}$, $S_{l-1}, S_l \in \mathbb{N}$ the *biases*. Furthermore, $a : \mathbb{R} \to \mathbb{R}$ describes a nonlinear *activation* function, which is applied element-wise on the output of each affine transformation. The size S_l denotes how many *neurons* are contained in the given layer. The *output layer* is defined by $A_{L+1}(y|\theta_{L+1} := (W_L^{L+1}, b_{L+1})) := W_L^{L+1} y + b_{L+1}$ with $W_L^{L+1} \in \mathbb{R}^{q\times d\times S_L}, b_{L+1} \in \mathbb{R}^{q\times d}$. The complete parameter space of such an architecture is therefore given by $\Theta := \left(\theta_1, \ldots, \theta_{L+1}\right) \in \mathbb{R}^{q\times d\times (S_L+1)+\sum_{l=1}^{L} S_{l-1}\times S_l + S_l}$. Widely common choices for the nonlinearity include: Rectified Linear Units (ReLU), sigmoid and the hyperbolic tangent activations. The optimal parameter space Θ^* is usually approximated by first formulating a *loss function*, which measures an abstract distance from the desired behavior, and then iteratively minimizing this loss through a *stochastic gradient descent* (SGD) type algorithm. For more details, we refer to Goodfellow *et al.* (2016).

The use of deep learning is often motivated by the so-called *Universal Approximation Theorems (UAT)*, which establish that neural networks can approximate a wide class of functions with arbitrary accuracy. The first version of the UAT property was proven by Cybenko (1989). However, as in the applications of this paper derivative approximations play an important role, we present the following extension of Hornik *et al.* (1990), which extends the UAT property to *Sobolev spaces*. In what follows, we use the common notations for $W^{k,p}(U) := \{f \in L^p(U) : \|f\|_{W^{k,p}} := (\sum_{|\alpha|\le k} \int_U |D^\alpha f|^p \, \mathrm{d}\lambda)^{1/p} < \infty\}$ for Sobolev spaces, where α denotes a multi-index, D^α is the differentiation operator in the weak sense and λ is the Lebesgue measure. In particular, we use $H^k(U) := W^{k,2}(U)$. Then the UAT in Sobolev spaces can be stated as follows—for a proof see Hornik *et al.* (1990, Corollary 3.6).

THEOREM 5.1 (Universal Approximation Theorem in Sobolev Spaces, Hornik *et al.*, 1990). Let $a : \mathbb{R} \to \mathbb{R}$ be an ℓ-finite activation function, i.e., $a \in C^l(\mathbb{R})$ and $\int_{\mathbb{R}} \left|D^\ell a\right| < \infty$. Let $U \subseteq \mathbb{R}^{d\times 1}$ be a compact subset. Denote the class of single hidden layer neural networks by $\Sigma(a) := \{\psi : \mathbb{R}^{d\times 1} \to \mathbb{R}^{1\times q} : \psi(x|\Theta = (W_0^1, b_1, W_1^2, b_2)) = W_1^2 a(W_0^1 x + b_1) + b_2, W_0^1 \in \mathbb{R}^{S_1\times d}, b_1 \in \mathbb{R}^{S_1\times 1}, W_1^2 \in \mathbb{R}^{1\times q\times S_1}, b_2 \in \mathbb{R}^{1\times q}, S_1 \in \mathbb{N}\}$. Then $\Sigma(a)$ is dense in $W^{m,p}(U)$ for each $0 \le m \le \ell$, i.e., for any $\epsilon > 0$ and $f \in W^{m,p}(U)$ there exists a $\psi \in \Sigma(a)$ such that $\|\psi - f\|_{W^{m,p}} < \epsilon$.

In particular, we have that for any $\ell = 1$-finite activation $a, f \in H^1(U)$ and $\epsilon > 0$, there exists a $\psi \in \Sigma(a)$ such that

$$\int_U |\psi - f|^2 \, \mathrm{d}\lambda + \int_U \left|\nabla_x \psi - Df\right|^2 \mathrm{d}\lambda < \epsilon. \tag{5.11}$$

The main implication of the UAT property is that given a compact domain on $\mathbb{R}^d$ and an appropriate activation function, one can approximate any Sobolev function by shallow neural networks[3] with arbitrary accuracy. It is worth highlighting that in the context of a regression Monte Carlo application, this does not establish an implementable *regression bias* due to the lack of bounds on the width of the hidden layer. We remark that the above version is not a state of the art result and refer to Pinkus (1999) for a classical survey on the subject.

Layer Normalization. Normalization is a standard tool to enhance the convergence of stochastic gradient descent like algorithms (Goodfellow *et al.*, 2016). In standard examples (Han *et al.*, 2018), this is usually done by a so-called *batch normalization* technique. However, as we shall see, in our setting, batch normalization is computationally intensive as it ruins batch independence and implies quadratic dependence of the Jacobian tensor on the chosen batch size. Hence, we instead deploy *layer normalization* (Ba *et al.*, 2016), where normalization takes place across the output activations of a given hidden layer. Therefore, the final network architecture considered in Section 6 is described by the sequence of compositions

$$\Psi(x|\bar{\Theta}) := a^{\text{out}} \circ A^{L+1}(\cdot|\theta^{L+1}) \circ a \circ A^L(\cdot|\theta^L) \circ \bar{n} \circ a \circ \cdots \circ \bar{n} \circ A^1(\cdot|\theta^1) \circ x, \tag{5.12}$$

with $\bar{n}(\cdot|\beta_l)$ and $\bar{\Theta} := (\Theta, \beta_1, \ldots, \beta_{L-1})$, where β_l denotes the lth normalization layer's parameters—see Ba *et al.* (2016).

5.3 *A Deep BSDE approach*

In what follows, we formulate a Deep BSDE approach similar to Huré *et al.* (2020), which scales well in high-dimensional settings and tackles the fully-implementable scheme (5.1) in a neural network least-squares Monte Carlo framework. The main difference between our approach and that of Huré *et al.* (2020) is that, unlike in the discretization problem (3.4), we solve the d-dimensional linear BSDE of the Malliavin derivatives in (3.1d)—on top of the scalar BSDE (3.1b). We separate the solutions of these two BSDEs and perform two distinct neural network regressions at each time step. We distinguish between two approaches. The first involves an additional layer of parametrization in which the matrix-valued Γ process is approximated by an $\mathbb{R}^{d\times d}$-valued neural network. In the second, we take advantage of neural networks being dense function approximators in Sobolev spaces provided by Theorem 5.1, circumvent parametrizing the Γ process and instead obtain it as the direct derivative of the Z process via automatic differentiation—in a way very similar to the second scheme (DBDP2) of Huré *et al.* (2020). In doing so, we require a so-called Jacobian training where the loss is dependent of the derivative of the neural network involved.

[3] It is clear that the above statement generalizes to deep neural networks containing multiple hidden layers.

In order to motivate the merged problem formulation, notice that by Assumption 4.1 on the coefficients of the BSDE, the arguments of the conditional expectations in (5.1) are all $\mathbb{L}^2$-integrable random variables. Consequently, (5.1c), combined with the martingale representation theorem, implies the existence of a unique random process $D_n\widetilde{Z}_r$ such that

$$D_n\widehat{Y}^\pi_{n+1} + \Delta t_n f^D\big(t_{n+1}, \widehat{\mathbf{X}}^\pi_{n+1}, \mathbf{D}_n\breve{\mathbf{X}}^\pi_{n+1,n}\big) = \breve{Z}^\pi_n + \int_{t_n}^{t_{n+1}} \big(\big(D_n\widetilde{Z}_r\big)^T \mathrm{d}W_r\big)^T. \tag{5.13}$$

Itô's isometry implies that the $\mathbb{L}^2$-projection of $D_n\widetilde{Z}_r$ coincides with $D_n\breve{Z}^\pi_n$ in (5.1b)

$$D_n\breve{Z}^\pi_n = \frac{1}{\Delta t_n}\mathbb{E}_n\left[\int_{t_n}^{t_{n+1}} D_n\widetilde{Z}_r \,\mathrm{d}r\right]. \tag{5.14}$$

Thereupon, $\breve{Z}^\pi_n + ((D_n\breve{Z}^\pi_n)^T\Delta W_n)^T$ is, not just the best $\mathbb{L}^2$-projection of the left-hand side of (5.13), but also of the arguments of the conditional expectations on the right-hand side of (5.1b). Hence, it simultaneously solves the discretization problems (5.1b) and (5.1c).

Motivated by these observations the Deep BSDE approach then goes as follows—the complete algorithm is collected in Algorithm 1. We set $\widehat{Y}^\pi_N = g\left(X^\pi_N\right)$, $\widehat{Z}^\pi_N = \nabla_x g\left(X^\pi_N\right)\sigma(T, X^\pi_N)$ and $\widehat{\Gamma}^\pi_N = \nabla_x(\nabla_x g\sigma)(T, X^\pi_N)$. Thereafter, each time step's Y, Z and Γ are parametrized by three independent fully-connected feedforward neural networks $\varphi(\cdot|\theta^y) : \mathbb{R}^{d\times 1} \to \mathbb{R}$, $\psi(\cdot|\theta^z) : \mathbb{R}^{d\times 1} \to \mathbb{R}^{1\times d}$ and $\chi(\cdot|\theta^\gamma) : \mathbb{R}^{d\times 1} \to \mathbb{R}^{d\times d}$ of the type (5.12). The parameter sets $(\theta^z, \theta^\gamma)$ and θ^y are trained in two separate regressions. First, in light of (5.13), we define the loss function of the regression problem corresponding to (5.1b)–(5.1c) by

$$\begin{aligned}\mathcal{L}^{z,\gamma}_n(\theta^z, \theta^\gamma) := \mathbb{E}\Big[\Big|&\big(1 + \Delta t_n \nabla_y f\big(t_{n+1}, \widehat{\mathbf{X}}^\pi_{n+1}\big)\big)D_n\widehat{Y}^\pi_{n+1} + \Delta t_n \nabla_x f\big(t_{n+1}, \widehat{\mathbf{X}}^\pi_{n+1}\big)D_n X^\pi_{n+1}\\ &- \psi\left(X^\pi_n|\theta^z\right) + \Delta t_n \nabla_z f\big(t_{n+1}, \widehat{\mathbf{X}}^\pi_{n+1}\big)\chi\left(X^\pi_n|\theta^\gamma\right)\sigma\left(t_n, X^\pi_n\right)\\ &- \left(\left(\chi\left(X^\pi_n|\theta^\gamma\right)\sigma\left(t_n, X^\pi_n\right)\right)^T \Delta W_n\right)^T\Big|^2\Big],\end{aligned} \tag{5.15}$$

where we approximate $D_n Z^\pi_n$ by $\chi\left(X^\pi_n|\theta^\gamma\right) D_n X^\pi_n$, according to the Malliavin chain rule. We gather an approximation of the minimal parameter set $(\theta^{z,*}_n, \theta^{\gamma,*}_n) \in \arg\min_{(\theta^z,\theta^\gamma)} \mathcal{L}^z_n(\theta^z, \theta^\gamma)$ after minimizing an empirically observed version of the loss function through a stochastic gradient descent optimization, resulting in approximations $\widehat{\theta}^z_n$ and $\widehat{\theta}^\gamma_n$—see Algorithm 1. The final approximations are given by $\widehat{Z}^\pi_n := \psi(X^\pi_n|\widehat{\theta}^z_n)$ and $\widehat{\Gamma}^\pi_n := \chi(X^\pi_n|\widehat{\theta}^\gamma_n)$.

Similarly to the second scheme in Huré *et al.* (2020), an alternative formulation can be given, which avoids parametrizing the Γ process, and instead approximates it as the direct derivative of the Z process provided by the Malliavin chain rule lemma Lemma 2.1. Eventually, this implies the direct connection $\chi\left(X^\pi_n|\theta^\gamma\right) \equiv \nabla_x\psi\left(X^\pi_n|\theta^z\right)$, with which the corresponding loss function becomes

$$\begin{aligned}\mathcal{L}^{z,\nabla z}_n(\theta^z) := \mathbb{E}\Big[\Big|&\big(1 + \Delta t_n \nabla_y f\big(t_{n+1}, \widehat{\mathbf{X}}^\pi_{n+1}\big)\big)D_n\widehat{Y}^\pi_{n+1} + \Delta t_n \nabla_x f\big(t_{n+1}, \widehat{\mathbf{X}}^\pi_{n+1}\big)D_n X^\pi_{n+1}\\ &- \psi\left(X^\pi_n|\theta^z\right) + \Delta t_n \nabla_z f\big(t_{n+1}, \widehat{\mathbf{X}}^\pi_{n+1}\big)\nabla_x\psi\left(X^\pi_n|\theta^z\right)\sigma\left(t_n, X^\pi_n\right)\\ &- \left(\left(\nabla_x\psi\left(X^\pi_n|\theta^z\right)\sigma\left(t_n, X^\pi_n\right)\right)^T \Delta W_n\right)^T\Big|^2\Big],\end{aligned} \tag{5.16}$$

where we exploited the relation between the Γ and Z processes, provided by the Malliavin chain rule, and set $D_n\widehat{Z}_n^\pi = \nabla_x \widehat{z}_n^\pi \left(X_n^\pi\right) D_n X_n^\pi$. The SGD approximation of the optimal parameter space $\theta_n^{z,*} \in \arg\min_{\theta^z} \mathcal{L}_n^{z,\nabla z}(\theta^z)$ is denoted by $\widehat{\theta}_n^z$, and the final approximations are of the form $\widehat{Z}_n^\pi := \psi(X_n^\pi|\widehat{\theta}_n^z)$ and $\widehat{\Gamma}_n^\pi := \nabla_x \psi(X_n^\pi|\widehat{\theta}_n^z)$.

Subsequently, these approximations are plugged into the regression problem of (5.1d). This step, apart from the additional theta-discretization, is identical to that of Huré *et al.* (2020) and the loss function reads as

$$\mathcal{L}_n^y(\theta^y) := \mathbb{E}\Bigg[\Big|\widehat{Y}_{n+1}^\pi + (1-\vartheta_y)\Delta t_n f\Big(t_{n+1}, \widehat{\mathbf{X}}_{n+1}^\pi\Big) - \varphi\Big(X_n^\pi|\theta^y\Big) + \vartheta_y \Delta t_n f\big(t_n, X_n^\pi, \varphi\big(X_n^\pi|\theta^y\big), \widehat{Z}_n^\pi\big) - \widehat{Z}_n^\pi \Delta W_n\Big|^2\Bigg]. \tag{5.17}$$

The stochastic gradient descent approximation of the optimal parameter space $\theta_n^{y,*} \in \arg\min_{\theta^y} \mathcal{L}_n^y(\theta^y)$ is denoted by $\widehat{\theta}_n^y$ and the final approximation is given by $\widehat{Y}_n^\pi := \varphi(X_n^\pi|\widehat{\theta}_n^y)$. At last, motivated by the continuity of the processes $\{(Y_t, Z_t)\}_{0\le t\le T}$ in the Malliavin framework, we initialize the parameters of the next time step's parametrizations according to

$$\theta^z = \widehat{\theta}_n^z, \qquad \theta^\gamma = \widehat{\theta}_n^\gamma, \qquad \theta^y = \widehat{\theta}_n^y. \tag{5.18}$$

Such a *transfer learning* trick guarantees a good initialization of the SGD iterations for $\check{Y}_{n-1}^\pi, \check{Z}_{n-1}^\pi, \check{\Gamma}_{n-1}^\pi$, simplifying the learning problem and reducing the number of iteration steps required for convergence. For an empirical assessment on the efficiency of this transfer learning trick, we refer to Chen & Wan (2021, Sec.5.3).

Dimensionality, linearity and vector-Jacobian products. The main reason why no numerical scheme has been proposed to solve the Malliavin BSDE in (3.1d) is related to dimensionality. Since the Γ process is an $\mathbb{R}^{d\times d}$-valued process, its computational complexity in a least-squares Monte Carlo method has a quadratic dependence on the number of dimensions d. Indeed, a least-squares Monte Carlo approach for the BSDE (1.1b) essentially comes down to the approximation of $d+1$-many conditional expectations. If, in addition, one would also like to solve the Malliavin BSDE (3.1d) this leads to d^2 additional conditional expectations to be approximated, induced by the Γ process. This observation justifies the use of deep neural network parametrizations, which enable good scalability in high-dimensions. Moreover, notice that the training of the loss function (5.16) through an SGD optimization requires differentiating the loss with respect to the parameters θ^z in each step. With the loss already depending on the Jacobian of the mapping $\psi(\cdot|\theta^z)$, this in particular implies that in each SGD step one needs to calculate the Hessian of a vector-valued mapping ψ with respect to the parameters θ^z. Consequently, for high-dimensional problems, the training of (5.16) becomes excessively intensive from a computational point of view. Nonetheless, what makes the Deep BSDE approach corresponding to (5.16) efficiently implementable is the linearity of the Malliavin BSDE (3.1d). In fact, due to linearity, one can circumvent explicitly calculating the Jacobian matrix of Z as it suffices to calculate the vector-Jacobian product

$$\nabla_z f\big(t_{n+1}, \widehat{\mathbf{X}}_{n+1}^\pi\big)\nabla_x \psi\left(X_n^\pi|\theta^z\right) = \nabla_x\left\langle v|\psi\left(X_n^\pi|\theta^z\right)\right\rangle, \qquad v := \nabla_z f\big(t_{n+1}, \widehat{\mathbf{X}}_{n+1}^\pi\big), \tag{5.21}$$

which boils down to computing a gradient instead. This mitigates the computational costs of minimizing the automatic differentiated loss function in (5.16) in an SGD iteration.

Algorithm 1: One-Step Malliavin Algorithm (OSM)

Input: $\pi(N)$, $\vartheta_y \in [0,1]$ – discretization parameters
Input: $B \in \mathbb{N}^+$, $I \in \mathbb{N}$, $\eta : \mathbb{N} \to \mathbb{R}$ – training parameters
Result: $\{(\widehat{Y}_n^\pi, \widehat{Z}_n^\pi, \widehat{\Gamma}_n^\pi)\}_{n=0,\dots,N}$ – discrete time approximations over π
$\widehat{Y}_N^\pi \leftarrow g(X_N^\pi)$, $\widehat{Z}_N^\pi \leftarrow \nabla_x g(X_N^\pi)\sigma(t_N, X_N^\pi)$, $\widehat{\Gamma}_N^\pi \leftarrow \nabla_x(\nabla_x g\sigma)(t_N, X_N^\pi)$ – collect terminal condition$\varphi(\cdot|\theta^y) : \mathbb{R}^{d\times 1} \to \mathbb{R}$, $\psi(\cdot|\theta^z) : \mathbb{R}^{d\times 1} \to \mathbb{R}^{1\times d}$, $\chi(\cdot|\theta^\gamma) : \mathbb{R}^{d\times 1} \to \mathbb{R}^{d\times d}$ – neural network parametrizations**for** $n = N-1, \dots, 0$ **do**
  if $n = N-1$ **then**
    $\theta^{z,(0)}, \theta^{y,(0)}$ – initialize parameter sets, according to [19]
  else
    $\theta^{z,(0)} \leftarrow \widehat{\theta}_{n+1}^z$, $\theta^{y,(0)} \leftarrow \widehat{\theta}_{n+1}^y$ – transfer learning initialization
  end
  Solve Equation 5.1c–Equation 5.1b.
  for $i = 0, \dots, I-1$ **do**
    $\{\{X_m^\pi(b)\}_{0\le m\le N}\}_{b=1}^B$ – Euler-Maruyama simulations by Equation 3.2$\{D_n X_{n+1}^\pi(b)\}_{b=1}^B$ – Euler-Maruyama approximations by Equation 3.5calculate empirical loss of Equation 5.15 or Equation 5.16

$$\begin{aligned}\widehat{\mathcal{L}}_n^{z,\gamma}(\theta^{z,(i)}, \theta^{\gamma,(i)}) = \frac{1}{B}\sum_{b=1}^B |&(1 + \Delta t_n \nabla_y f(t_{n+1}, \widehat{\mathbf{X}}_{n+1}^\pi(b)))D_n\widehat{Y}_{n+1}^\pi(b) \\ &+ \Delta t_n \nabla_x f(t_{n+1}, \widehat{\mathbf{X}}_{n+1}^\pi(b))D_n X_{n+1}^\pi(b) - \psi(X_n^\pi(b)|\theta^{z,(i)}) \\ &+ \Delta t_n \nabla_z f(t_{n+1}, \widehat{\mathbf{X}}_{n+1}^\pi(b))\chi(X_n^\pi(b)|\theta^{\gamma,(i)})\sigma(t_n, X_n^\pi) \\ &- ((\chi(X_n^\pi(b)|\theta^{\gamma,(i)})\sigma(t_n, X_n^\pi(b)))^T \Delta W_n(b))^T|^2\end{aligned} \tag{5.19}$$

    $(\theta^{z,(i+1)}, \theta^{\gamma,(i+1)}) \leftarrow (\theta^{z,(i)}, \theta^{\gamma,(i)}) - \eta(i)\nabla_{(\theta^z,\theta^\gamma)}\widehat{\mathcal{L}}_n^z(\theta^{z,(i)}, \theta^{\gamma,(i)})$ – stochastic gradient descent update
  end
  $\widehat{\theta}_n^z \leftarrow \theta^{z,(I)}$, $\widehat{\theta}_n^\gamma \leftarrow \theta^{\gamma,(I)}$ – collect optimal parameter estimations$\widehat{z}_n^\pi(\cdot) \leftarrow \psi(\cdot|\widehat{\theta}_n^z)$, $\widehat{\gamma}_n^\pi(\cdot) \leftarrow \chi(\cdot|\widehat{\theta}_n^\gamma)$ – collect approximations $\widehat{Z}_n^\pi, \widehat{\Gamma}_n^\pi$
  Solve Equation 5.1d.
  for $i = 0, \dots, I-1$ **do**
    $\{\{X_m^\pi(b)\}_{0\le m\le N}\}_{b=1}^B$ – Euler-Maruyama simulations by Equation 3.2calculate empirical loss of Equation 5.17

$$\begin{aligned}\widehat{\mathcal{L}}_n^y(\theta^{y,(i)}) = \frac{1}{B}\sum_{b=1}^B |&\widehat{Y}_{n+1}^\pi(b) + (1-\vartheta_y)\Delta t_n f(t_{n+1}, \widehat{\mathbf{X}}_{n+1}^\pi(b)) - \varphi(X_n^\pi(b)|\theta^{y,(i)}) \\ &+ \vartheta_y \Delta t_n f(t_n, X_n^\pi(b), \varphi(X_n^\pi(b)|\theta^{y,(i)}), \widehat{Z}_n^\pi(b)) - \widehat{Z}_n^\pi(b)\Delta W_n(b)|^2\end{aligned} \tag{5.20}$$

    $\theta^{y,(i+1)} \leftarrow \theta^{y,(i)} - \eta(i)\nabla_\theta \widehat{\mathcal{L}}_n^y(\theta^{y,(i)})$ – stochastic gradient descent step
  end
  $\widehat{\theta}_n^y \leftarrow \theta^{y,(I)}$ – collect optimal parameter estimations$\widehat{y}_n^\pi(\cdot) \leftarrow \varphi(\cdot|\widehat{\theta}_n^y)$ – collect approximations $\widehat{Y}_n^\pi$
end

5.4 *Regression error analysis*

In order to conclude the discussion on fully-implementable schemes for (5.1), we extend the discretization error results established by Theorem 4.3, so that it incorporates the approximation errors of the arising conditional expectations. Even though we focus on the Deep BSDE approach, our arguments naturally extend to the BCOS estimates. We consider shallow neural networks, with S_1-many hidden neurons and a hyperbolic tangent activation. While distinguishing between the parametrized and automatic differentiated Γ variants—see Equation 5.15 and (5.16), respectively—we rely on the following subclass of shallow neural networks introduced in Theorem 5.1

$$\Sigma_{C_b^2}(\tanh) := \left\{ \psi(x|\theta^z(S_1)) \in \Sigma(\tanh) : \sum_{i=1}^{d} \sum_{j=1}^{S_1} \left| \left[W_1^2(S_1) \right]_{1,i,j} \right| + \left| \left[W_0^1(S_1) \right]_{j,i} \right| \leq \Upsilon(S_1) \right\}, \quad (5.22)$$

for some dominating sequence $\Upsilon : \mathbb{N}_+ \to \mathbb{R}$. Then, due to the boundedness of the hyperbolic tangent function and its first two derivatives, the following upper bounds are in place for any $\psi(\cdot|\theta^z) \in \Sigma_{C_b^2}(\tanh)$

$$\sup_{x \in \mathbb{R}^{d\times 1}} \left| \psi(x|\theta^z) \right| \leq \Upsilon(S_1), \qquad \sup_{x \in \mathbb{R}^{d\times 1}} \left| \nabla_x \psi(x|\theta^z) \right| \leq \Upsilon^2(S_1), \qquad \sup_{x \in \mathbb{R}^{d\times 1}} \left| \mathrm{Hess}_x \psi(x|\theta^z) \right| \leq \Upsilon^3(S_1). \quad (5.23)$$

In light of Theorem 5.1, the hyperbolic tangent function is $\ell = 1$-finite. Subsequently, the family of shallow networks of the form (5.12) is dense in $H^1(U)$ for any compact subset $U \subset \mathbb{R}^{d\times 1}$.

The final approximations are denoted by $\widehat{Y}_n^\pi := \widehat{\mathcal{Y}}_n^\pi \left(X_n^\pi \right) =: \varphi(X_n^\pi|\widehat{\theta}_n^y), \widehat{Z}_n^\pi := \widehat{\mathcal{Z}}_n^\pi \left(X_n^\pi \right) =: \psi(X_n^\pi|\widehat{\theta}_n^z)$ and $\widehat{\Gamma}_n^\pi := \widehat{\mathcal{Y}}_n^\pi \left(X_n^\pi \right) =: \chi(X_n^\pi|\widehat{\theta}_n^\gamma)$. We introduce the notations $\Delta \breve{Y}_n^\pi := Y_{t_n} - \breve{Y}_n^\pi$, $\Delta \breve{Z}_n^\pi := Z_{t_n} - \breve{Z}_n^\pi$, $\Delta \breve{\Gamma}_n^\pi := \Gamma_{t_n} - \breve{\Gamma}_n^\pi$ and $\Delta \widehat{Y}_n^\pi := Y_{t_n} - \widehat{Y}_n^\pi$, $\Delta \widehat{Z}_n^\pi := Z_{t_n} - \widehat{Z}_n^\pi$, $\Delta \widehat{\Gamma}_n^\pi := \Gamma_{t_n} - \widehat{\Gamma}_n^\pi$. In light of the UAT property in Theorem 5.1, we define the *regression biases*

$$\begin{aligned}
\epsilon_n^y &:= \inf_{\theta^y} \mathbb{E}\left[\left| \breve{\mathcal{Y}}_n^\pi(X_n^\pi) - \varphi\left(X_n^\pi|\theta^y\right) \right|^2 \right], \\
\epsilon_n^z &:= \inf_{\theta^z} \mathbb{E}\left[\left| \breve{\mathcal{Z}}_n^\pi\left(X_n^\pi\right) - \psi\left(X_n^\pi|\theta^z\right) \right|^2 \right], \qquad \epsilon_n^\gamma := \inf_{\theta^\gamma} \mathbb{E}\left[\left| \left(\breve{\mathcal{Y}}_n^\pi\left(X_n^\pi\right) - \chi\left(X_n^\pi|\theta^\gamma\right) \right) \sigma\left(t_n, X_n^\pi\right) \right|^2 \right], \\
\epsilon_n^{z,\nabla z} &:= \inf_{\theta^z} \mathbb{E}\left[\left| \breve{\mathcal{Z}}_n^\pi\left(X_n^\pi\right) - \psi\left(X_n^\pi|\theta^z\right) \right|^2 + \Delta t_n \left| \left(\nabla_x \breve{\mathcal{Z}}_n^\pi\left(X_n^\pi\right) - \nabla_x \psi\left(X_n^\pi|\theta^z\right) \right) \sigma\left(t_n, X_n^\pi\right) \right|^2 \right].
\end{aligned} \quad (5.24)$$

The goal is to establish an upper bound for the total approximation error defined by

$$\widehat{\mathcal{E}}^\pi(|\pi|) := \max_{0\leq n\leq N} \mathbb{E}\left[\left| \Delta \widehat{Y}_n^\pi \right|^2 \right] + \max_{0\leq n\leq N} \mathbb{E}\left[\left| \Delta \widehat{Z}_n^\pi \right|^2 \right] + \mathbb{E}\left[\sum_{n=0}^{N-1} \int_{t_n}^{t_{n+1}} \left| \Gamma_r - \widehat{\Gamma}_n^\pi \right|^2 \mathrm{d}r \right], \quad (5.25)$$

depending on, not just the discretization, but also the regression errors arising from the approximations of the conditional expectations in (5.1).

THEOREM 5.2 Let the conditions of Assumption 4.1 be in place. Assume the time partition satisfies $N\Delta t_n \geq c$ for each $0 \leq n \leq N-1$, with some constant c. Then, for sufficiently small $|\pi|$, the total approximation error of the OSM scheme defined by the loss function Equation 5.15 admits to

$$\widehat{\mathcal{E}}^{\pi}(|\pi|) \leq C\left(|\pi| + N\sum_{n=0}^{N-1}\left\{\epsilon_n^y + \epsilon_n^z\right\} + \sum_{n=0}^{N-1}\epsilon_n^{\gamma}\right). \tag{5.26}$$

Furthermore, in case the Γ process is taken as the direct derivative of the Z process, as in (5.16), the total error can be bounded by

$$\widehat{\mathcal{E}}^{\pi}(|\pi|) \leq C\left(|\pi| + N\sum_{n=0}^{N-1}\{\epsilon_n^y + \epsilon_n^{z,\nabla z}\} + \frac{\Upsilon^6(S_1)}{N}\right), \tag{5.27}$$

where C is a constant independent of the time partition π^N.

Proof. Throughout the proof, C denotes a constant independent of the time partition, whose value may vary from line to line. We only highlight arguments that significantly differ from the ones of Theorem 4.3.

Step 1. *Regression errors induced by the loss functions.* Using (5.13), the relation (5.14) and the total law of probability, the loss function in Equation 5.15 can be rewritten as follows:

$$\begin{aligned}
\mathcal{L}_n^{z,\gamma}(\theta^z,\theta^\gamma) &= \mathbb{E}\left[\left|\check{Z}_n^{\pi} - \psi\left(X_n^{\pi}|\theta^z\right) + \Delta t_n \nabla_z f\left(t_{n+1}, \widehat{\mathbf{X}}_{n+1}^{\pi}\right)\left(\chi\left(X_n^{\pi}|\theta^\gamma\right) - \check{\Gamma}_n^{\pi}\right)\sigma\left(t_n, X_n^{\pi}\right)\right|^2\right] \\
&\quad + \Delta t_n \mathbb{E}\left[\left|\left(\check{\Gamma}_n^{\pi} - \chi\left(X_n^{\pi}|\theta^\gamma\right)\right)\sigma\left(t_n, X_n^{\pi}\right)\right|^2\right] + \mathbb{E}\left[\int_{t_n}^{t_{n+1}}\left|D_n\widetilde{Z}_r - D_n\check{Z}_n^{\pi}\right|^2 \mathrm{d}r\right] \\
&\quad + 2\Delta t_n \mathbb{E}\Bigg[\nabla_z f(t_{n+1}, \widehat{\mathbf{X}}_{n+1}^{\pi})\left(\chi\left(X_n^{\pi}|\theta^\gamma\right) - \check{\Gamma}_n^{\pi}\right)\sigma\left(t_n, X_n^{\pi}\right) \\
&\qquad \times \int_{t_n}^{t_{n+1}}\left(D_n\widetilde{Z}_r - \chi\left(X_n^{\pi}|\theta^\gamma\right)\sigma\left(t_n, X_n^{\pi}\right)\right)^T \mathrm{d}W_r\Bigg] \\
&=: \widetilde{\mathcal{L}}_n^{z,\gamma}(\theta^z,\theta^\gamma) + \mathbb{E}\left[\int_{t_n}^{t_{n+1}}\left|D_n\widetilde{Z}_r - D_n\check{Z}_n^{\pi}\right|^2 \mathrm{d}r\right] + \widetilde{I}_n^{\gamma}(\theta^\gamma).
\end{aligned} \tag{5.28}$$

The inequality $(a+b)^2 \leq (1+\varrho_1\Delta t_n)a^2 + (1+1/(\varrho_1\Delta t_n))b^2$, on top of the bounded differentiability of f provided by Assumption 4.1, implies

$$\begin{aligned}
\widetilde{\mathcal{L}}_n^{z,\gamma}(\theta^z,\theta^\gamma) &\leq (1+\varrho_1\Delta t_n)\mathbb{E}\left[\left|\check{Z}_n^{\pi} - \psi\left(X_n^{\pi}|\theta^z\right)\right|^2\right] \\
&\quad + \left[\frac{L_{\nabla f}^2}{\rho_1}(1+\varrho_1\Delta t_n) + 1\right]\Delta t_n \mathbb{E}\left[\left|\left(\check{\Gamma}_n^{\pi} - \chi\left(X_n^{\pi}|\theta^\gamma\right)\right)\sigma\left(t_n, X_n^{\pi}\right)\right|^2\right].
\end{aligned} \tag{5.29}$$

By the inequality $(a+b)^2 \geq (1-\varrho_2\Delta t_n)a^2 + (1-1/(\varrho_2\Delta t_n))b^2$, the following also holds:

$$\widetilde{\mathcal{L}}_n^{z,\gamma}(\theta^z,\theta^\gamma) \geq (1-\varrho_2\Delta t_n)\mathbb{E}\left[\left|\breve{Z}_n^\pi - \psi\left(X_n^\pi|\theta^z\right)\right|^2\right] + \left(1+\frac{L_{\nabla f}^2}{\varrho_2}(\varrho_2\Delta t_n - 1)\right)\Delta t_n\mathbb{E}\left[\left|\left(\breve{\Gamma}_n^\pi - \chi\left(X_n^\pi|\theta^\gamma\right)\right)\sigma\left(t_n, X_n^\pi\right)\right|^2\right]. \tag{5.30}$$

The Cauchy–Schwarz inequality, (5.14) and the ε-Young inequality $ab \leq a^2/(2\varepsilon) + \varepsilon b^2/2$, with $\varepsilon = 4L_{\nabla f}^2$, yield

$$\left|\widetilde{I}_n^\gamma(\theta^\gamma)\right| \leq \left(1/4 + 4L_{\nabla f}^2\Delta t_n\right)\Delta t_n\mathbb{E}\left[\left|\left(\breve{\Gamma}_n^\pi - \chi\left(X_n^\pi|\theta^\gamma\right)\right)\sigma\left(t_n, X_n^\pi\right)\right|^2\right] + 4L_{\nabla f}^2\Delta t_n\mathbb{E}\left[\int_{t_n}^{t_{n+1}}\left|D_n\widetilde{Z}_r - D_n\breve{Z}_n^\pi\right|^2 \mathrm{d}r\right]. \tag{5.31}$$

Implied by (5.28), minimizing $\mathcal{L}_n^{z,\gamma}(\theta^z,\theta^\gamma)$ is equivalent to minimizing $\bar{\mathcal{L}}_n^{z,\gamma} := \widetilde{\mathcal{L}}_n^{z,\gamma}(\theta^z,\theta^\gamma) + \widetilde{I}_n^\gamma(\theta^\gamma)$. Assuming that $(\widehat{\theta}_n^z, \widehat{\theta}_n^\gamma)$ is a perfect approximation—see Remark 5.3—of the minimal parameter space $(\theta_n^{z,*}, \theta_n^{\gamma,*}) \in \arg\min_{\theta^z,\theta^\gamma}\mathcal{L}_n^{z,\gamma}(\theta^z,\theta^\gamma)$, we have $\bar{\mathcal{L}}_n^{z,\gamma}(\widehat{\theta}_n^z, \widehat{\theta}_n^\gamma) \leq \bar{\mathcal{L}}_n^{z,\gamma}(\theta^z,\theta^\gamma)$ for any (θ^z,θ^γ). Combined with (5.29), (5.30), the triangle inequality and (5.31), this implies

$$\begin{aligned}(1-\varrho_2\Delta t_n)\mathbb{E}\left[\left|\breve{Z}_n^\pi - \widehat{Z}_n^\pi\right|^2\right] + \left(3/4 - L_{\nabla f}^2/\varrho_2 - 3L_{\nabla f}^2\Delta t_n\right)\Delta t_n\mathbb{E}\left[\left|\left(\breve{\Gamma}_n^\pi - \widehat{\Gamma}_n^\pi\right)\sigma\left(t_n, X_n^\pi\right)\right|^2\right] \\ \leq (1+\varrho_1\Delta t_n)\mathbb{E}\left[\left|\breve{Z}_n^\pi - \psi\left(X_n^\pi|\theta^z\right)\right|^2\right] \\ + \left[\frac{L_{\nabla f}^2}{\varrho_1}(1+\varrho_1\Delta t_n) + 5/4 + 4L_{\nabla f}^2\Delta t_n\right]\Delta t_n\mathbb{E}\left[\left|\left(\breve{\Gamma}_n^\pi - \chi\left(X_n^\pi|\theta^\gamma\right)\right)\sigma\left(t_n, X_n^\pi\right)\right|^2\right] \\ + 8L_{\nabla f}^2\Delta t_n\mathbb{E}\left[\int_{t_n}^{t_{n+1}}\left|D_n\widetilde{Z}_r - D_n\breve{Z}_n^\pi\right|^2 \mathrm{d}r\right],\end{aligned} \tag{5.32}$$

for any (θ^z,θ^γ), $\varrho_1,\varrho_2 > 0$. In particular, choosing $\varrho_2^* := 8L_{\nabla f}^2$, for any sufficiently small Δt_n such that $3L_{\nabla f}^2\Delta t_n < 1/8$ and $\varrho_2^*\Delta t_n \leq 1/2$, we derive

$$\mathbb{E}\left[\left|\breve{Z}_n^\pi - \widehat{Z}_n^\pi\right|^2\right] + \Delta t_n\mathbb{E}\left[\left|\left(\breve{\Gamma}_n^\pi - \widehat{\Gamma}_n^\pi\right)\sigma\left(t_n, X_n^\pi\right)\right|^2\right] \leq C\left(\epsilon_n^z + \Delta t_n\epsilon_n^\gamma\right) + 16L_{\nabla f}^2\Delta t_n\mathbb{E}\left[\int_{t_n}^{t_{n+1}}\left|D_n\widetilde{Z}_r - D_n\breve{Z}_n^\pi\right|^2 \mathrm{d}r\right], \tag{5.33}$$

recalling the definitions in (5.24). Through analogous steps to Huré *et al.* (2020, Thm. 4.1, steps 3–4), a similar estimate can be established for the loss function (5.17), ultimately giving

$$\mathbb{E}\left[\left|\check{Y}_n^\pi - \widehat{Y}_n^\pi\right|^2\right] \leq C \inf_{\theta^y} \mathbb{E}\left[\left|\check{Y}_n^\pi - \varphi\left(X_n^\pi | \theta^y\right)\right|^2\right] =: C\epsilon_n^y. \tag{5.34}$$

Step 2. $\mathbb{L}^2$*-regularity of* $D_n\widetilde{Z}_r$. In what follows, we will need an estimate controlling the so-called $\mathbb{L}^2$-regularity of the stochastic integrand $D_n\widetilde{Z}_r$, corresponding to the last term in (5.33). This term admits to the following bound:

$$\begin{aligned}\mathbb{E}\left[\int_{t_n}^{t_{n+1}} \left|D_n\widetilde{Z}_r - D_n\check{Z}_n^\pi\right|^2 \mathrm{d}r\right] &\leq 3\mathbb{E}\left[\int_{t_n}^{t_{n+1}} \left|D_n\widetilde{Z}_r - D_{t_n}Z_r\right|^2 \mathrm{d}r\right] + 3\mathbb{E}\left[\int_{t_n}^{t_{n+1}} \left|D_{t_n}Z_r - \overline{DZ}_n^{n+1}\right|^2 \mathrm{d}r\right] \\ &\quad + 3\Delta t_n \mathbb{E}\left[\left|\overline{DZ}_n^{n+1} - D_n\check{Z}_n^\pi\right|^2\right] =: 3(R_1 + R_2 + R_3).\end{aligned} \tag{5.35}$$

The second term of the right-hand side corresponds to the $\mathbb{L}^2$-regularity of DZ given by (4.6). For the first term, notice that by Itô's isometry, (5.13) and (3.6), we have

$$R_1 = \mathbb{E}\left[\left|\Delta\check{Z}_n^\pi - \Delta D_n\widehat{Y}_{n+1}^\pi + \int_{t_n}^{t_{n+1}} f^D\left(t_{n+1}, \widehat{\mathbf{X}}_{n+1}^\pi, \mathbf{D}_n\check{\mathbf{X}}_{n+1,n}^\pi\right) - f^D\left(r, \mathbf{X}_r, \mathbf{D}_{t_n}\mathbf{X}_r\right) \mathrm{d}r\right|^2\right]. \tag{5.36}$$

(5.1c) implies an identity similar to (4.19). Then, by the law of total probability, the $L^2([0,T];\mathbb{R}^d)$ Cauchy–Schwarz and Jensen inequalities, it follows that

$$\begin{aligned}R_1 &\leq 2\mathbb{E}\left[\left|\Delta D_n\widehat{Y}_{n+1}^\pi\right|^2 - \left|\mathbb{E}_n\left[\Delta D_n\widehat{Y}_{n+1}^\pi\right]\right|^2\right] \\ &\quad + 4\Delta t_n \mathbb{E}\left[\int_{t_n}^{t_{n+1}} \left|f^D\left(t_{n+1}, \widehat{\mathbf{X}}_{n+1}^\pi, \mathbf{D}_n\check{\mathbf{X}}_{n+1,n}^\pi\right) - f^D\left(r, \mathbf{X}_r, \mathbf{D}_{t_n}\mathbf{X}_r\right)\right|^2 \mathrm{d}r\right].\end{aligned} \tag{5.37}$$

Notice that the second term above implicitly depends on R_3. Similarly to step 1 in Theorem 4.3—see (4.15) in particular—we also gather the following estimate:

$$\begin{aligned}R_3 := \Delta t_n \mathbb{E}\left[\left|\overline{DZ}_n^{n+1} - D_n\check{Z}_n^\pi\right|^2\right] &\leq 4d\left(\mathbb{E}\left[\left|\Delta D_n\widehat{Y}_{n+1}^\pi\right|^2\right] - \mathbb{E}\left[\left|\mathbb{E}_n\left[\Delta D_n\widehat{Y}_{n+1}^\pi\right]\right|^2\right]\right) \\ &\quad + 28dL_{f^D}^2\Delta t_n\Bigg\{C\Delta t_n^2 + 2\Delta t_n\mathbb{E}\left[\left|\Delta X_{n+1}^\pi\right|^2\right] \\ &\qquad + 2\Delta t_n\left(\mathbb{E}\left[\left|\Delta\widehat{Y}_{n+1}^\pi\right|^2\right] + \mathbb{E}\left[\left|\Delta\widehat{Z}_{n+1}^\pi\right|^2\right]\right) \\ &\qquad + 2\Delta t_n\left(\mathbb{E}\left[\left|\Delta D_n X_{n+1}^\pi\right|^2\right] + \mathbb{E}\left[\left|\Delta D_n\widehat{Y}_{n+1}^\pi\right|^2\right]\right) \\ &\qquad + \mathbb{E}\left[\int_{t_n}^{t_{n+1}} \left|D_{t_n}Z_r - \overline{DZ}_n^{n+1}\right|^2 \mathrm{d}r\right]\Bigg\},\end{aligned} \tag{5.38}$$

for any sufficiently small time step satisfying $14dL_{f_D}^2\Delta t_n \leq 1/2$. Plugging the combined estimate resulting from (5.37) and (5.38) into (5.35) subsequently gives

$$\begin{aligned}\mathbb{E}\left[\int_{t_n}^{t_{n+1}}\left|D_n\widetilde{Z}_r - D_n\breve{Z}_n^\pi\right|^2\mathrm{d}r\right] &\leq 3(2+8d)\big(\mathbb{E}\big[\big|\Delta D_n\widehat{Y}_{n+1}^\pi\big|^2\big] - \mathbb{E}\big[\big|\mathbb{E}_n\left[\Delta D_n\widehat{Y}_{n+1}^\pi\right]\big|^2\big]\big)\\ &\quad + 84L_{f_D}^2(1+2d)\Delta t_n\Big\{C\Delta t_n^2 + 2\Delta t_n\mathbb{E}\big[\big|\Delta X_{n+1}^\pi\big|^2\big]\\ &\quad + 2\Delta t_n\big(\mathbb{E}\big[\big|\Delta\widehat{Y}_{n+1}^\pi\big|^2\big] + \mathbb{E}\big[\big|\Delta\widehat{Z}_{n+1}^\pi\big|^2\big]\big)\\ &\quad + 2\Delta t_n\mathbb{E}\big[\big|\Delta D_n X_{n+1}^\pi\big|^2\big]\\ &\quad + 2\Delta t_n\mathbb{E}\big[\big|\Delta D_n\widehat{Y}_{n+1}^\pi\big|^2\big]\Big\}\\ &\quad + C\mathbb{E}\left[\int_{t_n}^{t_{n+1}}\left|D_{t_n}Z_r - \overline{DZ}_n^{n+1}\right|^2\mathrm{d}r\right],\end{aligned} \tag{5.39}$$

establishing an upper bound for the $\mathbb{L}^2$-regularity of $D_n\widetilde{Z}_r$.

Step 3. *Approximation error bound in the parametrized case.* The total approximation errors can be decomposed into discretization and regression errors as follows:

$$\Delta t_n\mathbb{E}\left[\left|\overline{DZ}_n^{n+1} - D_n\widehat{Z}_n^\pi\right|^2\right] \leq 2\Delta t_n\mathbb{E}\left[\left|\overline{DZ}_n^{n+1} - D_n\breve{Z}_n^\pi\right|^2\right] + 2\Delta t_n\mathbb{E}\left[\left|D_n\breve{Z}_n^\pi - D_n\widehat{Z}_n^\pi\right|^2\right], \tag{5.40}$$

$$(1-\beta_z\Delta t_n)\mathbb{E}\left[\left|\Delta\widehat{Z}_n^\pi\right|^2\right] \leq \mathbb{E}\left[\left|\Delta\breve{Z}_n^\pi\right|^2\right] + \frac{1}{\beta_z\Delta t_n}\mathbb{E}\left[\left|\breve{Z}_n^\pi - \widehat{Z}_n^\pi\right|^2\right], \tag{5.41}$$

$$(1-\beta_y\Delta t_n)\mathbb{E}\left[\left|\Delta\widehat{Y}_n^\pi\right|^2\right] \leq \mathbb{E}\left[\left|\Delta\breve{Y}_n^\pi\right|^2\right] + \frac{1}{\beta_y\Delta t_n}\mathbb{E}\left[\left|\breve{Y}_n^\pi - \widehat{Y}_n^\pi\right|^2\right], \tag{5.42}$$

for any $\beta_z, \beta_y > 0$. Combined with (5.33), (5.41) leads to the following estimate:

$$(1-\beta_z\Delta t_n)\mathbb{E}\left[\left|\Delta\widehat{Z}_n^\pi\right|^2\right] \leq \mathbb{E}\left[\left|\Delta\breve{Z}_n^\pi\right|^2\right] + \frac{C}{\beta_z\Delta t_n}\left(\epsilon_n^z + \Delta t_n\epsilon_n^\gamma\right) + \frac{16L_{\nabla f}^2}{\beta_z}\mathbb{E}\left[\int_{t_n}^{t_{n+1}}\left|D_n\widetilde{Z}_r - D_n\breve{Z}_n^\pi\right|^2\mathrm{d}r\right], \tag{5.43}$$

for any $\beta_z > 0$. Similar arguments as in step 2 in Theorem 4.3 subsequently give

$$\begin{aligned}&(1-\beta_z\Delta t_n)\mathbb{E}\left[\left|\Delta\widehat{Z}_n^\pi\right|^2\right]\\ &\quad\leq (1+\varrho\Delta t_n)\mathbb{E}\left[\left|\mathbb{E}_n\left[\Delta D_n\widehat{Y}_{n+1}^\pi\right]\right|^2\right]\\ &\qquad + \frac{7L_{f_D}^2}{\varrho}(1+\varrho\Delta t_n)\Big\{C\Delta t_n^2 + 2\Delta t_n\left(\mathbb{E}\left[\left|\Delta X_{n+1}^\pi\right|^2 + \left|\Delta\widehat{Y}_{n+1}^\pi\right|^2 + \left|\Delta\widehat{Z}_{n+1}^\pi\right|^2\right]\right)\end{aligned}$$

$$+ 2\Delta t_n \left(\mathbb{E}\left[\left| \Delta D_n X_{n+1}^{\pi} \right|^2 + \left| \Delta D_n \widehat{Y}_{n+1}^{\pi} \right|^2 \right] \right)$$
$$+ \Delta t_n \mathbb{E}\left[\left| D_n \breve{Z}_n^{\pi} - \overline{DZ}_n^{n+1} \right|^2 \right]$$
$$+ \mathbb{E}\left[\int_{t_n}^{t_{n+1}} \left| D_{t_n} Z_r - \overline{DZ}_n^{n+1} \right|^2 \mathrm{d}r \right] \Bigg\}$$
$$+ \frac{C}{\beta_z \Delta t_n} \left(\epsilon_n^z + \Delta t_n \epsilon_n^{\gamma} \right) + \frac{16 L_{\nabla f}^2}{\beta_z} \mathbb{E}\left[\int_{t_n}^{t_{n+1}} \left| D_n \widetilde{Z}_r - D_n \breve{Z}_n^{\pi} \right|^2 \mathrm{d}r \right], \tag{5.44}$$

for any $\varrho > 0$. Plugging in the estimates established by (5.38) and (5.39), choosing $\varrho^* = 56 d L_{fD}^2$ and $\beta_z^* = 96(2 + 8d) L_{\nabla f}^2$, we derive

$$(1 - \beta_z^* \Delta t_n) \mathbb{E}\left[\left| \Delta \widehat{Z}_n^{\pi} \right|^2 \right] \leq (1 + C_z \Delta t_n) \mathbb{E}\left[\left| \Delta D_n \widehat{Y}_{n+1}^{\pi} \right|^2 \right]$$
$$+ C_z \Bigg\{ C \Delta t_n^2 + 2\Delta t_n \left(\mathbb{E}\left[\left| \Delta X_{n+1}^{\pi} \right|^2 + \left| \Delta \widehat{Y}_{n+1}^{\pi} \right|^2 + \left| \Delta \widehat{Z}_{n+1}^{\pi} \right|^2 \right] \right)$$
$$+ 2\Delta t_n \left(\mathbb{E}\left[\left| \Delta D_n X_{n+1}^{\pi} \right|^2 + \left| \Delta D_n \widehat{Y}_{n+1}^{\pi} \right|^2 \right] \right)$$
$$+ \mathbb{E}\left[\int_{t_n}^{t_{n+1}} \left| D_{t_n} Z_r - \overline{DZ}_n^{n+1} \right|^2 \mathrm{d}r \right]$$
$$+ \frac{\epsilon_n^z + \Delta t_n \epsilon_n^{\gamma}}{\beta_z^* \Delta t_n} \Bigg\}. \tag{5.45}$$

By analogous computations for Y, similar arguments as in Theorem 4.3 imply

$$(1 - \beta^* \Delta t_n) \left(\mathbb{E}\left[\left| \Delta \widehat{Y}_n^{\pi} \right|^2 \right] + \mathbb{E}\left[\left| \Delta \widehat{Z}_n^{\pi} \right|^2 \right] \right) (1 + C \Delta t_n) \left(\mathbb{E}\left[\left| \Delta \widehat{Y}_{n+1}^{\pi} \right|^2 \right] + \mathbb{E}\left[\left| \Delta \widehat{Z}_{n+1}^{\pi} \right|^2 \right] \right)$$
$$+ C \Bigg\{ \Delta t_n^2 + \mathbb{E}\left[\int_{t_n}^{t_{n+1}} \left| D_{t_n} Z_r - \overline{DZ}_n^{n+1} \right|^2 \mathrm{d}r \right]$$
$$+ \frac{\epsilon_n^y + \epsilon_n^z + \Delta t_n \epsilon_n^{\gamma}}{\Delta t_n} \Bigg\}, \tag{5.46}$$

with some $\beta^* > 0$, depending on both β_z^*, β_y^*. Thereafter, for any sufficiently small time step admitting to $\beta^* \Delta t_n < 1$, an application of the discrete Grönwall lemma implies the total approximation error of Y and Z in (5.26).

The Γ estimate then follows in a similar manner to step 5 in Theorem 4.3 using (5.40), (5.38), (5.33), (5.45); and observing that $(1 + C\Delta t_n)/(1 - \beta^* \Delta t_n) - 1$ is $\mathcal{O}(|\pi|)$ given $\beta^* \Delta t_n < 1$. This completes the total approximation error of (5.26).

Step 4. *Derivative representation error of Z and Γ*. In order to prove (5.27), we need to establish an error estimate bounding the difference between the spatial derivative of (5.1c) and the target of (5.1b). Notice

that, under the conditions of Assumption 4.1 and (5.23), the arguments of the conditional expectations are all C_b^2 in x. Then, formal differentiation of (5.1c) with the Leibniz rule and the integration-by-parts formula in (B.5) applied on (5.1b) gives

$$\begin{aligned}\left(\nabla_x \check{z}_n^\pi\left(X_n^\pi\right) - \check{\gamma}_n^\pi\left(X_n^\pi\right)\right)\sigma = \Delta t_n[\left(\widehat{\gamma}_n^\pi\left(X_n^\pi\right) - \check{\gamma}_n^\pi\left(X_n^\pi\right)\right)\sigma]^T \mathbb{E}_n\left[\nabla_x \nabla_z f(t_{n+1}, \widehat{\mathbf{X}}_{n+1}^\pi)\right]\sigma \\ + \Delta t_n \mathbb{E}_n\left[\nabla_z f(t_{n+1}, \widehat{\mathbf{X}}_{n+1}^\pi)\right]\nabla_x \widehat{\gamma}_n^\pi\left(X_n^\pi\right)\sigma^2. \end{aligned} \tag{5.47}$$

By the bounded differentiability conditions in $(\mathbf{A}_2^{f,g})$, we have that

$$\begin{aligned}\mathbb{E}\left[\left|\left(\nabla_x \check{z}_n^\pi\left(X_n^\pi\right) - \check{\gamma}_n^\pi\left(X_n^\pi\right)\right)\sigma\right|^2\right] \leq 2\Delta t_n^2 L_{\nabla^2 f}^2 |\sigma|^2 \, \mathbb{E}\left[\left|\left(\widehat{\gamma}_n^\pi\left(X_n^\pi\right) - \check{\gamma}_n^\pi\left(X_n^\pi\right)\right)\sigma\right|^2\right] \\ + 2\Delta t_n^2 L_{\nabla f}^2 |\sigma|^4 \, \mathbb{E}\left[\left|\nabla_x \widehat{\gamma}_n^\pi\left(X_n^\pi\right)\right|^2\right]. \end{aligned} \tag{5.48}$$

Splitting the first term according to $\widehat{\gamma}_n^\pi\left(X_n^\pi\right) - \check{\gamma}_n^\pi\left(X_n^\pi\right) = \widehat{\gamma}_n^\pi\left(X_n^\pi\right) - \nabla_x \check{z}_n^\pi\left(X_n^\pi\right) + \nabla_x \check{z}_n^\pi\left(X_n^\pi\right) - \check{\gamma}_n^\pi\left(X_n^\pi\right)$, using the direct estimate $\widehat{\gamma}_n^\pi\left(X_n^\pi\right) \equiv \nabla_x \widehat{z}_n^\pi\left(X_n^\pi\right)$ implied by (5.16) and recalling the bounds in (5.23) subsequently yields

$$\mathbb{E}\left[\left|\left(\nabla_x \check{z}_n^\pi\left(X_n^\pi\right) - \check{\gamma}_n^\pi\left(X_n^\pi\right)\right)\sigma\right|^2\right] \leq C\left(\mathbb{E}\left[\left|\left(\nabla_x \check{z}_n^\pi\left(X_n^\pi\right) - \nabla_x \widehat{z}_n^\pi\right)\sigma\right|^2\right] + \Upsilon^6(S_1)\right), \tag{5.49}$$

for small enough time steps admitting to $4\Delta t_n^2 L_{\nabla^2 f}^2 |\sigma|^2 < 1$. Combining this estimate with the upper bound (5.32), recalling the definition of $\epsilon_n^{z,\nabla z}$ in (5.24), we gather

$$\begin{aligned}\mathbb{E}\left[\left|\check{Z}_n^\pi - \widehat{Z}_n^\pi\right|^2\right] + \Delta t_n \mathbb{E}\left[\left|\left(\check{\Gamma}_n^\pi - \nabla_x \widehat{Z}_n^\pi\right)\sigma\left(t_n, X_n^\pi\right)\right|^2\right] \\ \leq C\left(\epsilon_n^{z,\nabla z} + \Delta t_n^3 \Upsilon^6(S_1)\right) + 16 L_{\nabla f}^2 \Delta t_n \mathbb{E}\left[\int_{t_n}^{t_{n+1}} \left|D_n \widetilde{Z}_r - D_n \check{z}_n^\pi\right|^2 \mathrm{d}r\right], \end{aligned} \tag{5.50}$$

for small enough time steps $\Delta t_n < 1$ and diverging $\Upsilon(S_1)$. The total approximation error estimate in (5.27) then follows in a similar manner, combining (5.50) with (5.34), (5.46) and the discrete Grönwall lemma, as in the previous step.

This completes the proof. □

Theorem 5.2 establishes the convergence of the Deep BSDE approach to (5.1), given the UAT property of neural networks provided by Theorem 5.1. The first terms in both (5.26) and (5.27) correspond to the discrete time approximation errors in Theorem 4.3. The second terms correspond to the approximations of the neural network regression Monte Carlo approach. Provided by Theorem 5.1, the corresponding regression biases defined by (5.24) can be made arbitrarily small with the choice of shallow neural networks already. In exchange, to avoid the parametrization in the automatic differentiation approach in (5.16), one needs to restrict the parametrization to the case of $\Sigma_{C_b^2}(\tanh)$ neural networks and subsequently deal with an additional error term in (5.27), which depends on the increasing sequence $\Upsilon(S_1)$, controlling the magnitude of the parameters. If this dominating sequence is such that $\Upsilon^6(S_1)/N \to 0$ while $S_1, N \to \infty$, this ensures the existence of neural networks $\varphi(\cdot|\theta^y), \psi(\cdot|\theta^z) \in \Sigma_{C_b^2}(\tanh)$ such that the total approximation error converges. We shall, however,

notice that the claim above guarantees nothing more, and in fact does not guarantee the convergence of the final approximations including regression errors, which we highlight in the remark below.

REMARK 5.3 (Limitations of Theorem 5.2). In the proof of Theorem 5.2, we neglected the presence of three additional error terms. These are the following.

1. First, the definitions in (5.24) only express the regression biases due to the choice of a finite number of parameters. The actual regression errors also incorporate the approximation error of the optimal parameter space $\widehat{\theta}_n^z$ and induce a term $\mathbb{E}\big[\big|\varphi(X_n^\pi|\theta_n^{y,*}) - \varphi(X_n^\pi|\widehat{\theta}_n^y)\big|^2\big]$, which stems from the fact that, unlike in a linear regression method—see, e.g., Bender & Steiner (2012)—one does not have a closed-form expression for the true minimizers $(\theta_n^{z,*}, \theta_n^{\gamma,*})$, $\theta_n^{y,*}$, but can only gather an approximation of them with a stochastic gradient descent (SGD) optimization. The present understanding of this term is poor, mainly due to the nonconvexity of the corresponding target function—see Jentzen *et al.* (2021) and the references therein. Currently, there exists no theoretical guarantee that would ensure the convergence of SGD iterations in the FBSDE context. Furthermore, the second term at the right-hand side of (5.26) (respectively, (5.27)) implies that, in order to preserve the convergence of the total approximation error $\widehat{\mathcal{E}}^\pi(|\pi|)$, one needs $\epsilon_n^y + \epsilon_n^z$ $(\epsilon_n^y + \epsilon_n^{z,\nabla z})$ to be at least $\mathcal{O}(N^{-2})$ for each $n = 0, \dots, N-1$. In case of the regression biases defined by (5.24), this can be achieved by the UAT property in Theorem 5.1. Establishing a similar theoretical guarantee for the regression *errors* stemming from SGD approximations is currently not possible due to the aforementioned reasons. Nonetheless, in Fig. 3, we provide empirical evidence that suggests that this condition may indeed be satisfied in practice, encouraging further research in this direction.

2. The second term arises due to the fact that in practice one can only calculate an empirical version of the expectations in $\mathcal{L}_n^y, \mathcal{L}_n^{z,\gamma}, \mathcal{L}_n^{z,\nabla z}$. This induces a Monte Carlo simulation error of finitely many samples. However, as we shall see in the upcoming numerical section, thanks to the soft memory limitation of a single SGD step, one can pass so many realizations of the underlying Brownian motion throughout the optimization cycle that the magnitude of the corresponding error term becomes negligible compared to other sources of error.

3. The final observation that needs to be highlighted is the compactness assumption on the domain in Theorem 5.1. This error term can be dealt with in a similar fashion to Huré *et al.* (2020, Remark 4.2), where a localization argument is constructed in such a way that—under suitable truncation ranges—convergence is ensured.

6. Numerical experiments

In order to show the accuracy and robustness of the proposed scheme, we present results of numerical experiments on three different types of problems. We distinguish between the two Deep BSDE approaches for the OSM scheme, based on whether the Γ process is parametrized with an $\mathbb{R}^{d\times d}$-valued neural network—see Equation 5.15—or it is obtained as the direct Jacobian of the parametrization of the Z process via automatic differentiation—as in (5.16). We label these variants by (P) and (D), respectively. As a reference method, we compare the results of the OSM scheme to the first scheme (DBDP1) of Huré *et al.* (2020), which corresponds to the Euler discretization of (3.4) when $\vartheta_y = \vartheta_z = 1$. In accordance with their findings, we found the parametrized version (DBDP1) more robust than the automatic differentiated one (DBDP2) in high-dimensional settings.

Each BSDE is discretized with N equidistant time intervals, giving $\Delta t_n = T/N$ for all $n = 0, \dots, N-1$. For the implicit ϑ_y parameter of the discretization in (3.13), we choose values $\vartheta_y \in \{0, 1/2, 1\}$. In all upcoming examples, we use fully-connected, feedforward neural networks of $L = 2$ hidden layers with $S_l = 100 + d$ neurons in each layer. In line with Theorem 5.2, a hyperbolic tangent activation is deployed, yielding continuously differentiable parametrizations. Layer normalization (Ba *et al.*, 2016) is applied in between the hidden layers. For the stochastic gradient descent iterations, we use the Adam optimizer with the adaptive learning rate strategy of Chen & Wan (2021)—see $\eta(i)$ in Algorithm 1. The optimization is done as follows: in each backward recursion we allow $I = 2^{15}$ SGD iterations for the $N-1$th time step. Thereafter, we make use of the transfer learning initialization given by (5.18), and reduce the number of iterations to $I = 2^{11}$ for all preceding time steps. In each iteration step, the optimization receives a new, independent sample of the underlying forward diffusion with $B = 2^{10}$ sample paths, meaning that in total the iteration processes 2^{25} and 2^{21} many realizations of the Brownian motion at time step $n = N-1$ and $n < N-1$, respectively. In order to speed up normalization, neural network trainings were carried out with single floating point precision. For the implementation of the BCOS method, we choose $K = 2^9$ Fourier coefficients, $P = 5$ Picard iterations and truncate the infinite integrals to a finite interval of $[a, b] = [x_0 + \kappa_\mu - L\sqrt{\kappa_\sigma}, x_0 + \kappa_\mu + L\sqrt{\kappa_\sigma}]$, where $\kappa_\mu = \mu(0, x_0)T$, $\kappa_\sigma = \sigma(0, x_0)T$. As in Ruijter & Oosterlee (2016), we fix $L = 10$.

The OSM method has been implemented in TensorFlow 2. In order to exploit static graph efficiency, all core methods are decorated with `tf.function` decorators. The library used in this paper will be publicly accessible under github. All experiments below were run on a DELL Alienware Aurora R10 machine, equipped with an AMD Ryzen 9 3950X CPU (16 cores, 64Mb cache, 4.7GHz) and an Nvidia GeForce RTX 3090 GPU (24Gb). In order to assess the inherent stochasticity of both the regression Monte Carlo method and the SGD iterations, we run each experiment 5 times and report on the mean and standard deviations of the resulting independent approximations. $\mathbb{L}^2$-errors are estimated over an independent sample of size $M = 2^{10}$ produced by the same machinery as the one used for the simulations. Hence, the final error estimates are calculated as

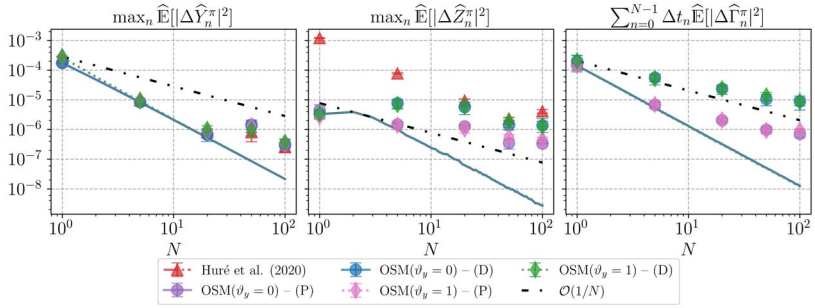
$$\begin{aligned} \widehat{\mathbb{E}}\big[\,|\Delta \widehat{Y}_n^\pi|^2\,\big] &= \frac{1}{M}\sum_{m=1}^{M} |\Delta \widehat{Y}_n^\pi(m)|^2, \quad \widehat{\mathbb{E}}\big[\,|\Delta \widehat{Z}_n^\pi|^2\,\big] = \frac{1}{M}\sum_{m=1}^{M} |\Delta \widehat{Z}_n^\pi(m)|^2, \\ \widehat{\mathbb{E}}\big[\,|\Delta \widehat{\Gamma}_n^\pi|^2\,\big] &= \frac{1}{M}\sum_{m=1}^{M} |\Delta \widehat{\Gamma}_n^\pi(m)|^2, \end{aligned} \tag{6.1}$$

where $\Delta Y_n^\pi(m)$ corresponds to the m'th path of test sample, and similarly for other error measures.

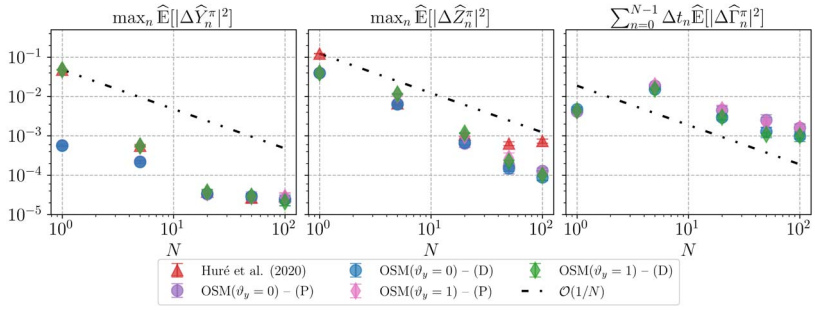
6.1 *Example 1: reaction-diffusion with diminishing control*

The first, *reaction-diffusion* type equation is taken from Gobet & Turkedjiev (2017, Example 2). Such equations are common in financial applications. The coefficients of the BSDE (1.1) are as follows:

$$\mu = \mathbf{0}_d, \quad \sigma = I_d, \quad f(t,x,y,z) = \frac{\omega(t,\lambda x)}{[1+\omega(t,\lambda x)]^2}\left[\lambda^2 d(y-\gamma) - 1 - \frac{\lambda^2}{2}d\right], \quad g(x) = \gamma + \frac{\omega(T,\lambda x)}{1+\omega(T,\lambda x)}, \tag{6.2}$$



(a) BCOS and Deep BSDE, $d = 1$. From left to right: maximum mean-squared approximation errors of Y and Z; average mean-squared approximation error of Γ. Lines correspond to BCOS estimates, scattered error bars to the means and standard deviations of 5 independent neural network regressions.



(b) Deep BSDE, $d = 10$. From left to right: maximum mean-squared approximation errors of Y and Z; average mean-squared approximation error of Γ. Means and standard deviations are calculated over 5 independent runs of the algorithm.

FIG. 1. Example 1 in (6.2). Convergence of approximation errors. Mean-squared errors are calculated over an independent sample of $M = 2^{10}$ realizations of the underlying Brownian motion.

where $\omega(t,x) = \exp(t + \sum_{i=1}^d x_i)$. These parameters satisfy Assumption 4.1. The driver is independent of Z and f^D does not depend on the Y process. Consequently, the solutions of (3.1b) and (3.1d) can be separated into two disjoint problems. The analytical solutions are given by

$$X_t = W_t, \quad y(t,x) = \frac{\omega(t,\lambda x)}{1+\omega(t,\lambda x)}, \quad z(t,x) = \lambda \frac{\omega(t,\lambda x)}{(1+\omega(t+\lambda x))^2}\mathbf{1}_d, \quad \gamma(t,x) = \lambda^2 \frac{\omega(t,\lambda x)(1-\omega(t,\lambda x))}{(1+\omega(t,\lambda x))^3}\mathbf{1}_{d,d}. \tag{6.3}$$

We choose $T = 0.5$, $\gamma = 0.6, \lambda = 1$ and fix $x_0 = \mathbf{1}_d$. We consider $d \in \{1, 10\}$ with $\vartheta_y \in \{0, 1\}$.

In Fig. 1, the convergence of the two fully-implementable schemes is assessed. Figure 1a depicts the convergence for $d = 1$. The BCOS estimates, drawn with lines, show the same order of convergence as in Theorem 4.3, confirming the theoretical findings of the discretization error analysis. The Deep BSDE approximations, depicted with scattered error bars, exhibit higher error figures, showcasing the presence of an additional regression component. Nevertheless, the complete approximation error of the corresponding regression estimates admit to the same order of convergence as in Theorem 5.2. The Γ approximations corresponding to the parametrized (P) and automatic differentiated (D) cases,

demonstrate the difference between the bounds in (5.26) and (5.27). Indeed, we observe an extra error stemming from the bounded differentiability component of the neural networks—see (5.23). The convergence of the regression approximations flattens out for the finest time partition $N = 100$—see the regression error of Y in particular—at a level of $\sim \mathcal{O}(10^{-7})$, indicating the presence of a regression bias induced by the restriction on a finite number of parameters. In Fig. 1b, the same dynamics are depicted for $d = 10$, where we observe the same order of convergence, in accordance with Theorem 5.2. Note that the regression estimates of the Z process converge until, and including, the finest time partition $N = 100$ in case of the OSM discretization. On the other hand, with the approach of Huré *et al.* (2020) the decay stops at $N = 50$, indicating the impact of diverging conditional variances, as anticipated in Remark 3.1. Table 1 contains the means and standard deviations of a collection of error measures with respect to 5 independent runs of the same regression Monte Carlo method. It can be seen that—regardless of the value of ϑ_y—the OSM scheme yields an order of magnitude improvement in the approximation of the Z process, while showing identical error figures in the Y process. Errors under the automatic differentiated case (D) with (5.16) are slightly better than in the parametrized approach (P). The Γ approximations show comparable accuracies. The total runtime of the OSM regressions is approximately double of that of Huré *et al.* (2020), which is intuitively explained by the fact that (5.1) solves two BSDEs at each point in time. Execution times under the automatic differentiated variant are slightly higher than in the parameterized case, confirming the extra computational complexity of Jacobian training in (5.16). The neural network regression Monte Carlo method yields sharp, robust estimates with small standard deviations over independent runs of the algorithm, in particular corresponding the Z process.

6.2 *Example 2: Hamilton–Jacobi–Bellman with LQG control*

The Hamilton–Jacobi–Bellman (HJB) equation is a nonlinear PDE derived from Bellman's dynamic programming principle, whose solution is the *value function* of a corresponding *stochastic control* problem. In what follows, we consider the linear-quadratic-Gaussian (LQG) control, which describes a linear system driven by additive noise (Han *et al.*, 2018). The FBSDE system (1.1), associated with the HJB equation has the following coefficients:

$$\mu = \mathbf{0}_d, \qquad \sigma = \sqrt{2} I_d, \qquad f(t,x,y,z) = |z|^2, \qquad g(x) = x^T A x + v^T x + c, \tag{6.4}$$

where $A \in \mathbb{R}^{d\times d}, v \in \mathbb{R}^{d\times 1}, c \in \mathbb{R}$. Unlike in Han *et al.* (2018), the hereby considered terminal condition is a quadratic mapping of space. This choice is made so that we have access to *semianalytical*, pathwise reference solutions $\{(Y_t, Z_t, \Gamma_t)\}_{0\le t\le T}$. Indeed, considering the associated parabolic problem (1.2), it is straightforward to show that the solution is given by

$$\begin{aligned} X_t = \sigma W_t, \quad y(t,x) = x^T P(t) x + Q^T(t) x + R(t), \\ z(t,x) = \sigma\big(\big[P(t) + P^T(t)\big]x + Q(t)\big), \quad \gamma(t,x) = \sigma\big[P(t) + P^T(t)\big], \end{aligned} \tag{6.5}$$

where the purely time dependent functions $P : [0,T] \to \mathbb{R}^{d\times d}, Q : [0,T] \to \mathbb{R}^{d\times 1}, R : [0,T] \to \mathbb{R}$ satisfy the following set of Riccati type ordinary differential equations (ODE)

$$\begin{gathered} \dot{P}(t) - \big[P(t) + P^T(t)\big]^2 = 0, \quad \dot{Q}(t) - 2\big[P(t) + P^T(t)\big]Q(t) = 0, \quad \dot{R}(t) + \mathrm{Tr}P(t) + P^T(t) - |Q(t)|^2 = 0, \\ P(T) = A, \quad Q(T) = v, \quad R(T) = c, \end{gathered} \tag{6.6}$$

TABLE 1 *Example* 1 *in (6.2), $d = 10$, $N = 100$. Summary of Deep BSDE estimates. Mean-squared errors are calculated over an independent sample of $M = 2^{10}$ realizations of the underlying Brownian motion. Means and standard deviations (in parentheses) obtained over* 5 *independent runs of the algorithm. Best estimates within one standard deviation highlighted in gray. Γ estimates from Huré et al. (2020) are obtained via automatic differentiation*

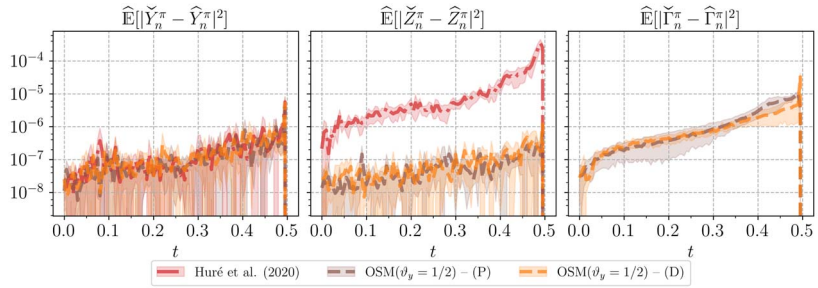
	OSM($\vartheta_y = 0$)		OSM($\vartheta_y = 1$)		Huré *et al.* (2020)
	(P)	(D)	(P)	(D)	
$\lvert\Delta\widehat{Y}_0^\pi\rvert/\lvert Y_0\rvert$	$3e-04\ (3e-04)$	$3e-04\ (2e-04)$	$6e-04\ (2e-04)$	$2e-04\ (2e-04)$	$1.1e-03\ (4e-04)$
$\lvert\Delta\widehat{Z}_0^\pi\rvert/\lvert Z_0\rvert$	$7e-03\ (3e-03)$	$8e-03\ (2e-03)$	$9e-03\ (2e-03)$	$9e-03\ (5e-03)$	$9e-03\ (2e-03)$
$\lvert\Delta\widehat{\Gamma}_0^\pi\rvert$	$1.2e-02\ (3e-03)$	$8e-03\ (3e-03)$	$9e-03\ (1e-03)$	$8e-03\ (2e-03)$	$9.9e+02\ (8e+01)$
$\max_n \widehat{\mathbb{E}}[\lvert\Delta\widehat{Y}_n^\pi\rvert^2]$	$2.4e-05\ (5e-06)$	$2.4e-05\ (7e-06)$	$2.7e-05\ (8e-06)$	$2.1e-05\ (4e-06)$	$2.9e-05\ (6e-06)$
$\max_n \widehat{\mathbb{E}}[\lvert\Delta\widehat{Z}_n^\pi\rvert^2]$	$1.3e-04\ (2e-05)$	$9e-05\ (1e-05)$	$1.1e-04\ (2e-05)$	$1.0e-04\ (3e-05)$	$7.4e-04\ (9e-05)$
$\sum_{n=0}^{N-1} \Delta t_n \widehat{\mathbb{E}}[\lvert\Delta\widehat{\Gamma}_n^\pi\rvert^2]$	$8e-04\ (2e-04)$	$5.0e-04\ (7e-05)$	$8e-04\ (2e-04)$	$5e-04\ (1e-04)$	$5.0e+03\ (8e+02)$
runtime (s)	$1.20e+03\ (1e+01)$	$1.44e+03\ (2e+01)$	$1.19e+03\ (1e+01)$	$1.43e+03\ (5e+01)$	$5.7e+02\ (3e+01)$

with $\dot{P} = \mathrm{d}P/\mathrm{d}t$, $\dot{Q} = \mathrm{d}Q/\mathrm{d}t$ and $\dot{R} = \mathrm{d}R/\mathrm{d}t$. The reference solution is then obtained by integrating (6.6) over a refined time grid of $N_{\mathrm{ODE}} = 10^4$ intervals.[4] We take $A = I_d$, $v = \mathbf{0}_d$, $c = 0$, $T = 0.5$ and fix $x_0 = \mathbf{1}_d$. An interesting feature of the FBSDE system defined by (6.4) is that the driver is independent of Y meaning that the Malliavin BSDE in (3.1d) can be solved separately from the backward equation. Consequently, the discrete time approximations of Z and Γ in (3.13) do not depend on ϑ_y. Moreover, the driver is quadratically growing in Z, in particular, it is only Lipschitz continuous over compact domains. Nevertheless, we include this problem to show promising results beyond Assumption 4.1. We pick $\vartheta_y = 1/2$ and investigate the solution in $d \in \{1, 50\}$.

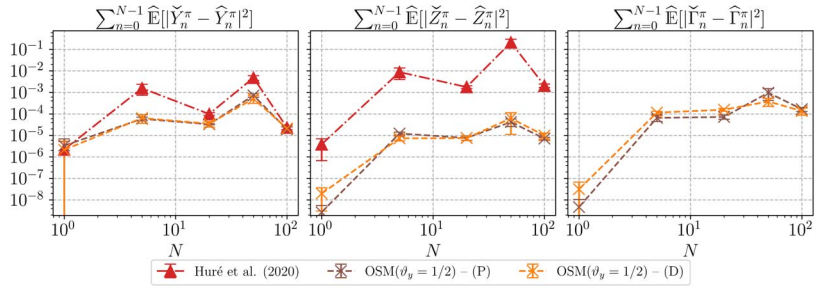
In Fig. 2, the regression errors of the Deep BSDE approach are assessed in $d = 1$. The true regression targets in (5.1) are benchmarked according to BCOS. In fact, at time step n, the corresponding cosine expansion coefficients are recovered by means of DCT, given neural network approximations $\widehat{Y}^{\pi}_{n+1}, \widehat{Z}^{\pi}_{n+1}, \widehat{\Gamma}^{\pi}_{n+1}$. These coefficients are subsequently plugged in (5.7) to gather BCOS estimates. For large enough Fourier domains and sufficiently many Picard iterations, the COS error becomes negligible compared to the discretization component and the resulting estimates approximate the true regression labels $\check{Y}^{\pi}_{n}, \check{Z}^{\pi}_{n}, \check{\Gamma}^{\pi}_{n}$. Hence, they can then be used to assess the regression errors induced by the Monte Carlo method. Figure 2a depicts these regression errors over time for $N = 100$. As it can be seen, the model of Huré *et al.* (2020) and the OSM scheme result in similar regression error components for the Y process. However, the regression errors of the Z process are three orders of magnitude worse in case of the reference method (Huré *et al.*, 2020), and in fact, dominate the total approximation error at $n = N-1$. In contrast, the OSM estimates—middle plot of Fig. 2a—exhibit the same order of regression error as for the Y process. This demonstrates the advantageous conditional variance behavior of the corresponding OSM estimates, as pointed out in Remark 3.1. The regression errors of the Γ process show comparable figures. The cumulative regression errors, corresponding to the second term in Theorem 5.2, are collected in Fig. 3b. In case of the model in Huré *et al.* (2020), the cumulative regression error of the Z process blows up as the mesh size $|\pi| = T/N$ decreases. On the contrary, the cumulative regression errors in all processes (Y, Z, Γ) are at a constant level of $\mathcal{O}(10^{-5})$ for the OSM scheme. In light of Remark 5.3, this indicates that the chosen, finite network architecture incorporates a regression bias that cannot be further reduced. In our experiments, we found that it is difficult to decrease this component by changing the number of hidden layers L or neurons per hidden layer S_l. Assessing this phenomenon requires a better understanding of both narrow UAT estimates and the convergence of SGD iterations.

In Fig. 3, the $d = 50$ dimensional case is depicted. In order to have dimension independent scales, *relative* mean-squared errors are reported. Figure 3a collects the relative approximation error over the discretized time window when $N = 100$. Compared to Huré *et al.* (2020), the OSM estimates yield a significant improvement in each part of the solution triple. In particular, the approximation errors of the Z process are three orders of magnitude better with both the parametrized (P) and automatic differentiated (D) approaches. In case of the Γ process, two observations can be made. First, the corresponding curve demonstrates that naive automatic differentiation of the Z approximations in Huré *et al.* (2020) does not provide reliable Γs. Moreover, it can be seen that the parametrized version (P) of the Deep BSDE approach given by Equation 5.15 provides an order of magnitude better average Γ errors. The convergence of the total approximation errors is depicted in Fig. 3b. The neural network regression estimates converge for both the parametrized (P) and the automatic differentiated (D) loss functions until $N = 50$, when the regression bias becomes apparent. Additionally, the convergence of the Γ

[4] This is done using scipy.integrate.odeint.



(a) Regression errors over time, $d = 1$, $N = 100$. From left to right: mean-squared regression errors of the Y, Z and Γ approximations over the discrete time window.

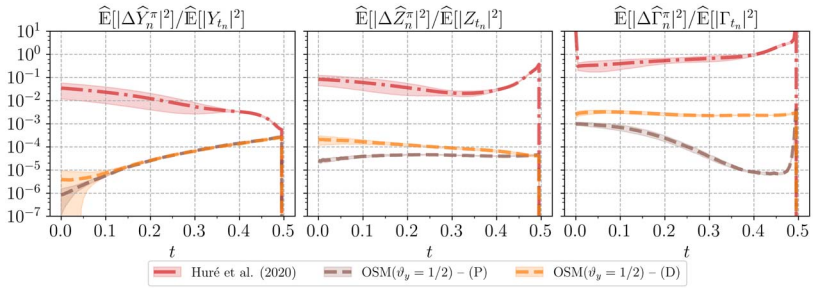


(b) Convergence of cumulative regression errors, $d = 1$. From left to right: cumulative regression errors of the Y, Z and Γ approximations over the number of time steps N.

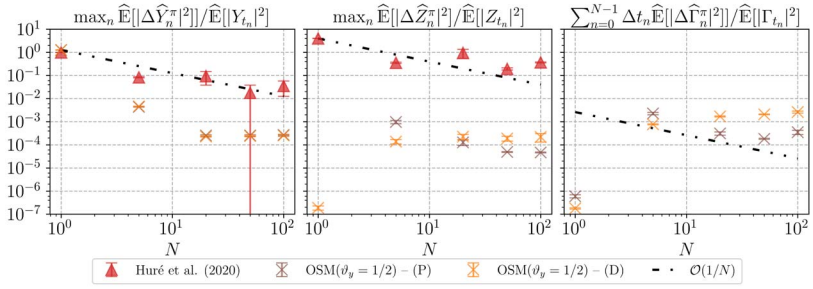
FIG. 2. Example 2 in (6.4). Neural network regression errors in $d = 1$. The true regression targets of (5.1) are identified by BCOS estimates. Mean-squared errors are calculated over an independent sample of $M = 2^{10}$ realizations of the underlying Brownian motion. Means and standard deviations are obtained over 5 independent runs of the algorithm.

approximations is significantly better in the parametrized case, suggesting that for such a quadratically scaling driver the last term of (5.27) is a driving error component.

In Table 2, means and standard deviations of a collection of error measures are gathered, with respect to 5 independent runs of the same regression Monte Carlo method, for both $d = 1$ and $d = 50$. The numbers are in line with the observations above. In particular, we highlight that the error terms corresponding to the Z and Γ approximations are four orders of magnitude better than in case of the reference method (Huré *et al.*, 2020). The parametrized version (P) of the Deep BSDE shows consistently better convergence. The total runtime of the neural network regression Monte Carlo approach is moderately increased between $d = 1$ and $d = 50$. In fact, the average execution time of a single SGD step for the parametrized (P) case in Equation 5.15 increases from $2.8e-3(4e-4)$ to $3.3e-3(4e-4)$ seconds in between $d = 1$ and $d = 50$. The same numbers for the automatic differentiated formulation (D) in (5.16) are $3.8(4e-4)$ and $4.4e-3(5e-4)$ seconds. These figures demonstrate the aforementioned methods' scalability for high-dimensional FBSDE systems. Finally, we point out that the OSM estimates are robust over independent runs of the algorithm as showcased by the small standard deviations in Table 2.



(a) Relative approximation errors over time, $d = 50$, $N = 100$. From left to right: relative mean-squared approximation errors of Y, Z and Γ over the discrete time window.



(b) Convergence of relative approximation errors, $d = 50$. From left to right: maximum relative mean-squared error of the Y, Z approximations; average relative mean-squared error of the Γ approximations.

FIG. 3. Example 2 in (6.4). $d = 50$. Relative approximation errors. Mean-squared errors are calculated over an independent sample of $M = 2^{10}$ realizations of the underlying Brownian motion. Means and standard deviations are obtained over 5 independent runs of the algorithm. Γ estimates from Huré *et al.* (2020) are obtained via automatic differentiation.

6.3 *Example 3: space-dependent diffusion coefficients*

Our final example is taken from Milstein & Tretyakov (2006); Ruijter & Oosterlee (2016) and it is meant to demonstrate that the conditions in Assumption 4.1 can be substantially relaxed. The FBSDE system (1.1) is defined by the following coefficients:

$$\mu_i(t,x) = \frac{\left(1+x_i^2\right)}{\left(2+x_i^2\right)^3}, \quad \sigma_{ij}(t,x) = \frac{1+x_i x_j}{2+x_i x_j}\delta_{ij},$$

$$f(t,x,y,z) = \frac{1}{\lambda(t+\tau)}\exp\left(-\frac{x^T x}{\lambda(t+\tau)}\right)\left[4\sum_{i=1}^{d}\frac{x_i^2\left(1+x_i^2\right)}{\left(2+x_i^2\right)^3} + \sum_{i=1}^{d}\frac{\left(1+x_i^2\right)^2}{\left(2+x_i^2\right)^2}\left(1-2\frac{x_i^2}{\lambda(t+\tau)}\right) - \sum_{i=1}^{d}\frac{x_i^2}{t+\tau}\right]$$

$$+\sqrt{\frac{1+y^2+\exp\left(-\frac{2x^T x}{\lambda(t+\tau)}\right)}{1+2y^2}}\sum_{i=1}^{d}\frac{z_i x_i}{\left(2+x_i^2\right)^2}, \qquad g(x) = \exp\left(-\frac{x^T x}{\lambda(T+\tau)}\right). \tag{6.7}$$

TABLE 2 *Example 2 in (6.4). Summary of Deep BSDE estimates. Mean-squared errors are calculated over an independent sample of $M = 2^{10}$ realizations of the underlying Brownian motion. Means and standard deviations (in parentheses) obtained over 5 independent runs of the algorithm. Best estimates within one standard deviation highlighted in gray. Γ estimates from Huré et al. (2020) are obtained via automatic differentiation*

	(a) $d = 1, N = 100$. OSM($\vartheta_y = 1/2$)		Huré *et al.* (2020)
	(P)	(D)	
$\lvert\Delta\widehat{Y}_0^\pi\rvert/\lvert Y_0\rvert$	$1.1e-03\,(5e-04)$	$2e-03\,(1e-03)$	$1.5e-03\,(3e-04)$
$\lvert\Delta\widehat{Z}_0^\pi\rvert/\lvert Z_0\rvert$	$1.3e-04\,(9e-05)$	$8e-05\,(9e-05)$	$1e-03\,(1e-03)$
$\lvert\Delta\widehat{\Gamma}_0^\pi\rvert/\lvert\Gamma_0\rvert$	$1.0e-04\,(5e-05)$	$2e-04\,(1e-04)$	$1.05e+00\,(7e-02)$
$\max_n \widehat{\mathbb{E}}[\lvert\Delta\widehat{Y}_n^\pi\rvert^2]$	$8e-06\,(2e-06)$	$8e-06\,(3e-06)$	$1.1e-04\,(1e-05)$
$\max_n \widehat{\mathbb{E}}[\lvert\Delta\widehat{Z}_n^\pi\rvert^2]$	$8e-07\,(3e-07)$	$1.4e-06\,(6e-07)$	$6.4e-03\,(3e-04)$
$\sum_{n=0}^{N-1} \Delta t_n \widehat{\mathbb{E}}[\lvert\Delta\widehat{\Gamma}_n^\pi\rvert^2]$	$8e-07\,(4e-07)$	$2.8e-06\,(9e-07)$	$5.5e-03\,(7e-04)$
runtime (s)	$1.18e+03\,(4e+01)$	$1.41e+03\,(3e+01)$	$5.7e+02\,(4e+01)$
	(b) $d = 50, N = 100$. OSM($\vartheta_y = 1/2$)		Huré *et al.* (2020)
	(P)	(D)	
$\lvert\Delta\widehat{Y}_0^\pi\rvert/\lvert Y_0\rvert$	$8e-04\,(5e-04)$	$1e-03\,(1e-03)$	$1.7e-01\,(8e-02)$
$\lvert\Delta\widehat{Z}_0^\pi\rvert/\lvert Z_0\rvert$	$5.0e-03\,(5e-04)$	$1.4e-02\,(3e-03)$	$2.8e-01\,(7e-02)$
$\lvert\Delta\widehat{\Gamma}_0^\pi\rvert/\lvert\Gamma_0\rvert$	$3.1e-02\,(2e-03)$	$4.9e-02\,(7e-03)$	$3.5e+00\,(1e-01)$
$\max_n \widehat{\mathbb{E}}[\lvert\Delta\widehat{Y}_n^\pi\rvert^2]$	$2.7e+00\,(1e-01)$	$2.5e+00\,(3e-01)$	$7e+01\,(4e+01)$
$\max_n \widehat{\mathbb{E}}[\lvert\Delta\widehat{Z}_n^\pi\rvert^2]$	$3.4e-02\,(1e-03)$	$3.1e-02\,(3e-03)$	$2.8e+02\,(1e+01)$
$\sum_{n=0}^{N-1} \Delta t_n \widehat{\mathbb{E}}[\lvert\Delta\widehat{\Gamma}_n^\pi\rvert^2]$	$4.1e-04\,(6e-05)$	$3.3e-03\,(2e-04)$	$2.9e+00\,(2e-01)$
runtime (s) $1.36e+03\,(1e+01)$	$1.62e+03\,(4e+01)$	$6.16e+02\,(1e+01)$	$6.16e+2\,(1e+1)$

The analytical solutions are given by

$$y(t,x) = \exp\left(-\frac{x^T x}{\lambda(t+\tau)}\right), \quad z_j(t,x) = -\frac{1+x_j^2}{2+x_j^2}\frac{2\exp\left(-\frac{x^T x}{\lambda(t+\tau)}\right)}{\lambda(t+\tau)}x_j, \quad \gamma_{ij}(t,x) = \partial_j z_i(t,x). \tag{6.8}$$

We use $T = 10, \lambda = 10, \tau = 1, d = 1$ and fix $x_0 = \mathbf{1}_d$. Notice that μ and σ are both C_b^2. In conjecture with Appendix A, this implies that the Euler–Maruyama schemes in (3.2) and (3.5) have an $\mathbb{L}^2$ convergence rate of order $1/2$. Additionally, by Itô's formula, the unique solution of the SDE is given by the closed form expression (Milstein & Tretyakov, 2006)

$$X_t = \Lambda(x_0 + \arctan(x_0) + W_t), \tag{6.9}$$

where $\Lambda : \mathbb{R} \to \mathbb{R}$ is defined implicitly $\Lambda(r) + \arctan(r) := r$ for any $r \in \mathbb{R}$, and applied element-wise. It is straightforward to check that $\Lambda \in C_b^1(\mathbb{R};\mathbb{R})$, in particular $\Lambda'(r) = \frac{1+\Lambda^2(r)}{2+\Lambda^2(r)}$ implying that Λ is a

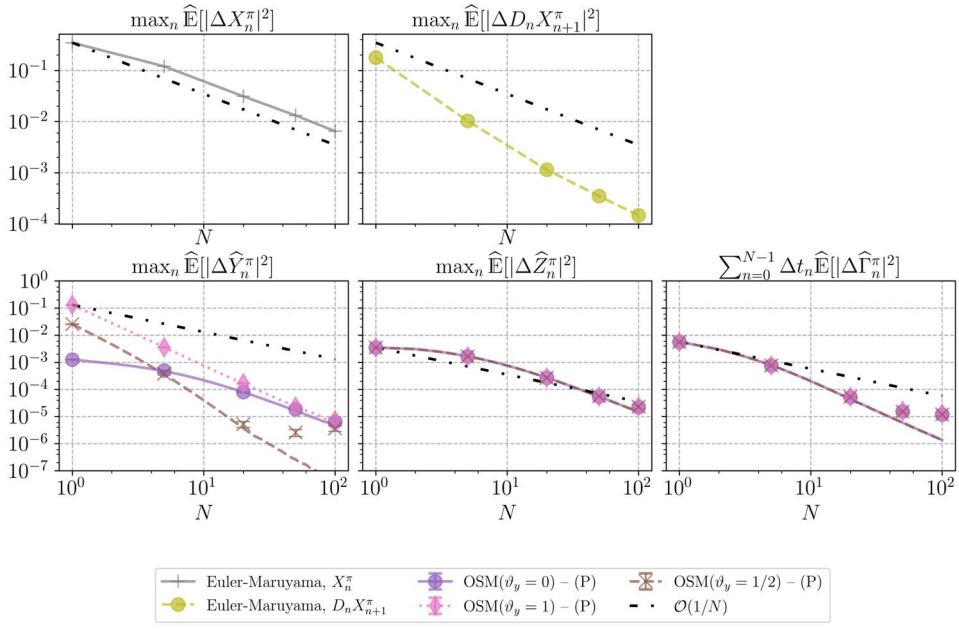


FIG. 4. Example 3 in (6.7). Convergence of approximation errors for $d = 1$. From left to right, top to bottom: maximum mean-squared errors of Euler–Maruyama approximations of X and DX; maximum mean-squared approximation errors of Y and Z; average mean-squared approximation error of Γ. Lines correspond to BCOS estimates, scattered error bars to the means and standard deviations of 5 independent neural network regressions. The mean errors are obtained over an independent sample of $M = 2^{10}$ trajectories of the underlying Brownian motion.

bijective. In light of the Malliavin chain rule formula in Lemma 2.1, we then also have

$$D_s X_t = \frac{1 + \Lambda^2(x + \arctan(x) + W_t)}{2 + \Lambda^2(x + \arctan(x) + W_t)} \mathbb{1}_{s \le t}. \tag{6.10}$$

We assess the convergence of the Euler–Maruyama estimates in (3.2)–(3.5) by solving the nonlinear equation in (6.9) for each realization of the Brownian motion.[5] The results of the numerical simulations in $d = 1$ are given in Fig. 4 for the parametrized Deep BSDE case and $\vartheta_y = 0, 1/2, 1$. We see that, in line with Appendix A, $D_n X_{n+1}^\pi$ inherits the convergence rate of X_n^π. The convergence rates of $(\widehat{Y}_n^\pi, \widehat{Z}_n^\pi, \widehat{\Gamma}_n^\pi)$ are of the same order as in Theorem 5.2. The BCOS estimates and the Deep BSDE approach exhibit coinciding error figures until a magnitude of $\mathcal{O}(10^{-6})$ is reached, when the regression bias becomes apparent. Similar convergence behavior is observed in high-dimensions. The results suggest that the convergence of the OSM scheme can be extended to the nonadditive noise case.

[5] This is done by `scipy.optimize.root`'s `df-sane` algorithm, which deploys the method in La Cruz *et al.* (2006).

7. Conclusion

In this paper, we introduced the OSM scheme, a new discretization for Malliavin differentiable FBSDE systems where the control process is estimated by solving the linear BSDE driving the Malliavin derivatives of the solution pair. The main contributions can be summarized as follows. The discretization in (3.13) includes Γ estimates, linked to the Hessian matrix of the associated parabolic problem. In Theorem 4.3, we have shown that under standard Lipschitz assumptions and additive noise in the forward diffusion, the aforementioned discrete time approximations admit to an $\mathbb{L}^2$ convergence of order $1/2$. We gave two fully-implementable schemes. In case of one-dimensional problems, we extended the BCOS method (Ruijter & Oosterlee, 2015), and gathered approximations via Fourier cosine expansions in (5.7). For high-dimensional equations, similarly to recent Deep BSDE methods (Han *et al.*, 2018; Huré *et al.*, 2020), we formulated a neural network regression Monte Carlo approach, where the corresponding processes of the solution triple are parametrized by fully-connected, feedforward neural networks. We carried out a complete regression error analysis in Theorem 5.2 and showed that the neural network parametrizations are consistent with the discretization, in terms of regression biases controlled by the universal approximation property. We supported our theoretical findings by numerical experiments and demonstrated the accuracy and robustness of the proposed approaches for a range of high-dimensional problems. Using BCOS estimates as benchmarks for one-dimensional equations, we empirically assessed the regression errors induced by stochastic gradient descent. Our findings with the Deep BSDE approach showcase accurate approximations for each process in (5.1), and in particular exhibit significantly improved approximations of the Z process for heavily control dependent equations.

Acknowledgements

The authors would like to thank the anonymous referees for their constructive and detailed feedback, which greatly improved the presentation of the paper. The first author would like to thank Adam Andersson for the fruitful discussions in the early stages of this work. The first author also acknowledges financial support from the Peter Paul Peterich Foundation via the TU Delft University Fund.

References

Alanko, S. & Avellaneda, M. (2013) Reducing variance in the numerical solution of BSDEs. *C. R. Math.*, **351**, 135–138.

Ba, J. L., Kiros, J. R. & Hinton, G. E. (2016) Layer normalization. arXiv:1607.06450 [cs, stat] arXiv: 1607.06450.

Bally, V. & Pagès, G. (2003) A quantization algorithm for solving multidimensional discrete-time optimal stopping problems. *Bernoulli*, **9**, 1003–1049.

Beck, C., Weinan, E. & Jentzen, A. (2019) Machine learning approximation algorithms for high-dimensional fully nonlinear partial differential equations and second-order backward stochastic differential equations. *J. Nonlinear Sci.*, **29**, 1563–1619.

Bender, C. & Denk, R. (2007) A forward scheme for backward SDEs. *Stochastic Process. Appl.*, **117**, 1793–1812.

Bender, C. & Steiner, J. (2012) Least-squares Monte Carlo for backward SDEs. *Numerical Methods in Finance* (R. A. Carmona, P. Del Moral, P. Hu & N. Oudjane eds). Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 257–289.

Bouchard, B. & Touzi, N. (2004) Discrete-time approximation and Monte-Carlo simulation of backward stochastic differential equations. *Stochastic Process. Appl.*, **111**, 175–206.

Briand, P. & Labart, C. (2014) Simulation of BSDEs by wiener chaos expansion. *Ann. Appl. Probab.*, **24**, 1129–1171.

Chassagneux, J.-F. & Richou, A. (2016) Numerical simulation of quadratic BSDEs. *Ann. Appl. Probab.*, **26**, 262–304.

Chen, Y. & Wan, J. W. L. (2021) Deep neural network framework based on backward stochastic differential equations for pricing and hedging American options in high dimensions. *Quant. Finance*, **21**, 45–67.

Cheridito, P., Soner, H. M., Touzi, N. & Victoir, N. (2007) Second-order backward stochastic differential equations and fully nonlinear parabolic PDEs. *Comm. Pure Appl. Math*.

Cybenko, G. (1989) Approximation by superpositions of a sigmoidal function. *Math. Control Signals Systems*, **2**, 303–314.

Delarue, F. & Menozzi, S. (2006) A forward–backward stochastic algorithm for quasi-linear PDES. *Ann. Appl. Probab.*, **16**, 140–184.

El Karoui, Peng, S. & Quenez, M. C. (1997) Backward stochastic differential equations in finance. *Math. Finance*, **7**, 1–71.

Fahim, A., Touzi, N. & Warin, X. (2011) A probabilistic numerical method for fully nonlinear parabolic PDEs. *Ann. Appl. Probab.*, **21**, 1322–1364.

Fang, F. & Oosterlee, C. W. (2009) A novel pricing method for European options based on Fourier-cosine series expansions. *SIAM J. Sci. Comput.*, **31**, 826–848.

Fujii, M., Takahashi, A. & Takahashi, M. (2019) Asymptotic expansion as prior knowledge in deep learning method for high dimensional BSDEs. *Asia-Pac. Financ. Mark.*, **26**, 391–408.

Germain, M., Pham, H. & Warin, X. (2021) Approximation error analysis of some deep backward schemes for nonlinear PDEs. arXiv:2006.01496 [math] arXiv: 2006.01496.

Glorot, X. & Bengio, Y. (2010) Understanding the difficulty of training deep feedforward neural networks. *Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics*. JMLR Workshop and Conference Proceedings, pp. 249–256. ISSN: 1938-7228.

Gobet, E., Lemor, J.-P. & Warin, X. (2005) A regression-based Monte Carlo method to solve backward stochastic differential equations. *Ann. Appl. Probab.*, **15**, 2172–2202.

Gobet, E. & Turkedjiev, P. (2017) Adaptive importance sampling in least-squares Monte Carlo algorithms for backward stochastic differential equations. *Stochastic Process. Appl.*, **127**, 1171–1203.

Goodfellow, I., Bengio, Y. & Courville, A. (2016) *Deep Learning*. Cambridge, Massachusetts: MIT Press.

Han, J., Jentzen, A. & Weinan, E. (2018) Solving high-dimensional partial differential equations using deep learning. *Proc. Natl. Acad. Sci.*, **115**, 8505–8510.

Han, J. & Long, J. (2020) Convergence of the deep BSDE method for coupled FBSDEs. *Probab. Uncertain. Quant. Risk*, **5**, 1–33.

Hornik, K., Stinchcombe, M. & White, H. (1990) Universal approximation of an unknown mapping and its derivatives using multilayer feedforward networks. *Neural Netw.*, **3**, 551–560.

Hu, Y., Nualart, D. & Song, X. (2011) Malliavin calculus for backward stochastic differential equations and application to numerical solutions. *Ann. Appl. Probab.*, **21**, 2379–2423.

Huré, C., Pham, H. & Warin, X. (2020) Deep backward schemes for high-dimensional nonlinear PDEs. *Math. Comp.*, **89**, 1547–1579.

Imkeller, P. & Dos Reis, G. (2010) Path regularity and explicit convergence rate for BSDE with truncated quadratic growth. *Stochastic Process. Appl.*, **120**, 348–379.

Jentzen, A., Kuckuck, B., Neufeld, A. & von Wurstemberger, P. (2021) Strong error analysis for stochastic gradient descent optimization algorithms. *IMA J. Numer. Anal.*, **41**, 455–492.

Karatzas, I. & Shreve, S. (1998) *Brownian Motion and Stochastic Calculus*, 2nd edn. Graduate Texts in Mathematics. New York: Springer.

Kloeden, P. E. & Platen, E. (1992) *Numerical Solution of Stochastic Differential Equations*. Berlin, Heidelberg: Springer.

La Cruz, Martínez, J. & Raydan, M. (2006) Spectral residual method without gradient information for solving large-scale nonlinear systems of equations. *Math. Comp.*, **75**, 1429–1449.

Ma, J., Protter, P. & Yong, J. (1994) Solving forward–backward stochastic differential equations explicitly—a four step scheme. *Probab. Theory Related Fields*, **98**, 339–359.

Ma, J. & Zhang, J. (2002) Representation theorems for backward stochastic differential equations. *Ann. Appl. Probab.*, **12**, 1390–1418.

Mastrolia, T., Possamaï, D. & Réveillac, A. (2017) On the Malliavin differentiability of BSDEs. *Ann. Inst. H. Poincaré Probab. Statist.*, **53**, 464–492.

Milstein, G. N. & Tretyakov, M. V. (2006) Numerical algorithms for forward–backward stochastic differential equations. *SIAM J. Sci. Comput.*, **28**, 561–582.

Nualart, D. (2006) *The Malliavin Calculus and Related Topics*, 2nd edn. Probability and Its Applications. Berlin Heidelberg: Springer.

Pardoux, E. & Peng, S. (1992) Backward stochastic differential equations and quasilinear parabolic partial differential equations. *Stochastic Partial Differential Equations and Their Applications* (B. L. Rozovskii & R. B. Sowers eds). Lecture Notes in Control and Information Sciences. Berlin, Heidelberg: Springer, pp. 200–217.

Pinkus, A. (1999) Approximation theory of the MLP model in neural networks. *Acta Numer.*, **8**, 143–195.

Ruijter, M. J. & Oosterlee, C. W. (2015) A Fourier cosine method for an efficient computation of solutions to BSDEs. *SIAM J. Sci. Comput.*, **37**, A859–A889.

Ruijter, M. J. & Oosterlee, C. W. (2016) Numerical Fourier method and second-order Taylor scheme for backward SDEs in finance. *Appl. Numer. Math.*, **103**, 1–26.

Turkedjiev, P. (2015) Two algorithms for the discrete time approximation of Markovian backward stochastic differential equations under local conditions. *Electron. J. Probab.*, **20**, 49.

Zhang, J. (2004) A numerical scheme for BSDEs. *Ann. Appl. Probab.*, **14**, 459–488.

Appendix A. Convergence of $D_n X^{\pi}_{n+1}$

We show the convergence of $D_n X^{\pi}_{n+1}$ estimates of the Euler–Maruyama discretization (3.5) under the assumptions

($\widetilde{\mathbf{A}}_1^{\sigma,\mu}$) σ is uniformly bounded;

($\widetilde{\mathbf{A}}_2^{\sigma,\mu}$) $\mu \in C_b^{0,1}(\mathbb{R}^{d\times 1};\mathbb{R})$, $\sigma \in C_b^{0,1}(\mathbb{R}^{d\times 1};\mathbb{R}^{d\times d})$. In particular, both of them are Lipschitz continuous in x.

From the estimation (3.5) and the linear SDE of the Malliavin derivative in (3.1c)—using the inequality $(a+b+c)^2 \le 3(a^2+b^2+c^2)$, on top of the $L^2([0,T];\mathbb{R}^{d\times d})$ Cauchy–Schwarz inequality and Itô's isometry—it follows

$$\begin{aligned}\mathbb{E}\left[\left|D_{t_n}X_{t_{n+1}} - D_n X^{\pi}_{n+1}\right|^2\right] &\le 3\mathbb{E}\left[\left|\sigma\left(t_n, X_{t_n}\right) - \sigma\left(t_n, X^{\pi}_n\right)\right|^2\right] \\ &\quad + 3\Delta t_n \mathbb{E}\left[\int_{t_n}^{t_{n+1}} \left|\nabla_x \mu(r, X_r) D_{t_n} X_r - \nabla_x \mu\left(t_n, X^{\pi}_n\right)\sigma\left(t_n, X^{\pi}_n\right)\right|^2 \mathrm{d}r\right] \\ &\quad + 3\mathbb{E}\left[\int_{t_n}^{t_{n+1}} \left|\nabla_x \sigma(r, X_r) D_{t_n} X_r - \nabla_x \sigma\left(t_n, X^{\pi}_n\right)\sigma\left(t_n, X^{\pi}_n\right)\right|^2 \mathrm{d}r\right]. \quad \text{(A.1)}\end{aligned}$$

Bounded continuous differentiability in ($\widetilde{\mathbf{A}}_2^{\sigma,\mu}$), in particular, implies Lipschitz continuity. Furthermore, by the uniform boundedness of the diffusion coefficient and the mean-squared continuity of $D_{t_n}X$ in (2.6), we gather

$$\mathbb{E}\left[\left|D_{t_n}X_{t_{n+1}} - D_n X^{\pi}_{n+1}\right|^2\right] \le 3L_\sigma^2 \mathbb{E}\left[\left|X_{t_n} - X^{\pi}_n\right|^2\right] + C\Delta t_n, \quad \text{(A.2)}$$

for any $\Delta t_n < 1$. Then, due to the discretization error of the Euler–Maruyama estimates given by (3.3), we conclude $\limsup_{|\pi|\to 0} \frac{1}{|\pi|}\mathbb{E}\left[\left|D_{t_n}X_{t_{n+1}} - D_n X^{\pi}_{n+1}\right|^2\right] < \infty$.

Appendix B. Integration by parts formulas

For the formula in (5.5), we refer to Ruijter & Oosterlee (2015, A.1). In order to prove (5.6), let $v : [0,T] \times \mathbb{R} \to \mathbb{R}$ and consider

$$\mathbb{E}_n^x\big[v\left(t_{n+1}, X_{n+1}^{\pi}(\Delta W_n)\right)\Delta W_n^2\big] = \mathbb{E}_n^x\left[\frac{1}{\sqrt{2\pi\,\Delta t_n}}\int_{\mathbb{R}} v\left(t_{n+1}, X_{n+1}^{\pi}(\nu)\right)\nu^2 e^{-\frac{1}{2\Delta t_n}\nu^2}\mathrm{d}\nu\right], \tag{B.1}$$

with the Euler–Maruyama approximations $X_{n+1}^{\pi}(\Delta W_n) = x + \mu(t_n,x)\Delta t_n + \sigma(t_n,x)\Delta W_n$. For a sufficiently smooth v, integration by parts implies

$$\begin{aligned}
&\mathbb{E}_n^x\left[\frac{1}{\sqrt{2\pi\,\Delta t_n}}\int_{\mathbb{R}} v\left(t_{n+1}, X_{n+1}^{\pi}(\nu)\right)\nu^2 e^{-\frac{1}{2\Delta t_n}\nu^2}\mathrm{d}\nu\right]\\
&\quad= \frac{1}{\sqrt{2\pi\,\Delta t_n}}\mathbb{E}_n^x\Big[-\Delta t_n\big[\nu v\left(t_{n+1}, X_{n+1}^{\pi}(\nu)\right)e^{-\nu^2/(2\Delta t_n)}\big]_{-\infty}^{+\infty} + \Delta t_n\int_{\mathbb{R}} v\left(t_{n+1}, X_{n+1}^{\pi}(\nu)\right)e^{-\frac{1}{2\Delta t_n}\nu^2}\mathrm{d}\nu\\
&\qquad\qquad + \Delta t_n\sigma(t_n,x)\int_{\mathbb{R}}\partial_x v\left(t_{n+1}, X_{n+1}^{\pi}(\nu)\right)\nu e^{-\frac{1}{2\Delta t_n}\nu^2}\mathrm{d}\nu\Big].
\end{aligned} \tag{B.2}$$

For a v with sufficient radial decay, we therefore conclude that

$$\mathbb{E}_n^x\big[v\left(t_{n+1}, X_{n+1}^{\pi}\right)\Delta W_n^2\big] = \Delta t_n\mathbb{E}_n^x\left[v\left(t_{n+1}, X_{n+1}^{\pi}\right)\right] + \Delta t_n^2\sigma^2(t_n,x)\mathbb{E}_n^x\big[\partial_{xx}^2 v\left(t_{n+1}, X_{n+1}^{\pi}\right)\big], \tag{B.3}$$

by the estimate in (5.5).

Thereupon, given a cosine expansion approximation $v(t_{n+1},\rho) \approx {\sum}'_{k=0}^{K-1}\mathcal{V}_k(t_{n+1})\cos(k\pi\frac{\rho-a}{b-a})$, the corresponding spatial derivative approximations are given by $\partial_x v(t_{n+1},\rho) \approx {\sum}'_{k=0}^{K-1} -\mathcal{V}_k(t_{n+1})\frac{k\pi}{b-a}\sin(k\pi\frac{\rho-a}{b-a})$, $\partial_{xx}^2 v(t_{n+1},\rho) \approx {\sum}'_{k=0}^{K-1} -\mathcal{V}_k(t_{n+1})\left(\frac{k\pi}{b-a}\right)^2\cos(k\pi\frac{\rho-a}{b-a})$. Then (5.5)–(5.6) follow from the expressions $\mathbb{E}_n^x\big[\sin\left(k\pi\frac{X_{n+1}^{\pi}-a}{b-a}\right)\big] = \Im\{\Phi(k|x)\}$, $\mathbb{E}_n^x\big[\cos\left(k\pi\frac{X_{n+1}^{\pi}-a}{b-a}\right)\big] = \Re\{\Phi(k|x)\}$, where $\Phi(k|x)$ is defined as in Section 5.1.

Multi-dimensional extensions. In case the underlying forward process is an $\mathbb{R}^{d\times 1}$-dimensional Brownian motion, the following extension can be given. Let $v : [0,T]\times\mathbb{R}^{d\times 1} \to \mathbb{R}$ be a scalar-valued. Then reasoning similar to Ruijter & Oosterlee (2015, A.1) shows that $\mathbb{E}_n\left[\left(\Delta W_n\right)_{i1} v(t_{n+1}, X_{n+1}^{\pi})\right] = \sum_{k=1}^{d}\Delta t_n\mathbb{E}_n\left[\partial_k v(t_{n+1}, X_{n+1}^{\pi})\right]\left(\sigma(t_n, X_n^{\pi})\right)_{ki}$. In matrix notation

$$\left(\mathbb{E}_n\left[\Delta W_n v(t_{n+1}, X_{n+1}^{\pi})\right]\right)^T = \Delta t_n\mathbb{E}_n\left[\nabla_x v(t_{n+1}, X_{n+1}^{\pi})\right]\sigma(t_n, X_n^{\pi}). \tag{B.4}$$

Alternatively, for a vector-valued mapping $\psi : [0,T]\times\mathbb{R}^{d\times 1} \to \mathbb{R}^{1\times d}$, similar arguments give the following, component-wise formula $\mathbb{E}_n\big[\left(\Delta W_n\right)_{i1}\left(\psi\left(t_{n+1}, X_{n+1}\right)\right)_{1j}\big] = \sum_{k=1}^{d}\Delta t_n\mathbb{E}_n\big[\partial_k\left(\psi\left(t_{n+1}, X_{n+1}^{\pi}\right)\right)_{1j}\big]\left(\sigma(t_n, X_n^{\pi})\right)_{ki}$. In matrix notation

$$\left(\mathbb{E}_n\left[\Delta W_n\psi\left(t_{n+1}, X_{n+1}^{\pi}\right)\right]\right)^T = \Delta t_n\mathbb{E}_n\left[\nabla_x\psi\left(t_{n+1}, X_{n+1}^{\pi}\right)\right]\sigma\left(t_n, X_n^{\pi}\right), \tag{B.5}$$

where $\nabla_x\psi$ is the Jacobian matrix of ψ.

Appendix C. BCOS estimates

Let us fix $d = 1$. The BCOS approximations of the OSM scheme in (5.7) can be derived as follows. Using the definition in (5.8) and the Euler–Maruyama estimates in (3.5), the Γ estimates in (5.1b) can be written according to

$$\begin{aligned} D_n\breve{Z}_n^\pi = \breve{\gamma}_n^\pi(x)\sigma(t_n,x) = {} & \frac{1}{\Delta t_n}\sigma(t_n,x)\left(1+\Delta t_n\partial_x\mu(t_n,x)\right)\mathbb{E}_n^x\left[\Delta W_n w_{n+1}^\pi\left(\widehat{\mathbf{X}}_{n+1}^\pi\right)\right] \\ & + \frac{1}{\Delta t_n}\sigma(t_n,x)\partial_x\sigma(t_n,x)\mathbb{E}_n^x\left[\Delta W_n^2 w_{n+1}^\pi\left(\widehat{\mathbf{X}}_{n+1}^\pi\right)\right] \\ & + \mathbb{E}_n^x\left[\Delta W_n\partial_z f\left(t_{n+1},\widehat{\mathbf{X}}_{n+1}^\pi\right)\right]\breve{\gamma}_n^\pi(x)\sigma(t_n,x). \end{aligned} \tag{C.1}$$

A cosine expansion approximation for $w_{n+1}^\pi(\widehat{\mathbf{X}}_{n+1}^\pi)$ and $\partial_z f(t_{n+1},\widehat{\mathbf{X}}_{n+1}^\pi)$ can be obtained by means of DCT, yielding approximations $\{\widehat{\mathcal{W}}_k(t_{n+1})\}_{k=0,\dots,K-1}$, $\{\widehat{\mathcal{F}}_k^z(t_{n+1})\}_{k=0,\dots,K-1}$, respectively. Consequently, plugging these approximations combined with the integration by parts formulas in (5.5)–(5.6) in the estimate above yields

$$\begin{aligned} \widehat{\gamma}_n^\pi(x)\sigma(t_n,x) = {} & -\sigma^2(t_n,x)(1+\partial_x\mu(t_n,x)\Delta t_n)\sum_{k=0}^{K-1}{}'\frac{k\pi}{b-a}\widehat{\mathcal{W}}_k(t_{n+1})\Im\{\Phi(k|x)\} \\ & + \sigma(t_n,x)\partial_x\sigma(t_n,x)\sum_{k=0}^{K-1}{}'\widehat{\mathcal{W}}_k(t_{n+1})\Re\{\Phi(k|x)\} \\ & - \Delta t_n\sigma^3(t_n,x)\partial_x\sigma(t_n,x)\sum_{k=0}^{K-1}{}'\left(\frac{k\pi}{b-a}\right)^2\widehat{\mathcal{W}}_k(t_{n+1})\Re\{\Phi(k|x)\} \\ & - \widehat{\gamma}_n^\pi(x)\Delta t_n\sigma^2(t_n,x)\sum_{k=0}^{K-1}{}'\frac{k\pi}{b-a}\widehat{\mathcal{F}}_k^z(t_{n+1})\Im\{\Phi(k|x)\}. \end{aligned} \tag{C.2}$$

The approximation $D_n\widehat{Z}_n^\pi = \widehat{\Gamma}_n^\pi\sigma\left(t_n,X_n^\pi\right)$ subsequently follows. The coefficients $\{\widehat{\mathcal{DZ}}_k(t_{n+1})\}_{k=0,\dots,K-1}$ are calculated by DCT and subsequently plugged into the approximations of the Z process, which follows analogously using the formulas in (5.4)–(5.5). The approximation of the Y process in (5.1d) is identical to Ruijter & Oosterlee (2015) and therefore omitted.