THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

Convergence and qualitative properties of modified explicit schemes for BSDEs with polynomial growth

OPEN ACCESS

CONVERGENCE AND QUALITATIVE PROPERTIES OF MODIFIED EXPLICIT SCHEMES FOR BSDES WITH POLYNOMIAL GROWTH

BY ARNAUD LIONNET[*,1], GONCALO DOS REIS[†,2] AND LUKASZ SZPRUCH[†]

INRIA Paris—Ecole Nationale des Ponts et Chaussées and
University of Edinburgh[†]

The theory of Forward–Backward Stochastic Differential Equations (FBSDEs) paves a way to probabilistic numerical methods for nonlinear parabolic PDEs. The majority of the results on the numerical methods for FBSDEs relies on the global Lipschitz assumption, which is not satisfied for a number of important cases such as the Fisher–KPP or the FitzHugh–Nagumo equations. Furthermore, it has been shown in [*Ann. Appl. Probab.* **25** (2015) 2563–2625] that for BSDEs with monotone drivers having polynomial growth in the primary variable y, only the (sufficiently) implicit schemes converge. But these require an additional computational effort compared to explicit schemes.

This article develops a general framework that allows the analysis, in a systematic fashion, of the integrability properties, convergence and qualitative properties (e.g., comparison theorem) for whole families of modified explicit schemes. The framework yields the convergence of some modified explicit scheme with the same rate as implicit schemes and with the computational cost of the standard explicit scheme.

To illustrate our theory, we present several classes of easily implementable modified explicit schemes that can computationally outperform the implicit one and preserve the qualitative properties of the solution to the BSDE. These classes fit into our developed framework and are tested in computational experiments.

1. Introduction. Since the initial papers of Zhang [22] and Bouchard and Touzi [2], an important literature has developed concerning the methods for approximating numerically the solution to a nonlinear backward stochastic differential equation (BSDE thereafter). The importance of BSDEs with nonlinear drivers is due to their frequent use in stochastic control problems, and to their deep connection with parabolic partial differential equations (PDEs), which are used to de-

scribe many biological and physical phenomena as well as the solution to many decision problems. Indeed, the solution v to the PDE

$$\left(\partial_t v + \frac{1}{2}(\sigma\sigma^*) \cdot v_{xx} + b \cdot v_x + f(\cdot, v, v_x\sigma)\right)(t, x) = 0 \qquad \text{with } v(T, x) = g(x),$$

is given by solving, for $(s, x) \in [0, T] \times \mathbb{R}^d$, for $t \in [t, T]$, the forward-backward SDE,

$$(1.1) \qquad X_t^{s,x} = x + \int_s^t b(r, X_r^{s,x}) \, dr + \int_s^t \sigma(r, X_r^{s,x}) \, dW_r,$$

$$(1.2) \qquad Y_t^{s,x} = g(X_T^{s,x}) + \int_t^T f(r, Y_r^{s,x}, Z_r^{s,x}) \, dr - \int_t^T Z_r^{t,x} \, dW_r,$$

and setting $v(s, x) = Y_s^{s,x}$ (see, e.g., [11]). In solving (1.2), one seeks a pair of processes $\mathcal{S} = (Y, Z)$, adapted to the filtration $\mathcal{F}$ of the Brownian motion W. The data of the BSDE are the $\mathcal{F}_T$-measurable random variable $\xi = g(X_T^{s,x})$, called terminal condition, and the function f, generally referred to as driver and which will depend only on time, Y and Z for the simplicity of exposition.

Consider a discretization of the time interval $[0, T]$ by a regular subdivision π^N: $0 = t_0 < t_1 < \cdots < t_N = T$, where $t_i = ih$ for all $i \in \{0, \ldots, N\}$ and $h = T/N$. To construct a numerical methods for BSDEs, one typically begins by discretizing the time dynamics for Y over $[t_i, t_{i+1}]$ as

$$(1.3) \qquad Y_i^N = \mathbb{E}_i[Y_{i+1}^N + (1 - \theta)f(t_i, Y_{i+1}^N, Z_i^N)h] + \theta f(t_i, Y_i^N, Z_i^N)h,$$

where $\mathbb{E}_i$ is $\mathbb{E}[\cdot|\mathcal{F}_{t_i}]$ and the approximation Z_i^N is suitably computed. When the parameter $\theta = 0$, this is the explicit scheme while $\theta = 1$ is the implicit scheme,[3] both dubbed BTZ schemes in [5]. Such a time-discretization scheme is first initialized by setting Y_N^N to be a numerical approximation ξ^N of the terminal condition ξ and then applied in a backward recursive fashion, producing a family $(Y_i^N, Z_i^N)_{i=0,\ldots,N}$ approximating the solution $(Y_t, Z_t)_{t \in [0,T]}$ of the BSDE.

There has been a significant progress in the analysis of variants of scheme (1.3) or the ways to approximation the conditional expectations $\mathbb{E}_i$, although the vast majority of works impose a restrictive Lipschitz condition on driver f in both its Y and Z variables (see [1, 4, 10, 12] and references therein).

However, in many cases of interest, the driver is not Lipschitz but instead has superlinear growth in one of its variables. For instance, for PDEs of reaction-diffusion type such as the Allen–Cahn equation, the FitzHugh–Nagumo equations, the Fisher–KPP equation or the standard nonlinear heat and Schrödinger equation (see [13, 20] and references), the function f is a polynomial in Y. Meanwhile,

---

[3]Numerical schemes most often compute first Z_i^N explicitly from the input Y_{i+1}^N, and then use this to compute Y_i^N. In this paper, all mentions of implicit and explicit scheme in the context of BSDEs refer the to Y-component.

in stochastic control problems, the driver typically has quadratic growth in the Z variable.

Chassagneux and Richou obtained in [8] the convergence of an implicit scheme in the case where the terminal condition ξ is bounded and the driver f has quadratic growth in Z. In [19], we studied the case where the terminal condition has all moments, the driver has polynomial growth in Y and satisfies a so-called monotonicity condition (also known as one-sided Lipschitz condition). The monotonicity condition is a structure property which states, in the scalar case, that for all y, y' in the domain $\mathbb{R}$ and for all z,

$$(1.4) \qquad \big(f(y, z) - f(y', z)\big)(y' - y) \le M_y |y' - y|^2,$$

where $M_y \in \mathbb{R}$. For instance, the driver $f(y, z) = y - y^3$ typical of the FitzHugh–Nagumo equation is one-sided Lipschitz over the domain $\mathbb{R}$ with $M_y = 1$.

As explained in [19], the explicit scheme described in (1.3) can explode. This is due to the superlinear growth of the driver f and the unboundedness of the terminal condition ξ. As a remedy, [19] proposed to use the implicit scheme, which was shown to converge, or an explicit scheme with a truncated numerical terminal condition $T^N(\xi^N)$, where the truncation function T^N fades to the identity when the number N of time-steps goes to $+\infty$. However, the implicit scheme requires an extra computational effort to solve the nonlinear equation (1.3) defining Y_i^N when $\theta = 1$. And for the explicit scheme with the truncated terminal conditions, severe restrictions on the size of the time-step had to be imposed, and the tuning/performance of the algorithm depends on knowledge of f, in particular on its growth. The purpose of this paper is to obtain converging explicit schemes by working instead on the dynamics of the scheme itself, replacing the driver f by a modified driver f^h, with no time-step restriction. In addition, we can obtain some numerical schemes of black-box type, where no a priori knowledge on the structure of the driver is required.

Explosion problems of naïve explicit schemes were already stressed in [16] in the context of the numerical methods for stochastic differential equations (SDEs thereafter). A significant body of works considered various modifications to the explicit schemes for SDEs, dubbed "tamed schemes," to recover integrability and convergence in the non-Lipschitz setting; see [6, 14–18, 21]. In the context of BSDEs, very particular instances of modified drivers were already used by [8] to deal with the quadratic growth in Z, and in [19] to deal with the dependence of the driver of (1.2) in the solution X to the SDE (1.1). But this was only used as an ad hoc tool to handle a particular issue.

This motivates us to study systematically the family of modified explicit schemes where the BSDE driver f is replaced by a "tamed" driver f^h. Provided the modified driver f^h is appropriately "tamed," the explosion of the scheme is prevented. One then expects that such a modified scheme will converge to the continuous-time solution provided $f^h \to f$. In addition, if this convergence can

happen fast enough, the usual convergence rate of the implicit schemes can be recovered. Our approach in this work is to identify the essential properties of these modified drivers f^h which guarantee the convergence of the corresponding modified explicit scheme for BSDEs. As a consequence, we show at once the convergence of a whole range of modified explicit BSDE schemes.

In a certain sense, this question is reminiscent of the stability with respect to the driver for solutions of continuous-time BSDEs. Given the fixed set of times $[0, T]$, let us denote by $\mathcal{S}^f$ the solution to the BSDE with driver f and by $\mathcal{S}^{f^\varepsilon}$ the solution for the driver f^ε, such that f^ε converges to f in some sense (uniformly on compact, typically). The stability theorem, valid for monotone drivers as well as for Lipschitz drivers, states that $\mathcal{S}^{f^\varepsilon}$ converges to $\mathcal{S}^f$, in the appropriate norm, and gives an upper bound on the distance. Here, for the set of times π^N (which "converges to $[0, T]$" as $N \to +\infty$), let us denote by $S^{N, f^h} = (Y_i^{N, f^h}, Z_i^{N, f^h})_{i=0...N}$ the output of the modified explicit scheme with drivers f^h, and by $S^{N, f}$ the output of the standard (BTZ) explicit scheme, with driver f. One may then be tempted to say that, as f^h converges to f when $h = T/N \to 0$, S^{N, f^h} should be close to $S^{N, f}$ for large N. However, the BTZ explicit scheme $S^{N, f}$, in general, does not converge to $\mathcal{S}^f$ when f is not Lipschitz, while in this paper it is proved that S^{N, f^h} converges to $\mathcal{S}^f$. Therefore, convergence of S^{N, f^h} to $S^{N, f}$ cannot hold (at least not uniformly over π^N: for fixed π^N and $f^\varepsilon \to f$, one should have $S^{N, f^\varepsilon} \to S^{N, f}$). We depict the situation on Figure 1 where by (!) we marked the convergence results that, in general, do not hold in the non-Lipschitz setting.

One could nonetheless say that the BSDE solution $\mathcal{S}^f$ with driver f and the BSDE solution $\mathcal{S}^{f^h}$ with driver f^h must be close by the stability result for continuous time BSDEs. This depicted on the Figure 1 by the arrow (b). Then, if f^h is Lipschitz for each N (albeit with Lipschitz and growth constants which explode when $h \to 0$), the standard estimate for the convergence of the explicit BTZ scheme with Lipschitz drivers should allow to bound the distance between $\mathcal{S}^{f^h}$ and S^{N, f^h}. This depicted on the Figure 1 by the arrow (a). However, this strategy will not work since the upper bound in that estimate involves (the exponential of) a constant that grows with the regularity and growth constants of f^h.

$$
\begin{array}{ccc}
S^{N, f^h} & \xrightarrow{\ \ (!)\ \ } & S^{N, f} \\
\left\downarrow{\scriptstyle (a)}\vphantom{\Big|}\right. & \searrow\!\!\!\searrow_{(b)} & \left\downarrow{\scriptstyle (!)}\vphantom{\Big|}\right. \\
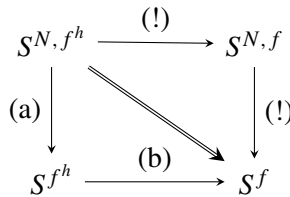S^{f^h} & \xrightarrow[\ \ (b)\ \ ]{} & S^f
\end{array}
$$

FIG. 1. *Schematic diagram of key findings of this paper. Double arrow marks the conclusion of Theorem* 2.4. (!) *marks convergence that in general does not hold.*

One must therefore carefully analyze the modified explicit schemes and work with modifications f^h that, on one hand, allow to tame the superlinear growth and, on the other hand, preserve certain structural properties of f such as the monotonicity condition. In addition, we manage to show our convergence $S^{N,f^h} \to \mathcal{S}^f$ for modified drivers f^h that are not necessarily Lipschitz, but only almost-Lipschitz and of linear growth [cf. the assumptions **(TRegY)** and **(TGrowth)** below, when $\mathcal{R}^{(\text{regY})} \neq 0$].

Taming the driver does not merely allow to recover the convergence of the explicit scheme for BSDEs in the regime $h \to 0$. It also makes it more robust and qualitatively more satisfying for finite h. An important qualitative property is the comparison theorem, which plays a major role in the study of continuous-time BSDEs. In the context of numerical schemes for BSDEs with driver that is quadratic in Z, this property was obtained in [8] and used to prove the convergence of their scheme. In this paper, we show in the case of drivers with superlinear growth in Y, that a comparison theorem holds for suitably modified explicit schemes. This qualitative property is of interest for its own sake, but also allows to deduce that the scheme remains in some domain D of the space when the continuous-time solution does. This is important if the monotonicity condition for f is only satisfied on D, as is the case for instance for the Fisher–KPP equation where the driver is $f(y, z) = y - y^2$. Here, the monotonicity condition for f is satisfied on the domain $D = [0, +\infty[$, and the solution stays positive. To the best of our knowledge, this is the first result on a comparison theorem for explicit schemes for BSDEs, even for a driver f that is Lipschitz in both its Y and Z variables.

The paper is organized as follows. In Section 2, we state the assumptions under which we work, the numerical schemes considered and our main results. We give in Section 2.4 several examples of driver modifications that fit our assumptions. In Section 2.5, we study the discrete comparison and the preservation of positivity (of the explicit schemes); and Section 2.6 illustrates with some numerical examples of tamed drivers the results we obtained.

Sections 3 and 4 are concerned with proving the convergence of the scheme. Specifically, in Section 3, we show how taming the driver prevents the scheme from exploding, deriving almost-sure local and then global bounds, which lead to moment estimates. In Section 4, we then prove that the scheme converges. For this, we first prove that the scheme is almost L^2-stable over one time-step, which allows to control the propagation of errors. We then analyze the discretization error created at each time-step, and show that their sum converges with rate at least $1/2$.

The paper closes with two appendices. Appendix A provides proofs regarding our main results, while Appendix B provides proofs for the examples studied in Section 2.4.

2. Assumptions, schemes and main results. *Notation.* Fix $T > 0$. We work on a canonical Wiener space $(\Omega, \mathcal{F}, \mathbb{P})$ carrying a d-dimensional Wiener process $W = (W^1, \ldots, W^d)$ restricted to the time interval $[0, T]$. We denote by

$\mathcal{F} = (\mathcal{F}_t)_{t \in [0,T]}$ its natural filtration enlarged in the usual way by the $\mathbb{P}$-zero sets and by $\mathbb{E}$ and $\mathbb{E}[\cdot|\mathcal{F}_t] = \mathbb{E}_t[\cdot]$ the usual expectation and conditional expectation operator respectively.

We denote by $\langle \cdot, \cdot \rangle$ and $|\cdot|$ the canonical inner product and Euclidean norm on $\mathbb{R}^d$, and by x^* the transposed of $x \in \mathbb{R}^d$ when seen as a $\mathbb{R}^{d \times 1}$ matrix (column-vector). I_d denotes the d-dimensional identity matrix. $L^p = L^p(\mathcal{F}_T, \mathbb{R}^d)$ is the space of $\mathbb{R}^d$-valued $\mathcal{F}_T$-measurable random variables X with norm $\|X\|_{L^p} = \mathbb{E}[|X|^p]^{1/p} < \infty$. $\mathcal{S}^p$ is the space of d-dimensional $\mathcal{F}$-adapted processes Y satisfying $\|Y\|_{\mathcal{S}^p} = \mathbb{E}[\sup_{t \in [0,T]} |Y_t|^p]^{1/p} < \infty$. $\mathcal{H}^p$ is the space of d-dimensional $\mathcal{F}$-adapted processes Z satisfying $\|Z\|_{\mathcal{H}^p} = \mathbb{E}[(\int_0^T |Z_s|^2 \, ds)^{p/2}]^{1/p} < \infty$.

2.1. *Assumptions on the continuous-time dynamics.* We make the following running assumptions.

The SDE and the terminal condition. We assume that the functions $b : [0, T] \times \mathbb{R}^d \to \mathbb{R}^d$, $\sigma : [0, T] \times \mathbb{R}^d \to \mathbb{R}^{d \times d}$ in (1.1) are 1/2-Hölder continuous in their time variable, Lipschitz continuous and of linear growth in their spatial variables. The terminal condition $g : \mathbb{R}^d \to \mathbb{R}^n$ is a Lipschitz function. The terminal condition $\xi := g(X_T) \in L^p$ for $p \geq 1$.

The driver of the BSDE. We work with drivers $f : [0, T] \times \mathbb{R}^n \times \mathbb{R}^{n \times d} \to \mathbb{R}^n$ having polynomial growth in y and satisfying the so-called monotonicity condition (also known as one-sided Lipschitz condition), while being Lipschitz functions of z. Specifically, our drivers f satisfy the growth, monotonicity and regularity conditions stated below. We choose not to include the possibility of a Lipschitz dependence of f on the variable x. It is seen in [19] that this can easily be dealt with. For clarity of our results, we exclude it here.

(Growth) There exist $m \in \mathbb{N}^*$ and $K_t, K_y, K_z \geq 0$ such that, for all $(t, y, z) \in [0, T] \times \mathbb{R}^n \times \mathbb{R}^{n \times d}$,

$$\left| f(t, y, z) \right| \leq K_t + K_y |y|^m + K_z |z|.$$

That is, f has polynomial growth in y of degree m and linear growth in z.

(Mon) There exist a constant $M_y \in \mathbb{R}$ such that for all t, y, y', z,

$$\langle y' - y, f(t, y', z) - f(t, y, z) \rangle \leq M_y |y' - y|^2.$$

That is, f is monotone ("decreasing") in the variable y. The monotonicity constant M_y can be, but is not necessarily, strictly negative.

(Reg) There exist constants $L_t, L_z \geq 0$ such that for all t, t', y, z',

$$\left| f(t', y, z') - f(t, y, z) \right| \leq L_t |t' - t|^{\frac{1}{2}} + L_z |z' - z|.$$

That is, f is $\frac{1}{2}$-Hölder in time and Lipschitz in z.

In a number of places, we need to know about the regularity of f in the variable y [notice that **(Mon)** does not even imply continuity]. We assume that f satisfies the following.

(RegY) There exists a constant $L_y \geq 0$ such that for all t, y, y', z,

$$|f(t, y', z) - f(t, y, z)| \leq L_y(1 + |y'|^{m-1} + |y|^{m-1})|y' - y|.$$

That is, f is locally Lipschitz in y with local Lipschitz constant growing polynomially with degree $m - 1$.

We are primarily interested in drivers f that are polynomials in y, and for these we see that **(RegY)** is clearly satisfied indeed. Finally, we introduce the following monotone growth assumption.

(MonGr) There exist constants $\bar{M}_t, \bar{M}_z \geq 0$ and $\bar{M}_y \in \mathbb{R}$ such that for all t, y, z,

$$\langle y, f(t, y, z) \rangle \leq \bar{M}_t + \bar{M}_y|y|^2 + \bar{M}_z|z|^2.$$

REMARK 2.1. **(MonGr)** is a direct consequence of **(Mon)** and **(Growth)** and $\bar{M}_z$ (as well as $\bar{M}_t$) can be chosen arbitrarily small, as proved below. We single out this property because it controls the growth of the driver and therefore the integrability of the solution Y. When running the scheme with a tamed driver f^h, we only need assumptions similar to **(Growth)** and **(MonGr)** to guarantee moment bounds for the scheme, and thereby nonexplosion.

Let f satisfies **(Mon)** and **(Growth)**. For all t, y, z and for any $\alpha > 0$, we have

$$\langle y, f(t, y, z) \rangle = \langle y - 0, f(t, y, z) - f(t, 0, z) \rangle + \langle y, f(t, 0, z) \rangle$$

$$\leq M_y|y - 0|^2 + |y|(K_t + K_z|z|) \leq (M_y + \alpha)|y|^2 + \frac{K_t^2}{2\alpha} + \frac{K_z^2}{2\alpha}|z|^2.$$

Hence, we can take $\bar{M}_t = \frac{K_t^2}{2\alpha}$ and $\bar{M}_z = \frac{K_z^2}{2\alpha}|z|^2$ arbitrarily small, while taking $\bar{M}_y = M_y + \alpha$. We also note that by combining **(Mon)** and **(Reg)** we obtain the general estimate

$$\langle y' - y, f(t, y', z') - f(t, y, z) \rangle = \langle y' - y, f(t, y', z') - f(t, y, z') \rangle$$

$$+ \langle y' - y, f(t, y, z') - f(t, y, z) \rangle$$

$$\leq (M_y + \alpha)|y' - y|^2 + \frac{L_z^2}{4\alpha}|z' - z|^2.$$

Results from BSDE theory. The assumptions **(Growth)**, **(Mon)**, **(Reg)** and **(RegY)** are a more detailed version of assumptions (HY0) and (HY0$_{loc}$) in [19], Section 2.2, and they imply the fundamental BSDE results of Sections 2 and 3 of [19]. Essentially, those results are the existence and uniqueness of the solution, a priori bound estimates and the path-regularity theorem. We recall them in Appendix A.1. Throughout we denote by $(Y_t, Z_t)_{t \in [0,T]}$ the unique solution to (1.2), which we aim at approximating numerically.

2.2. *Time-discretization*: *Dynamics and assumptions.* We discretize the time-interval $[0, T]$ using a partition $\pi : 0 = t_0 < t_1 < \cdots < t_{N-1} < t_N = T$ with N intervals. The modulus of the partition is $|\pi| = \max_{i=0,\dots,N-1} h_{i+1}$ where $h_{i+1} = t_{i+1} - t_i$. While our results would hold for more general partitions, we restrict ourselves to regular partitions for notational simplicity. Consequently, given the number N of time-intervals, we work with the partition π^N where $t_i = ih$, $h = T/N \in (0, T]$ being the modulus of the partition.

We wish to focus on the numerical approximation of the backward SDE (1.2). So we do not discuss the numerical approximation of the forward SDE (1.1) and that of the terminal condition. We work with the following assumption regarding the numerical approximation ξ^N of the terminal condition $\xi = g(X_T)$.

(Aξ^N) For any $p \geq 2$, $\xi^N \in L^p(\mathcal{F}_T)$. There exists a constant c (independent of N) such that

$$\mathrm{ERR}_h(\xi) := \mathbb{E}[|\xi - \xi^N|^2]^{\frac{1}{2}} \leq c h^{\frac{1}{2}}.$$

Given the assumptions made on b, σ and g, one can use the standard Euler scheme for SDEs to produce an approximation $X^N = (X_i^N)_{i=0,\dots,N}$ of X and set $\xi^N = g(X_N^N)$. This ξ^N satisfies **(Aξ^N)**.

2.2.1. *The modified explicit schemes.* For $i \in \{0, \dots, N-1\}$, we denote the Brownian increments by $\Delta W_{t_{i+1}} := W_{t_{i+1}} - W_{t_i}$. We also take a family of $\mathbb{R}^d$-valued random variables $(H_{i+1})_{i=0,\dots,N-1}$ which approximate $\Delta W_{t_{i+1}}/h$ and satisfy the following assumption:

(AH)   1. For any $i = 0, \dots, N-1$, H_{i+1} is independent from $\mathcal{F}_i$ and satisfies $\mathbb{E}_i[H_{i+1}] = 0$ (H_{i+1} is a martingale increment).

2. For any $i = 0, \dots, N-1$, $\mathbb{E}_i[(H_{i+1}h)(H_{i+1}h)^*] = \Lambda h I_d$, where $\frac{1}{2} \leq \Lambda \leq 1$. As a consequence, $\mathbb{E}_i[|H_{i+1}h|^2] = \Lambda dh$.

3. There exists a constant $C \geq 0$ (independent of N) such that

$$\max_{i=0,\dots,N-1} \mathbb{E}\left[\left|\frac{\Delta W_{t_{i+1}}}{h} - H_{i+1}\right|^2\right] \leq C.$$

In works on numerical methods for BSDEs, H_{i+1} is often defined as $\Delta W_{t_{i+1}}/h$ (in which case $\Lambda = 1$). However, as will be seen in Section 2.5, in order to have an explicit scheme which is numerically stable (i.e., reproduces qualitative properties of the continuous-time BSDE, such as the positivity of the solution), the H_{i+1} are required to be bounded. One way to do this is to truncate the Brownian increment $\Delta W_{t_{i+1}}$ to $\Delta W_{t_{i+1}}^h$, by projecting it on the centered ball of radius R^h, where $R^h \to +\infty$ as $h \to 0$. In that case, $H_{i+1} = \Delta W_{t_{i+1}}^h/h$. Also, working with H_{i+1} rather than $\Delta W_{t_{i+1}}/h$ also allows to include in the analysis tree-based methods such as cubature [10].

We work with the following scheme. It is initialized with $Y_N = \xi^N$ (and $Z_N = 0$). Then, for $i = N - 1$ to 0, the *output of one step of the scheme when the input is Y_{i+1} is $(Y_i, Z_i) = S_i(Y_{i+1})$* defined by

$$(2.1) \quad \begin{cases} Y_i = \mathbb{E}_i[Y_{i+1} + f^h(t_i, Y_{i+1}, Z_i)h], \\ Z_i = \mathbb{E}_i[(Y_{i+1} + (1 - \theta')f^h(t_i, Y_{i+1}, 0)h)H_{i+1}^*], \end{cases}$$

where $\theta' \in [0, 1]$ and the driver f^h is a modification of f. The precise assumptions on f^h are described later. The global output of the scheme is the sequence of random variables $S^{N, f^h} = ((Y_i, Z_i))_{i=0,\ldots,N-1}$ valued in $\mathbb{R}^n \times \mathbb{R}^{n \times d}$. The superscript N is omitted since the discrete subscript $i \in \{0, \ldots, N\}$ already indicates that Y_i (say) refers to the numerical approximation, while Y_{t_i} is simply the solution to the continuous-time BSDE at time t_i.

This (explicit) scheme corresponds to the case where the parameter $\theta = 0$ in (1.3), θ' being another parameter here. Most schemes for BSDEs proposed in the literature choose $\theta' = 1$. We show in this work that the scheme converges for any $\theta' \in [0, 1]$. However, a reader familiar with the analysis of continuous-time BSDEs will easily see in Sections 3 and 4 that having $\theta' = \theta$ (so, $\theta' = 0$) allows to analyze the scheme (2.1) by mimicking more closely the continuous-time analysis, as suggested below.

Discrete-time martingale representation. While Y_i is defined as a conditional expectation in (2.1), it is useful to rewrite it using a martingale increment. Note that such a representation was already used in [7, 8] and [3], although the way we use it in the estimates of Sections 3 and 4 differs.

LEMMA 2.2. *Given $i \in \{0, \ldots, N - 1\}$, consider a $\mathcal{F}_{i+1}$-measurable $\mathbb{R}^n$-valued random variable $\mathcal{Y}_{i+1}$ as well as a $\mathcal{F}_i$-measurable $\mathbb{R}^{n \times d}$-valued random variable $\mathcal{Z}_i$, such that $\mathcal{Y}_{i+1} + f^h(t_i, \mathcal{Y}_{i+1}, \mathcal{Z}_i)h \in L^1$. Define $\mathcal{Y}_i = \mathbb{E}_i[\mathcal{Y}_{i+1} + f^h(t_i, \mathcal{Y}_{i+1}, \mathcal{Z}_i)h]$. Then $\mathcal{Y}_i$ can be written*

$$(2.2) \quad \mathcal{Y}_i = \mathcal{Y}_{i+1} + f^h(t_i, \mathcal{Y}_{i+1}, \mathcal{Z}_i)h - \Delta M_{i+1},$$

where $\mathbb{E}_i[\Delta M_{i+1}] = 0$. Moreover, there exists a unique pair $(\zeta_i, \Delta N_{i+1})$, with ζ_i a $\mathcal{F}_i$-measurable $\mathbb{R}^{n \times d}$-valued random variable and ΔN_{i+1} a martingale increment orthogonal to $H_{i+1}h$, such that

$$(2.3) \quad \Delta M_{i+1} = \zeta_i \Lambda^{-1} H_{i+1} h + \Delta N_{i+1}.$$

ζ_i is given by

$$(2.4) \quad \zeta_i = \mathbb{E}_i[(\mathcal{Y}_{i+1} + f^h(t_i, \mathcal{Y}_{i+1}, \mathcal{Z}_i)h)H_{i+1}^*].$$

The proof of this statement is relatively straightforward. For completeness, it is is presented in Appendix A.2.

Derivation of the scheme. Lemma 2.2 suggests that, if one has already chosen a time-discretization for the process (Y_t) over $[t_i, t_{i+1}]$ given by

$$(2.5) \qquad \check{Y}_i = \mathbb{E}_i[\check{Y}_{i+1} + f^h(t_i, \check{Y}_{i+1}, \check{Z}_i)h],$$

where the input $\check{Y}_{i+1}$ already approximates $Y_{t_{i+1}}$, and $\check{Z}_i$ is yet to-be-determined, it is more natural to choose subsequently

$$(2.6) \qquad \begin{aligned} \check{Z}_i &= \mathbb{E}_i\big[(\check{Y}_{i+1} + f^h(t_i, \check{Y}_{i+1}, \check{Z}_i)h)H_{i+1}^*\big] \\ &= \mathbb{E}_i\big[(\check{Y}_{i+1} + (1-\theta')f^h(t_i, \check{Y}_{i+1}, \check{Z}_i)h)H_{i+1}^*\big], \end{aligned}$$

with $\theta' = 0$. Indeed, since "$\check{Z}_i = \zeta_i$," this means that $(\check{Y}_i, \check{Z}_i) =: \check{S}_i(\check{Y}_{i+1})$ is defined by

$$\check{Y}_i = \check{Y}_{i+1} + f^h(t_i, \check{Y}_{i+1}, \check{Z}_i)h - (\check{Z}_i\Lambda^{-1}H_{i+1}h + \Delta N_{i+1}).$$

Due to the resemblance between the above equation and the continuous-time BSDE, this scheme has the advantage that the analysis carried out in the later Sections 3 and 4 is more natural to the reader accustomed with the continuous-time analysis. The above explains why it would be natural, and later convenient for the analysis of the scheme, to take $\theta' = \theta$ (hence $\theta' = 0$). But due to the prevalence and convenience of implementation of the choice $\theta' = 1$, we study the scheme for a general $\theta' \in [0, 1]$, thus covering both cases.

However, in the scheme (2.5)–(2.6), $\check{Z}_i$ is defined implicitly (unless $\theta' = 1$), which leads to solving a nonlinear equation to compute $\check{Z}_i$ at each step of the scheme. In theory, this could be dealt with, seeing as $z \mapsto f^h(t_i, \check{Y}_{i+1}, z)h$ is Lipschitz and, therefore, a contraction for h small enough, so $\check{Z}_i$ could be approximated by iteration. But this defeats the purpose of this work, as we aim at fully explicit schemes. Also, we want to avoid imposing restrictions on the size of the time-steps. So we replace $f^h(t_i, \check{Y}_{i+1}, \check{Z}_i)h$ by $f^h(t_i, \check{Y}_{i+1}, 0)h$ in our scheme (2.1). Given the Lipschitz dependence of f in z, this creates an error that should not affect the convergence rate of the scheme, which we aim to be $\frac{1}{2}$, like for the implicit scheme.

To be able to take advantage of Lemma 2.2 and the natural character of scheme (2.5)–(2.6) when analyzing scheme (2.1), we introduce the random variable $D_i := Z_i - \zeta_i$, where ζ_i is given by (2.4) with $\mathcal{Y}_{i+1} = Y_{i+1}$ and $\mathcal{Z}_i = Z_i$. This D_i measures the difference between the theoretically-natural scheme and the scheme used in practice. The proofs of Lemma 3.1 and Proposition 3.2 are written in such a way that the interested reader can easily track the consequence of $D_i \neq 0$. For the scheme (2.5)–(2.6), the analogue $\check{D}_i := \check{Z}_i - \check{\zeta}_i$, where $\check{\zeta}_i$ is given by (2.4) with $\mathcal{Y}_{i+1} = \check{Y}_{i+1}$ and $\mathcal{Z}_i = \check{Z}_i$, is null.

2.2.2. *Assumptions on the tamed driver and comments.* We now introduce the general assumptions on the tamed driver f^h. They summarize the fact that we

want f^h to enjoy most of the properties of f, in particular to preserve as much as possible the monotonicity of **(Mon)** and **(MonGr)**, but with the polynomial growth of f and its local Lipschitz constant tamed; see **(Growth)** and **(RegY)**.

These abstract assumptions can be split in three categories. First, we have the growth conditions **(TGrowth)** and **(TMonGr)** which ensure the nonexplosion of the scheme. Second, the monotonicity and regularity conditions **(TReg)**, **(TRegY)** and **(TMon)** ensure the stability of the scheme. Finally, **(TCvg)** ensures the convergence of the scheme. Recall that $h \in (0, T]$ and we are interested in doing $h \to 0$.

Assumptions on the growth.

(TGrowth) There exist K_t^h, K_y^h and $K_z^h \geq 0$ such that, for all $(t, y, z) \in [0, T] \times \mathbb{R}^n \times \mathbb{R}^{n \times d}$,

$$\left| f^h(t, y, z) \right| \leq K_t^h + K_y^h |y| + K_z^h |z|.$$

The constants K_t^h, K_y^h and K_z^h may depend on h but in such a way that $(K_t^h)^2 h$, $(K_y^h)^2 h$ and K_z^h are bounded in h, that is, $K_t^h, K_y^h = O(1/\sqrt{h})$. Also, $|f^h(t, y, z)| \leq |f(t, y, z)|$.

(TMonGr) There exist $\bar{M}_t^h, \bar{M}_z^h \geq 0$ and $\bar{M}_y^h \in \mathbb{R}$ such that, for all $(t, y, z) \in [0, T] \times \mathbb{R}^n \times \mathbb{R}^{n \times d}$,

$$\langle y, f(t, y, z) \rangle \leq \bar{M}_t^h + \bar{M}_y^h |y|^2 + \bar{M}_z^h |z|^2.$$

The constants may depend on h and satisfy $\sup_h \{\bar{M}_t^h, \bar{M}_z^h, \bar{M}_y^h\} < \infty$.

Assumptions on the regularity.

(TReg) There exist $L_t^h, L_z^h \geq 0$ such that, for all t, t', y, z, z',

$$\left| f^h(t', y, z') - f^h(t, y, z) \right| \leq L_t^h |t' - t|^{\frac{1}{2}} + L_z^h |z' - z|.$$

The constants may depend on h and satisfy $\sup_h \{L_t^h, L_z^h\} < \infty$.

(TRegY) There exist $L_y^h \geq 0$ and a positive function $\mathcal{R}^{(\text{regY})}$ satisfying **(TCvg)** (defined below; it ensure that $\mathcal{R}^{(\text{regY})}$ goes to 0) such that for all t, y, y', z,

$$\left| f^h(t, y', z) - f^h(t, y, z) \right| \leq L_y^h |y' - y| + \mathcal{R}^{(\text{regY})}(t, y', y, z).$$

L_y^h may depend on h s.th. $(L_y^h)^2 h$ is bounded in h, that is, $L_y^h = O(1/\sqrt{h})$.

(TMon) There exists $M_y^h \in \mathbb{R}$ and a function $\mathcal{R}^{(\text{mon})}$ satisfying **(TCvg)** (defined below; it ensure that $\mathcal{R}^{(\text{mon})}$ goes to 0) such that for all t, y, y', z,

$$\langle y' - y, f^h(t, y', z) - f^h(t, y, z) \rangle \leq M_y^h |y' - y|^2 + \mathcal{R}^{(\text{mon})}(t, y', y, z).$$

M_y^h may depend on h, but in a bounded way s.th. $\sup_h \{M_y^h\} < \infty$.

Assumptions on the convergence. We need to ensure that $f^h \to f$ as $h \to 0$. This is in some sense a *consistency condition*, ensuring that the output of the scheme converges to the solution to the correct BSDE, and not a BSDE with a different driver. We introduce for this $R^h = f - f^h$. Also, we need the remainders $\mathcal{R}^{(\mathrm{reg}Y)}$ and $\mathcal{R}^{(\mathrm{mon})}$ to vanish sufficiently fast, so as not to prevent convergence of the scheme. The following assumption guarantees that R^h, $\mathcal{R}^{(\mathrm{reg}Y)}$ and $\mathcal{R}^{(\mathrm{mon})}$ converge to zero. In its statement, $\mathcal{R}$ stands for both of the remainders $\mathcal{R}^{(\mathrm{reg}Y)}$ and $\mathcal{R}^{(\mathrm{mon})}$.

(TCvg) One of the following holds:

1. There exist constants $C \geq 0$, $p, q \geq 1$ and $\alpha \geq 1/2$ such that for any y', y, z

$$\left| R^h(t, y, z) \right| \leq C \left(1 + |y|^q + |z|^p \right) h^\alpha,$$

$$\mathcal{R}(t, y', y, z) \leq C \left(1 + |y'|^q + |y|^q + |z|^p \right) h^\alpha.$$

2. There exist constants $C \geq 0$, $p, q \geq 1$, $r_0 > 0$ and $\beta > 0$ such that, with $r(h) = r_0 h^{-\beta}$, for any y', y, z

$$\left| R^h(t, y, z) \right| \leq C \left(1 + |y|^q + |z|^q \right) \mathbb{1}_{\{|f(t,y,z)| > r(h)\}},$$

$$\mathcal{R}(t, y', y, z) \leq C \left(1 + |y'|^q + |y|^q + |z|^p \right) \mathbb{1}_{\{|f(t,y',z)| > r(h) \text{ or } |f(t,y,z)| > r(h)\}}.$$

3. There exist constants $C \geq 0$, $p, q \geq 1$, $r_0 > 0$ and $\gamma > 0$ such that, with $r(h) = r_0 h^{-\gamma}$, for any y', y, z

$$\left| R^h(t, y, z) \right| \leq C \left(1 + |y|^q + |z|^p \right) \mathbb{1}_{\{|y| > r(h)\}},$$

$$\mathcal{R}(t, y', y, z) \leq C \left(1 + |y'|^q + |y|^q + |z|^p \right) \mathbb{1}_{\{|y'| > r(h) \text{ or } |y| > r(h)\}}.$$

REMARK 2.3. The above **(TCvg)** implies that $f^h \to f$ as $h \to 0$ pointwise. We would expect from the stability theorems for continuous-time BSDEs that the solutions would converge if $f^h \to f$ uniformly on compacts, which **(TCvg)** also implies. However, in order to obtain convergence rates for the scheme, stronger assumptions are needed.

2.3. *Main result and outline of the proof.* The path-regularity theorem (see Theorem A.2) implies that the distance between the solution $(Y_t, Z_t)_{t \in [0,T]}$ and its projection on the grid, $(Y_{t_i}, \overline{Z}_{t_i})_{i=0,\dots,N-1}$, is of order $h^{1/2}$, where $\overline{Z}_{t_i}$ is defined as

$$\overline{Z}_{t_i} = \mathbb{E}_i \left[\frac{1}{h} \int_{t_i}^{t_{i+1}} Z_u \, du \right] = \mathbb{E}_i \left[\int_{t_i}^{t_{i+1}} Z_u \, dW_u \frac{\Delta W^*_{t_{i+1}}}{h} \right]$$

$$= \mathbb{E}_i \left[\left(Y_{i+1} + \int_{t_i}^{t_{i+1}} f(u, Y_u, Z_u) \, du \right) \frac{\Delta W^*_{t_{i+1}}}{h} \right].$$

We measure the distance between the numerical approximation $(Y_i, Z_i)_{i=0,\ldots,N}$ and the solution of the BSDE with the following error criterion:

$$\mathrm{ERR}_N = \left(\sup_{i=0,\ldots,N} \mathbb{E}\big[|Y_{t_i} - Y_i|^2\big] + \mathbb{E}\left[\sum_{i=0}^{N-1} |\overline{Z}_{t_i} - Z_i|^2 h \right] \right)^{1/2}.$$

Main result. Our principal result, which Sections 3 and 4 are devoted to proving, is that if the driver f is tamed is such a way that the assumption of the previous subsection are satisfied, then the resulting scheme converges. Specifically, we define the rate μ as follows:

- If **(TCvg)**.1 is satisfied, with $q, p \geq 1$ and $\alpha > 0$, then $\mu = \alpha$.
- If **(TCvg)**.2 is satisfied, with $q, p \geq 1$ and $\beta > 0$, then $\mu = \frac{\beta l}{2}$, for arbitrary $l \geq 1$. Since $\beta > 0$, we will take l such that $\mu \geq 1$.
- If **(TCvg)**.3 is satisfied, with $q, p \geq 1$ and $\gamma > 0$, then $\mu = \frac{\gamma l}{2}$, for arbitrary $l \geq 1$. Since $\gamma > 0$, we will take l such that $\mu \geq 1$.

THEOREM 2.4. *Assume that f^h satisfies* **(TGrowth)**, **(TMonGr)**, **(TReg)**, **(TRegY)**, **(TMon)** *and* **(TCvg)**. *Then the scheme* (2.1) *converges, that is,* $\mathrm{ERR}_N \to 0$ *as* $N \to +\infty$.

More precisely, there exists a constant C (independent of N) such that

$$\mathrm{ERR}_N^2 \leq Ch + Ch^\mu.$$

Having $\mu > 0$ guarantees convergence of the scheme. Having $\mu \geq 1$ guarantees that taming the driver does not slow down the convergence and that the standard convergence rate of the implicit scheme can be recovered.

Outline of the proof. We follow a standard strategy which consists in seeing the error at time t_i as resulting from the one-step time-discretization error—by how much the BSDE and the scheme differ over one time-step when initialized with the same input—and the propagation to time t_i of the error already present at time t_{i+1}—the control of this latter error coming from the stability of the scheme.

To express this, we introduce the family of random variables $(\widehat{Y}_i, \widehat{Z}_i)_{i=0,\ldots,N-1}$ defined, for all i, by

$$(2.7) \qquad \widehat{Y}_i = \mathbb{E}_i\big[Y_{t_{i+1}} + f^h(t_i, Y_{t_{i+1}}, \widehat{Z}_i)h\big],$$

$$(2.8) \qquad \widehat{Z}_i = \mathbb{E}_i\big[\big(Y_{t_{i+1}} + (1 - \theta')f^h(t_i, Y_{t_{i+1}}, 0)h\big)H_{i+1}^*\big].$$

Otherwise said, $(\widehat{Y}_i, \widehat{Z}_i)$ is the output of one step of the scheme (2.1) when the input is $Y_{t_{i+1}}$, the value of the solution at the time t_{i+1}. Then the above mentioned decomposition of the error at time t_i can be written as

$$Y_{t_i} - Y_i = \underbrace{Y_{t_i} - \widehat{Y}_i}_{\text{one-step error}} + \underbrace{\widehat{Y}_i - Y_i}_{\text{propagation of error}}$$

and

$$\bar{Z}_{t_i} - Z_i = \underbrace{\bar{Z}_{t_i} - \widehat{Z}_i}_{\text{one-step error}} + \underbrace{\widehat{Z}_i - Z_i}_{\text{propagation of error}}.$$

For the time-discretization errors, we define

$$\tau_i(Y) = \mathbb{E}\big[|Y_{t_i} - \widehat{Y}_i|^2\big] \quad \text{and} \quad \tau_i(Z) = \mathbb{E}\big[|\bar{Z}_{t_i} - \widehat{Z}_i|^2\big]h.$$

To study the propagation of errors, we introduce the following notion of stability.

DEFINITION 2.5.   We say that the scheme (2.1) is *almost-stable* if there exist constants c and C, independent of N, such that for all $i \in \{0, \ldots, N-1\}$

$$\mathbb{E}\big[|\widehat{Y}_i - Y_i|^2\big] + \frac{1}{4}\mathbb{E}\big[|\widehat{Z}_i - Z_i|^2\big]h \leq (1 + ch)\mathbb{E}\big[|Y_{t_{i+1}} - Y_{i+1}|^2\big] + Ch^{\mu+1}.$$

The terms $Ch^{\mu+1}$ are the *stability imperfections*.

From the fundamental lemma below, the global error is then controlled by three terms. The first is the error made on approximating the terminal condition. The second is the sum, essentially, of the one-step discretization errors, $\sum_{i=0}^{N-1} \frac{\tau_i(Y)}{h} + \tau_i(Z)$. The third is the total contribution of the stability imperfections, and is of order h^{μ}.

LEMMA 2.6 (Fundamental lemma).   *Assume that the scheme* (2.1) *is almost-stable. Then there exist a constant $C \geq 0$ such that, for all $N \geq 1$,*

$$(\text{ERR}_N)^2 \leq C\mathbb{E}\big[|\xi - \xi^N|^2\big] + C\left(\sum_{i=0}^{N-1} \frac{\tau_i(Y)}{h} + \tau_i(Z)\right) + Ch^{\mu}.$$

We place the proof in Appendix A.3. From $(\mathbf{A}\xi^N)$, the first term is known to be of order h. So Theorem 2.4 will be proved if we can prove that the sum of discretization errors if of order h and that the scheme is almost-stable.

REMARK 2.7 (On the almost-stability).   If one could take $C = 0$ in the Definition 2.5 of the almost-stability, then the scheme would be *stable*, in the usual sense. This is the case for the standard explicit and implicit (BTZ) schemes for BSDEs with Lipchitz drivers.

2.4. *Examples of modified drivers, and truncated Brownian increments.*   We show in Sections 3 and 4 how the set of general assumptions of Section 2.2.2 about the modified drivers f^h ensure the convergence of the scheme. As we stated in the Introduction, our aim was to isolate a set of properties that would guarantee the convergence of the modified explicit scheme (2.1) for a large class of modified

drivers f^h, rather than just treating one particular case. To use this result, we then have to show that various modifications of f result in drivers f^h which fit in our framework.

We present here several natural ways to modify the driver f to avoid explosion of the scheme and obtain its convergence. We organize them in three categories, based on how the driver is modified. We also point out that the multiplicative taming, outer taming and inner taming below correspond to (**TCvg**).1, (**TCvg**).2 and (**TCvg**).3 (resp.) of assumption (**TCvg**) in Section 2.2.2.

2.4.1. *Multiplicative taming.* Consider a radius $r(h) = r_0 h^{-\alpha}$, with $r_0 > 0$ and $\alpha > 0$. So $r(h) \to +\infty$ as $h \to 0$. We can tame the high values of f by multiplicating it by a damping factor,

$$f^h(t, y, z) = \chi^h(y) f(t, y, z) \qquad \text{where } \chi^h(y) = \frac{1}{1 + F(y) r(h)^{-1}}.$$

Several choices are possible for the function F. We only consider the four following ones:

  (a) $F(y) = |f(0, y, 0)|$,
  (b) $F(y) = \frac{|f(0, y, 0) - f(0, 0, 0)|}{|y|} \mathbb{1}_{\{y \neq 0\}}$,
  (c) $F(y) = |y|^m$,
  (d) $F(y) = |y|^{m-1}$.

The choices (a) and (b) use only the outputs of f and require no detailed knowledge of f (black-box taming), while the choices (c) and (d) use the input y and require knowing the degree m of the polynomial growth. Also, as can be readily seen, the choices (a) and (c) result in a bounded driver (in the variable y, for fixed h) while (b) and (d) result in a driver with linear growth in y.

2.4.2. *Outer taming.* Consider a radius $r(h) = r_0 h^{-\beta}$, with $r_0 > 0$ and $\beta > 0$. The outer taming is given by

$$f^h(t, y, z) = T^h(f(t, y, z)),$$

where T^h is essentially a projection on the ball of $\mathbb{R}^n$ of center 0 and radius $r(h)$. Specifically, we can consider at least the following two choices:

  (A) The projection (or truncation): $T^h(f) = \frac{f}{\max(1, |f| r(h)^{-1})} = \frac{r(h) f}{\max(r(h), |f|)}$.
  (B) A smoothed projection: $T^h(f) = \frac{f}{1 + |f| r(h)^{-1}} = \frac{r(h) f}{r(h) + |f|}$.

One could also consider a projection on the ball of radius $r(h) + 1$ that is smoothed only at the transition region from identity to constant, so that it would remain the identity on the ball of radius $r(h)$ and be constant in any y-direction beyond $r(h) + 1$.

Notice some general properties of T^h: $|T^h(f)| \le |f|$ and $|T^h(f)| \le r(h)$, for all $f \in \mathbb{R}^n$. The projection also satisfies $T^h(f) = f$ when $|f| \le r(h)$.

Notice that for both the case of the standard projection (A) and the case of the particular smoothed projection (B), the taming can be written multiplicatively, $f^h(t, y, z) = \chi^h(t, y, z) f(t, y, z)$. Indeed, we have

$$T^h(f) = \frac{1}{\max(1, |f| r(h)^{-1})} f \quad \text{and} \quad T^h(f) = \frac{1}{1 + |f| r(h)^{-1}} f$$

in cases (A) and (B), respectively.

Case (A), the standard projection, can therefore be viewed as a generalization-variation of the multiplicatively tamed driver, case (a), the generalization being that we consider a damping factor $\chi^h(t, y, z)$ instead of just $\chi^h(y)$, the variation being that we have to deal with $\max(1, x)$ instead of $1 + x$. For case (B), it is only a generalization of (a). If we consider a driver depending only on y, as we do in all our examples, we see that the outer taming (B) was already treated as the multiplicative taming (a). So let us ignore this case and focus only on the standard projection (A), in this section. From now on, for the outer taming, T^h is the standard projection on the ball of radius $r = r(h)$.

2.4.3. *Inner taming.* Another way to avoid values of f that are too high is to limit the size of the inputs entered in f. Consider a radius $r(h) = r_0 h^{-\gamma}$, with $r_0 > 0$ and $\gamma > 0$. The inner taming is given by

$$f^h(t, y, z) = f(t, T^h(y), z),$$

where T^h is the projection on the ball of $\mathbb{R}^n$ of center 0 and radius $r(h)$. Here, again, as for the outer taming, one could consider a number of variations for the "projection" T^h. We will only study the example of the standard inner projection. Recall the basic properties of T: $|T(y)| \le r$ and $|T(y)| \le |y|$ for all y, and $T(y) = y$ if $|y| \le r$.

2.4.4. *The Brownian increment.* We give an example of H_{i+1} that satisfies (**HH**).3. Namely, $H_{i+1} := \frac{T_{R(h)}(\Delta W_{t_{i+1}})}{h}$ where T_R is the projection (for the distance induced by the infinity-norm $|\cdot|_\infty$ on $\mathbb{R}^d$) on the ball of radius R centered at the origin (for the norm $|\cdot|_\infty$). Otherwise said, each coordinate of $\Delta W_{t_{i+1}}$ is capped at R. Note already that the norm $|\cdot|$ appearing in the expectation (and indeed in all the computations that were done so far) is the Euclidian norm $|\cdot|_2$.

2.5. *Qualitative properties*: *Discrete comparison and positivity preservation.* In this section, we discuss, in the 1-dimensional case, the preservation of order and (consequently) of positivity by the discretization scheme. As was pointed out in [19], the standard explicit scheme may fail to preserve positivity. Here, we are interested in determining conditions under which one can guarantee that the Y_i approximation remains positive when continuous-time solutions is positive. This problem is of significant importance, qualitatively and for numerical stability. On

the one hand, prices should be nonnegative as well as sizes of populations and chemical quantities. On the other hand, the driver may be monotone only on $D = \mathbb{R}_+$ and should one input Y_{i+1} not be a.s. positive, the scheme may explode. We are aware of only two other works stating similar comparison results (see [9] and [8]) but, even in the Lipschitz setting, they do not deal with explicit schemes. The discussion on the explicit scheme is, to the best of our knowledge, new.

The analysis below is based on a linearization technique, as in [8]. We first show a generic discrete comparison result and then positivity follows as a corollary. Essentially due to technicalities arising by the explicit component in the scheme, the comparison result is not as general as one might expect from the implicit scheme. The several result below require at least the following assumption.

ASSUMPTION 2.8. Assume $n = 1$, that **(TReg)**, **(TRegY)** holds with $\mathcal{R}^{(\text{regY})} = 0$ and

$$\text{for } \theta' \in [0, 1] \qquad \sup_{i=0,\dots,N-1} h\big(L_y^h + L_z^h |H_{i+1}| + (1 - \theta')h L_z^h |H_{i+1}| L_y^h\big) < 1.$$

Since we work in the one-dimensional setting, the above assumption is not a drawback.

PROPOSITION 2.9 (Discrete comparison for the explicit scheme). *Let Assumption* 2.8 *hold. For* $j \in \{1, 2\}$ *take the modified drivers* $f^{h,j}$ *and numerical terminal conditions* $\xi^{N,j}$, *as well as the outputs* $(Y_i^j, Z_i^j)_{i=0\dots N}$ *obtained through scheme* (2.1). *Define, for* $0 \le i \le N - 1$,

$$(2.9) \quad \beta_{i+1} := \frac{f^{h,1}(t_i, Y_{i+1}^1, Z_i^1) - f^{h,1}(t_i, Y_{i+1}^2, Z_i^1)}{Y_{i+1}^1 - Y_{i+1}^2} \mathbb{1}_{\{Y_{i+1}^1 - Y_{i+1}^2 \neq 0\}},$$

$$(2.10) \quad \widehat{\beta}_{i+1} := \frac{f^{h,1}(t_i, Y_{i+1}^1, 0) - f^{h,1}(t_i, Y_{i+1}^2, 0)}{Y_{i+1}^1 - Y_{i+1}^2} \mathbb{1}_{\{Y_{i+1}^1 - Y_{i+1}^2 \neq 0\}} \quad and$$

$$(2.11) \quad \gamma_{i+1} := \frac{f^{h,1}(t_i, Y_{i+1}^2, Z_i^1) - f^{h,1}(t_i, Y_{i+1}^2, Z_i^2)}{|Z_i^1 - Z_i^2|^2} (Z_i^1 - Z_i^2)^* \mathbb{1}_{\{Z_i^1 - Z_i^2 \neq 0\}},$$

and $B_{i+1} := 1 + h\beta_{i+1} + h\gamma_{i+1}(1 + (1 - \theta')h\widehat{\beta}_{i+1})H_{i+1}^*$.

Assume further that γ_{i+1} *is* $\mathcal{F}_i$-*measurable for all* i; *and that either* $(f^{h,1} - f^{h,2})(t_i, Y_{i+1}^2, 0)$ *is* $\mathcal{F}_i$-*measurable* $\forall i$ *or* $\theta' = 1$ *or* $\forall i$ $\gamma_{i+1} = 0$. *Then, with the convention* $\prod_{j=k}^l \cdot = 1$ *for* $l < k$,

$$(2.12) \quad Y_i^1 - Y_i^2 = \mathbb{E}_i\Bigg[(\xi^{N,1} - \xi^{N,2}) \prod_{j=i}^{N-1} B_{j+1}$$

$$+ h \sum_{j=i}^{N-1} (f^{h,1} - f^{h,2})(t_j, Y_{j+1}^2, Z_j^2) \prod_{k=i}^{j-1} B_{k+1}\Bigg].$$

If $\xi^{N,1} - \xi^{N,2} \geq 0$ and $(f^{h,1} - f^{h,2})(t_i, Y_{i+1}^2, Z_i^2) \geq 0$ for all $0 \leq i \leq N - 1$ then $Y_i^1 \geq Y_i^2$ for all $0 \leq i \leq N$.

REMARK 2.10 (On the assumptions of the comparison theorem).   Two of the assumptions stand as nontrivial, and perhaps slightly opaque, namely that γ_{i+1} is $\mathcal{F}_i$-measurable and the $\mathcal{F}_i$-measurability of $(f^{h,1} - f^{h,2})(t_i, Y_{i+1}^2, 0)$. The reason for both is of technical nature and due to the presence of the Y_{i+1}-term in the scheme.

Concerning the first, it happens for instance when one is able to write $f(t, y, z) = \tilde{f}(t, y) + \hat{f}(t, z) \, \forall t, y, z$ or when the Y_{i+1}^2's are deterministic. When f does not depend on z then $\gamma_{i+1} = 0$, this is a case of interest for reaction-diffusion equations. When $\theta' \neq 1$, the restriction $(f^{h,1} - f^{h,2})(Y_{i+1}^2, 0)$ is $\mathcal{F}_i$-measurable is a real limitation to the comparison result, as in general Y_{i+1}^2 is not $\mathcal{F}_i$-measurable. However, one is often interested in comparing the scheme against a constant: this will be case case when we prove a corollary on preservation of positivity.

Lastly, one may ask how realistic Assumption 2.8 is, in particular, which choices of H_{i+1} satisfy the assumption. The example in Section 2.4.4 satisfies the assumption. The verification is omitted but follows the ideas presented in [8] (see also Appendix B).

PROOF OF PROPOSITION 2.9.   Let $i \in \{0, \ldots, N - 1\}$. Under assumptions (**TReg**) and (**TRegY**) (with $\mathcal{R}^{(\mathrm{regY})} = 0$) the random variables γ, β and $\hat{\beta}$ are well defined and satisfy: $|\beta_{i+1}| \leq L_y^h$, $|\hat{\beta}_{i+1}| \leq L_y^h$ and $|\gamma_{i+1}| \leq L_z$. Moreover, $\beta_{i+1}, \hat{\beta}_{i+1}$ and γ_{i+1} are $\mathcal{F}_{i+1}$-adapted.

Define $\delta Y_i := Y_i^1 - Y_i^2$, $\delta Z_i := Z_i^1 - Z_i^2$ and $\widehat{\delta f}_{i+1} := (f^{h,1} - f^{h,2})(t_i, Y_{i+1}^2, 0)$. Recalling (2.1), and the notation (2.9), (2.10) and (2.11), we can write

$$\delta Z_i = \mathbb{E}_i\big[\big(\delta Y_{i+1} + (1 - \theta')[f^{h,1}(t_i, Y_{i+1}^1, 0) - f^{h,2}(t_i, Y_{i+1}^2, 0)]h\big)H_{i+1}^*\big]$$
$$= \mathbb{E}_i\big[\delta Y_{i+1}(1 + (1 - \theta')h\hat{\beta}_{i+1})H_{i+1}^* + (1 - \theta')h\widehat{\delta f}_{i+1}H_{i+1}^*\big].$$

For the δY component, we first linearize the driver terms then inject the above expression for δZ. Namely, set $\delta f_{i+1} := (f^{h,1} - f^{h,2})(t_i, Y_{i+1}^2, Z_i^2)$ and using δZ_i from above yields

$$\delta Y_i = \mathbb{E}_i\big[\delta Y_{i+1}(1 + h\beta_{i+1}) + h\gamma_{i+1}\delta Z_i + h\delta f_{i+1}\big] = \mathbb{E}_i[\delta Y_{i+1}B_{i+1} + h\delta f_{i+1}],$$

where we used the assumption that γ_{i+1} is $\mathcal{F}_i$-measurable and then that either $\theta' = 1$ or $\gamma_{i+1} = 0$ or $\widehat{\delta f}_{i+1}$ is $\mathcal{F}_i$-measurable, so the last term vanishes due to the fact that $\mathbb{E}_i[H_{i+1}] = 0$. Furthermore, it is clear that without this second assumption, it not possible to have a comparison result as it is not possible to control the sign of H_{i+1}. Iterating the last inequality from i to N yields (2.12).

The comparison statement: from Assumption 2.8 it follows that all B_i terms are positive, and hence the comparison statement follows provided $\delta f_{i+1} \geq 0$ and $\delta Y_N \geq 0$.   □

REMARK 2.11.   One may also want to consider instead of (2.1), a scheme using as input (Y_{i+1}, Z_{i+1}), for instance setting $Y_i = \mathbb{E}_i[Y_{i+1} + f_i^h(Y_{i+1}, Z_{i+1})h]$. However, due to the presence of Z_{i+1} in the scheme, one cannot guarantee a comparison result for this scheme.

As a corollary of the previous result, we have a preservation of positivity result.

COROLLARY 2.12 (Preservation of positivity).   *Let Assumption* 2.8 *hold. If* $\xi^N \geq 0$ *and for all* $0 \leq i \leq N - 1$ $f^h(t_i, 0, 0) \geq 0$, *then* $Y_i \geq 0$ *for any* $1 \leq i \leq N$, *in other words the tamed explicit scheme* (2.1) *is positivity preserving.*
Moreover, if $\xi^N > 0$ *and* $f^h(t_i, 0, 0) \geq 0$ *for any* $i \in \{0, \ldots, N - 1\}$ *then* $Y_i > 0$ $\forall i$.

PROOF.   In Proposition 2.9, take $(Y_i^2, Z_i^2)_{i=0\ldots N} = 0$ and $f^{h,2} = 0$. Under this setting, the random variables γ_{i+1} defined in (2.11) are $\mathcal{F}_i$-measurable, as is $\widehat{\delta f}_{i+1} = f^{h,1}(t_i, 0, 0) - 0$. The expression (2.12) then simplifies to

$$Y_i^1 = \mathbb{E}_i \left[Y_N^1 \prod_{j=i}^{N-1} B_{j+1} + h \sum_{j=i}^{N-1} f^{h,1}(t_i, 0, 0) \prod_{k=i}^{j-1} B_{k+1} \right].$$

Under Assumption (2.8), the random variables B_{i+1} are all positive and since $\xi^N \geq 0$ and $f^h(t_i, 0, 0) \geq 0$ the statement follows. The case of the strict inequalities follow trivially.   □

2.6. *Numerical simulations.*   In this section, we illustrate with numerical simulations the study performed above regarding the convergence and qualitative properties of explicit schemes with tamed drivers.

Our analysis is concerned with the time-discretization. In practice, we also need to approximate the conditional expectations in the scheme (2.1). We do so using the method of regression on a family of function; see, for instance, [12]. Given a uniform time-discretization grid with N time-steps, $\pi^N = (t_i)_{i=0\ldots N}$ with $t_i = ih$ and $h = T/N$, we therefore simulate a sample of M paths (which are an approximation) of the forward process X: $(X_{m,i}^N)$, for $m = 1, \ldots, M$ and $i = 0, \ldots, N$. In our examples, X will be an arithmetic Brownian motion and simulated exactly (not using a numerical scheme for SDEs). We then use this sample for the regression, at each time t_i, on a family of K functions, which we take to be the first K Hermite polynomials.

2.6.1. *Observing the convergence of the schemes.*   We start by observing the convergence of the schemes. We consider here the following FBSDE, with time-horizon $T = 1$. The forward process X in (1.1) is an Brownian motion started at $x = 0$, with drift $b = 0$ and diffusion coefficient $\sigma = 1$. The BSDE (1.2) has driver $f(t, y, z) = f(y) = -y^3$, which is monotone (decreasing) on $\mathbb{R}$, and terminal condition $\xi = g(X_T)$ where $g(x) = \mathrm{id}(x) = x$ is unbounded.

As explained above, we can consider many generic ways of obtaining a modified driver f^h from f. Here, we consider an example of inner-taming, $f^h(y) = f(T^h(y))$, an example of outer-taming $f^h(y) = T^h(f(y))$, and an example of multiplicative taming $f^h(y) = \chi^h(y) f(y)$. We take the functions T^h to be projections on a ball centered in 0 with growing radius, and we take the taming factor $\chi^h(y)$ of the form $\frac{1}{1+F(y)\tilde{R}(h)^{-1}}$. An analysis of these examples and how they fit in our framework is provided in Appendix B. Specifically, we compute approximations of the solutions using the following schemes:

1. (black) The implicit scheme.
2. (blue) A modified explicit scheme with driver f^h tamed from inside by a projection, $f^h(y) = f(T^h(y))$, where T^h is the projection of the centered ball of radius $r^h = r_0 h^{-\gamma}$, with $r_0 = 1$ and $\gamma = \frac{1}{2(m-1)}$, $m = 3$ being the degree of the polynomial f (see Appendix B for the choice of γ).
3. (green) A modified explicit scheme with driver f^h tamed from outside by a projection, $f^h(y) = T^h(f(y))$, where T^h is the projection of the centered ball of radius $R^h = R_0 h^{-\beta}$, with $R_0 = 1.5$ and $\beta = \frac{1}{2}$.
4. (cyan) A modified explicit scheme with driver f^h given by $f^h(y) = \frac{f(y)}{1+|y|^{m-1}\tilde{R}_0^{-1}h^\alpha}$, where $\tilde{R}_0 = 1$ and $\alpha = \frac{1}{2}$.

We point out the relation between the 2nd case and (**TCvg**).3; the 3rd case and (**TCvg**).2; and the 4th case and (**TCvg**).1.

We generate a sequence of uniform partitions $(\pi^N)_N$ of $[0, T]$ with mesh $h = T/N$ for $N \in \{2, 4, 8, 16, \ldots, 2048\}$ (we simulate first the Brownian paths on the finest partition, and then use these to compute the forward paths X^N on all partition). Since we do not know the exact solution to the FBSDE, we use as a proxy the average Y^{proxy} of the results returned by the schemes 1 and 2 for the finest time-grid. We measure as error the distance $\text{dist}(Y, Y') = \max_{i=0,\ldots,N} \mathbb{E}[|Y_i - Y_i'|^2]^{\frac{1}{2}}$ between, on the one hand, the output $Y^N = (Y_i^N)_{i=0,\ldots,N}$ of one of the schemes (1 to 4) and, on the other hand, the proxy Y^{proxy} for the solution.

On Figure 2, we plot first the error versus the number of time-steps (top picture) and then the computation time versus the error (bottom picture), both in log-log scales.

We observe that the modified explicit schemes 2 and 3 provide errors comparable to the implicit scheme. However, as they are of explicit type, they benefit from a lower computation time. Scheme 4 however does not perform as well. We took $\alpha = \frac{1}{2}$, as suggested by Appendix B so that the scheme fits in our framework. Then Theorem 2.4 guarantees convergence but since $\mu = \alpha < 1$ the modified scheme 4 has a convergence rate lower than the usual rate. Essentially, what slows down the multiplicative schemes with $\chi^h(y)$ of the form $1/(1 + F(y)\tilde{R}(h)^{-1})$ is that even when $F(y)$ [which can be $|f(y)|$, $|y|^m$, $|y|^{m-1}$, $\ldots$] is small, that is, $F(y) \le \tilde{R}(h)$, one does not have $f^h(y) = f(y)$. This creates an error compared to the true dy-
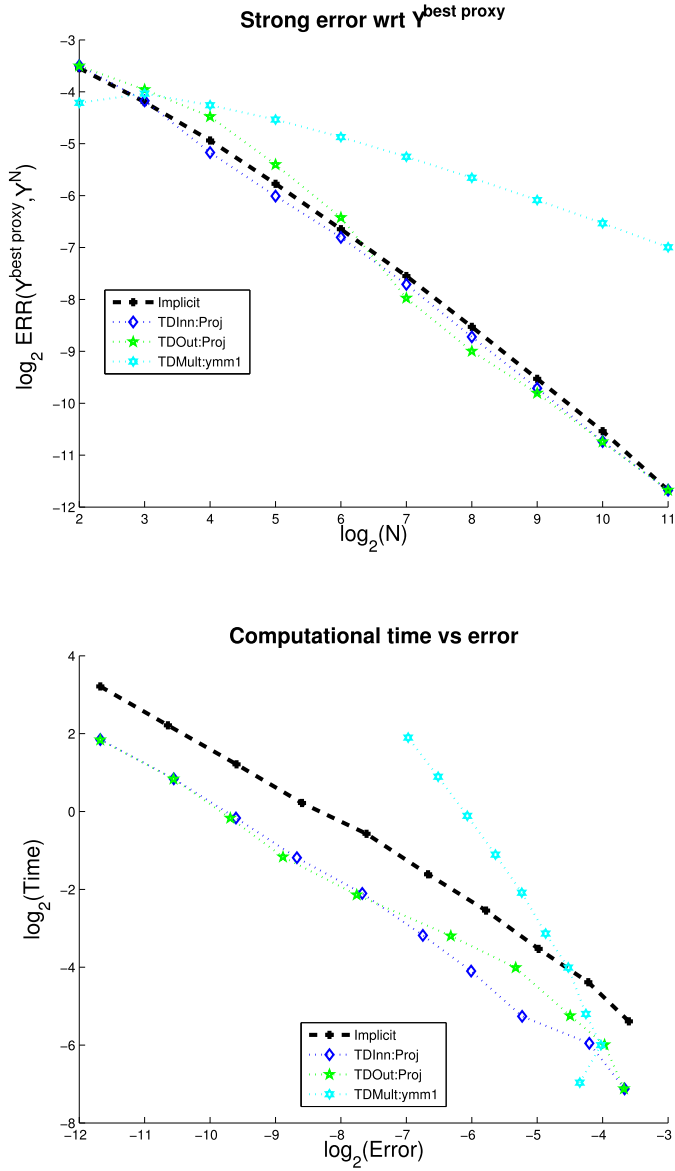
FIG. 2.   *Errors computed for $N \in \{4, 8, 16, \ldots, 2048\}$, with $M = 100{,}000$ and $K = 10$.*

namics, which is not necessary [since $f(y)$ is not big and does not need to be modified], and that error vanishes but too slowly.

2.6.2. *Numerical stability and preservation of positivity.*   We now look at the qualitative behavior of the modified explicit schemes used above.
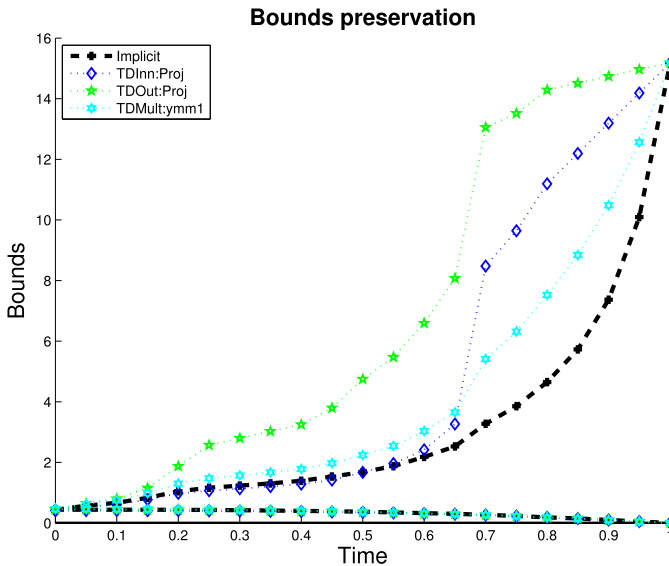
FIG. 3. *Computed with $N = 10$, $M = 50{,}000$ and $K = 30$.*

We consider the following FBSDE, with $T = 1$. The forward component X in (1.1) is a Brownian motion started at $x = 0$, with drift $b = 0$ and diffusion coefficient $\sigma = 1.25$. The BSDE (1.2) has driver $f(t, y, z) = f(y) = -y^2$, which is monotone decreasing on the domain $D = [0, +\infty[$, and terminal condition $\xi = g(X_T)$ where $g(x) = x^2$ is positive.

The solution of the continuous-time BSDE remains positive (i.e., in the domain D) and we have proven in Section 2.5 that the modified explicit schemes should reproduce this property, at least under certain sufficient conditions—Assumption 2.8 reduces in our case to $hL_y^h < 1$. Figure 3 shows the empirical maximum and minimum of Y_i^N, as t_i goes from T to 0.

We observe the desired preservation of positivity, and note that N is only equal to 10. We also observe, when looking at the upper bounds, a regular decay on the implicit scheme. This is due to the strict monotonicity of the driver f. The modified explicit schemes appear to preserve this qualitative behavior as well, though in a limited way.

It must be stressed that such a numerical stability properties, the preservation of positivity and monotonicity, can fail due to an imperfect approximation of the conditional expectations. Our results in Section 2.5 hold only for the time-discretization scheme and rely on the order-preserving property of the conditional expectation operators. In practice, when we tried to approximate the conditional expectations with $K = 4$ functions, we found that frequently (the numerical method we used is a Monte-Carlo one) positivity was violated at some of the grid-points. Due to the taming of the driver, explosion was prevented, but the

qualitative behavior is not satisfactory if the conditional expectations are poorly approximated.

3. Size estimates and nonexplosion of the schemes. In this section, we undertake the size-analysis for the modified explicit scheme (2.1) and show that, under our assumptions, it cannot explode. Specifically, we obtain bounds on the p-moments of the scheme that are uniform in N. For this, we first carry out the size-analysis for *one step* of the scheme. Thanks to the linear growth *and* the monotone growth of the driver f^h, we obtain an estimate which, unlike that for the explicit BTZ scheme (see [19]), can satisfactorily be iterated. This then leads to an almost-sure and uniform-in-N global bound, which in turn leads to bounds on the moments.

3.1. *The one-step almost-sure estimate.* The first results are useful estimates about the size of the output of the scheme over a single time-step. For clarity of the computations, we first prove the result for scheme (2.5)–(2.6) (which has $\hat{D}_i = 0$) in Lemma 3.1 then extend it to the scheme (2.1) (for which $D_i = Z_i - \zeta_i \neq 0$) in Lemma 3.2. We do this in order to show the differences in the estimations. Also, the first result needs a smallness assumption on the step size h while the second, due to the enhanced estimation, *does not*.

LEMMA 3.1. *Assume that the driver f^h satisfies* (**TGrowth**) *and* (**TMonGr**) *with $\bar{M}_z^h \leq \frac{d}{8}$. Let h be such that $h \leq h_0 := \frac{d}{24(K_z^h)^2}$. Then there exist constants $c, C \geq 0$ such that, for any $i \in \{0, \ldots, N-1\}$, and for any random variable $Y_{i+1} \in L^2(\mathcal{F}_{i+1})$, with (Y_i, Z_i) the output of scheme (2.5)–(2.6) for the input Y_{i+1}, one has*

$$|Y_i|^2 + \frac{1}{8}|Z_i|^2 dh + \mathbb{E}_i[\Delta N_{i+1}^2] \leq (1 + ch)\mathbb{E}_i[|Y_{i+1}|^2] + Ch.$$

The constants c and C are uniform in N.

PROOF. By Lemma 2.2, we write

$$Y_i = Y_{i+1} + f^h(t_i, Y_{i+1}, Z_i)h - \Delta M_{i+1}$$
$$= Y_{i+1} + f^h(t_i, Y_{i+1}, Z_i)h - ((Z_i - D_i)\Lambda^{-1}H_{i+1}h + \Delta N_{i+1}).$$

Squaring $Y_i + \Delta M_{i+1} = Y_{i+1} + f^h(t_i, Y_{i+1}, Z_i)h$ and taking $\mathbb{E}_i$, we have

$$|Y_i|^2 + \mathbb{E}_i[|\Delta M_{i+1}|^2] = \mathbb{E}_i[|Y_{i+1}|^2] + \mathbb{E}_i[2\langle Y_{i+1}, f^h(t_i, Y_{i+1}, Z_i)\rangle]h$$
$$+ \mathbb{E}_i[|f^h(t_i, Y_{i+1}, Z_i)|^2]h^2.$$

Using (**TMonGr**) leads to

$$|Y_i|^2 + \mathbb{E}_i[|\Delta M_{i+1}|^2] \leq (1 + 2\bar{M}_y^h h)\mathbb{E}_i[|Y_{i+1}|^2] + 2\bar{M}_t^h h + 2\bar{M}_z^h|Z_i|^2 h$$
$$+ \mathbb{E}_i[|f^h(t_i, Y_{i+1}, Z_i)|^2]h^2.$$

Due to the orthogonality of $H_{i+1}h$ and ΔN_{i+1} and using a Young inequality, we have

$$\mathbb{E}_i\big[|\Delta M_{i+1}|^2\big] = |\zeta_i|^2 \Lambda^{-2} \mathbb{E}_i\big[|H_{i+1}h|^2\big] + \mathbb{E}_i\big[|\Delta N_{i+1}|^2\big]$$

$$= |Z_i - D_i|^2 \Lambda^{-1} dh + \mathbb{E}_i\big[|\Delta N_{i+1}|^2\big]$$

$$\geq \frac{1}{2}|Z_i|^2 \Lambda^{-1} dh - |D_i|^2 \Lambda^{-1} dh + \mathbb{E}_i\big[|\Delta N_{i+1}|^2\big].$$

So the estimate currently yields

$$|Y_i|^2 + \frac{1}{2}|Z_i|^2 \Lambda^{-1} dh - |D_i|^2 \Lambda^{-1} dh + \mathbb{E}_i\big[|\Delta N_{i+1}|^2\big]$$

$$\leq (1 + 2\bar{M}_y^h h)\mathbb{E}_i\big[|Y_{i+1}|^2\big] + 2\bar{M}_t^h h + 2\bar{M}_z^h |Z_i|^2 h$$

$$+ \mathbb{E}_i\big[|f^h(t_i, Y_{i+1}, Z_i)|^2\big]h^2,$$

hence, since $1 \leq \Lambda^{-1} \leq 2$,

$$|Y_i|^2 + \left(\frac{1}{2} - \frac{2\bar{M}_z^h}{d}\right)|Z_i|^2 dh + \mathbb{E}_i\big[|\Delta N_{i+1}|^2\big]$$

$$\leq (1 + 2\bar{M}_y^h h)\mathbb{E}_i\big[|Y_{i+1}|^2\big] + 2\bar{M}_t^h h + 2|D_i|^2 dh$$

$$+ \mathbb{E}_i\big[|f^h(t_i, Y_{i+1}, Z_i)|^2\big]h^2.$$

Now, using the growth of f^h given by **(TGrowth)**, we obtain that

$$\mathbb{E}_i\big[|f^h(t_i, Y_{i+1}, Z_i)|^2\big]h^2 \leq 3(K_t^h)^2 h^2 + 3(K_y^h)^2 h^2 \mathbb{E}_i\big[|Y_{i+1}|^2\big] + 3(K_z^h)^2 |Z_i|^2 h^2.$$

This immediately implies the core estimate:

(3.1)
$$|Y_i|^2 + \left(\frac{1}{2} - \frac{2\bar{M}_z^h}{d}\right)|Z_i|^2 dh + \mathbb{E}_i\big[|\Delta N_{i+1}|^2\big]$$

$$\leq \big(1 + [2\bar{M}_y^h + 3(K_y^h)^2 h]h\big)\mathbb{E}_i\big[|Y_{i+1}|^2\big] + \big(2\bar{M}_t^h + 3(K_t^h)^2 h\big)h$$

$$+ 2|D_i|^2 dh + 3(K_z^h)^2 |Z_i|^2 h^2.$$

Now, recall that (Y_i, Z_i) are produced by the scheme (2.6)–(2.6) with input Y_{i+1}, and we have therefore $D_i = 0$ here. Also, from the assumptions we have

$$\frac{2\bar{M}_z^h}{d} \leq \frac{1}{4} \quad \text{and} \quad 3\frac{(K_z^h)^2}{d}h \leq \frac{1}{8}.$$

This allows to pass the term in $|Z_i|^2 h^2$ on the RHS of the inequality to its LHS. □

We now state the one-step estimate for (2.1), for which $D_i \neq 0$ a priori. We emphasize the additional assumption **(TReg)**, the absence of a smallness condition

on h, and that our handling of the end of the proof is slightly different (even in the case $D_i = 0$).

PROPOSITION 3.2. *Let the driver f^h satisfy* **(TGrowth)**, **(TMonGr)** *with* $\bar{M}_z^h \leq \frac{d}{8}$ *and* **(TReg)**. *Then there exist $c, C \geq 0$ such that, for any $i \in \{0, \ldots, N-1\}$, and for any random variable $Y_{i+1} \in L^2(\mathcal{F}_{i+1})$, with (Y_i, Z_i) the output of scheme* (2.1) *for the input Y_{i+1}, one has*

$$|Y_i|^2 + \frac{1}{4}|Z_i|^2 dh + \mathbb{E}_i\big[|\Delta N_{i+1}|^2\big] \leq (1 + ch)\mathbb{E}_i\big[|Y_{i+1}|^2\big] + Ch.$$

The constants c and C are uniform in N.

PROOF.   We resume from estimate (3.1) of the previous proof, before passing the $|Z_i|^2 h^2$ to the LHS. We start by estimating $|D_i|^2 dh$ as a function of $|Z_i|^2 h^2$.

Following the definition of Z_i in (2.1) and that of ζ_i in (2.4), we have

$$
\begin{aligned}
-D_i &= \zeta_i - Z_i \\
&= \mathbb{E}_i\big[(Y_{i+1} + f^h(t_i, Y_{i+1}, Z_i)h)H_{i+1}^*\big] \\
&\quad - \mathbb{E}_i\big[(Y_{i+1} + (1-\theta')f^h(t_i, Y_{i+1}, 0)h)H_{i+1}^*\big] \\
&= \mathbb{E}_i\big[(f^h(t_i, Y_{i+1}, Z_i) - (1-\theta')f^h(t_i, Y_{i+1}, 0))hH_{i+1}^*\big].
\end{aligned}
$$

Using the Cauchy–Schwarz inequality,

$$
\begin{aligned}
|D_i|^2 dh &\leq dh\mathbb{E}_i\big[|H_{i+1}|^2\big] \times \mathbb{E}_i\big[|(f^h(t_i, Y_{i+1}, Z_i) - (1-\theta')f^h(t_i, Y_{i+1}, 0))h|^2\big] \\
&= \Lambda d^2 \times \mathbb{E}_i\big[|(f^h(t_i, Y_{i+1}, Z_i) - f^h(t_i, Y_{i+1}, 0)) + \theta' f^h(t_i, Y_{i+1}, 0)|^2\big]h^2 \\
&\leq 2d^2(L_z^h)^2|Z_i|^2 h^2 + 2d^2\theta'^2\mathbb{E}_i\big[|f^h(t_i, Y_{i+1}, 0)|^2\big]h^2,
\end{aligned}
$$

where we used the z-regularity of f^h from **(TReg)**, $(a+b)^2 \leq 2a^2 + 2b^2$ and $\Lambda \leq 1$.

Injecting this in (3.1), we obtain

$$
\begin{aligned}
|Y_i|^2 &+ \left(\frac{1}{2} - \frac{2\bar{M}_z^h}{d}\right)|Z_i|^2 dh + \mathbb{E}_i\big[|\Delta N_{i+1}|^2\big] \\
&\leq (1 + [2\bar{M}_y^h + 3(K_y^h)^2 h]h)\mathbb{E}_i\big[|Y_{i+1}|^2\big] + (2\bar{M}_t^h + 3(K_t^h)^2 h)h \\
&\quad + (3(K_z^h)^2 + 4d^2(L_z^h)^2)|Z_i|^2 h^2 + 4d^2\theta'^2\mathbb{E}_i\big[|f^h(t_i, Y_{i+1}, 0)|^2\big]h^2.
\end{aligned}
$$

Using again the growth assumption **(TGrowth)**, we have

$$\mathbb{E}_i\big[|f^h(t_i, Y_{i+1}, 0)|^2\big]h^2 \leq 2(K_t^h)^2 h^2 + 2(K_y^h)^2 h^2 \mathbb{E}_i\big[|Y_{i+1}|^2\big].$$

Hence, the previous estimate implies that

$$|Y_i|^2 + \left(\frac{1}{2} - \frac{2\bar{M}_z^h}{d}\right)|Z_i|^2 dh + \mathbb{E}_i\big[|\Delta N_{i+1}|^2\big]$$
$$\leq \left(1 + [2\bar{M}_y^h + 3(K_y^h)^2 h + 8d^2\theta'^2(K_y^h)^2 h]h\right)\mathbb{E}_i\big[|Y_{i+1}|^2\big]$$
$$+ \left(2\bar{M}_t^h + 3(K_t^h)^2 h + 8d^2\theta'^2(K_t^h)^2 h\right)h + \left(3(K_z^h)^2 + 4d^2(L_z^h)^2\right)|Z_i|^2 h^2.$$

At this stage, instead of assuming h small enough and passing the term in $|Z_i|^2 h^2$ from the RHS to the LHS as in the previous proof, we estimate Z_i directly from its explicit definition in (2.1). From the Cauchy–Schwarz inequality, **(AH)**, $\Lambda \leq 1$ and **(TGrowth)**,

$$|Z_i|^2 h = \big|\mathbb{E}_i\big[(Y_{i+1} + (1 - \theta')f^h(t_i, Y_{i+1}, 0)h)H_{i+1}^*\big]\big|^2 h$$
$$\leq 2d\mathbb{E}_i\big[|Y_{i+1}|^2\big] + 2d(1 - \theta')^2\mathbb{E}_i\big[|f^h(t_i, Y_{i+1}, 0)|^2\big]h^2$$
$$\leq \left(2d + 4d(1 - \theta')^2(K_y^h)^2 h^2\right)\mathbb{E}_i\big[|Y_{i+1}|^2\big] + 4d(1 - \theta')^2(K_t^h)^2 h^2.$$

Injecting this estimate into the previous one and rearranging the terms, we obtain

$$|Y_i|^2 + \left(\frac{1}{2} - \frac{2\bar{M}_z^h}{d}\right)|Z_i|^2 dh + \mathbb{E}_i\big[|\Delta N_{i+1}|^2\big] \leq (1 + c^h h)\mathbb{E}_i\big[|Y_{i+1}|^2\big] + C^h h,$$

where c^h and C^h are given by

$$c^h := 2\bar{M}_y^h + 3(K_y^h)^2 h + 8d^2\theta'^2(K_y^h)^2 h$$
$$+ \left(3(K_z^h)^2 + 4d^2(L_z^h)^2\right)\left(2d + 4d(1 - \theta')^2(K_y^h)^2 h^2\right),$$
$$C^h := 2\bar{M}_t^h + 3(K_t^h)^2 h + 8d^2\theta'^2(K_t^h)^2 h$$
$$+ \left(3(K_z^h)^2 + 4d^2(L_z^h)^2\right)4d(1 - \theta')^2(K_t^h)^2 h^2.$$

Now, first, observe that since we have assumed $\frac{2\bar{M}_z^h}{d} \leq \frac{1}{4}$, the LHS simplifies. Second, from the assumptions on the constants that are made in **(TGrowth)**, **(TMonGr)** and **(TReg)**, and the fact that $h \leq T$, there exist c and C such that $c^h \leq c$ and $C^h \leq C$. □

REMARK 3.3 (On the constants c^h and C^h).   The constants c^h and C^h defined in the above proof depend on h in a bounded way.

First, since $(K_t^h)^2 h$ and $(K_y^h)^2 h$ are bounded, we see that $(K_t^h)^2 h^2$ and $(K_y^h)^2 h^2$ vanish as $h \to 0$. So the last term in C^h and most of the last term in c^h vanish. Here, our handling of the estimates allows to conclude without restrictions on h, at the price of having bigger constants. As $h \to 0$, these bigger constant decrease to, essentially, what they would have been if we had assumed h small and handled the estimates as in Lemma 3.1.

Second, we note that if $\theta' = \theta = 0$ and f does not depend on z (in which case $K_z^h = L_z^h = 0$), then the scheme (2.1) coincides with the scheme (2.5)–(2.6) and the constants c^h and C^h are the same as those found in the previous lemma. As will be clearer in the proof of the almost-stability of the scheme (Proposition 4.2), using scheme (2.1) (not taking $\theta = 0$ and replacing Z_i by zero in the definition of Z_i) creates some errors that vanish at the same rate as the standard error, and only make for bigger constants.

3.2. *The global almost-sure estimate.* The one-step size estimate of Proposition 3.2 can be readily iterated to yield an informative almost-sure bound on the size of the Y_i's, where $(Y_i, Z_i)_{i=0,\ldots,N-1}$ is the output from the iteration of scheme (2.1) with terminal condition initialized to ξ^N.

PROPOSITION 3.4. *Under* (**TGrowth**), (**TMonGr**) *and* (**TReg**), *for any* $i \in \{0, \ldots, N-1\}$,

$$|Y_i|^2 + \mathbb{E}_i\left[\frac{1}{4}\sum_{j=i}^{N-1}|Z_j|^2 dh + \sum_{j=i}^{N-1}|\Delta N_{j+1}|^2\right]$$

$$\leq e^{c(T-t_i)}\mathbb{E}_i\left[|\xi^N|^2\right] + e^{c(T-t_i)}C(T-t_i).$$

PROOF. This proof follows by directly iterating Proposition 3.2; see Lemma A.3. □

3.3. *Moment estimates.* We now show that $(Y_i, Z_i)_{i=0,\ldots,N}$ has p-moments which are bounded uniformly in N. This is crucial in the next section to show that, while the scheme might not be strictly stable, the stability imperfections are small enough that the scheme is almost-stable.

PROPOSITION 3.5. *Assume* (**Aξ^N**), (**TGrowth**), (**TMonGr**) *and* (**TReg**). *For every* $p \geq 1$, *there exists a constant* C (*independent on* N) *such that*

$$\sup_{i=0,\ldots,N-1}\mathbb{E}\left[|Y_i|^{2p}\right] \leq C \quad and \quad \mathbb{E}\left[\sum_{i=0}^{N-1}\left(|Z_i|^2 h\right)^p\right] \leq C.$$

The proof of this result is somewhat similar to that in [19], Proposition 5.1. For the convenience of the reader, we place it in Appendix A.4.

We have analogue results for $(\widehat{Y}_i, \widehat{Z}_i)_{i=0,\ldots,N-1}$ defined in (2.7)–(2.8).

PROPOSITION 3.6 (Integrability of $\widehat{Z}_i$). *Let* (**TGrowth**) *hold. Then, for any* $p \geq 1$ *there exists* $C \geq 0$ *such that, for any* $N \geq 1$,

$$\sup_{i=0,\ldots,N-1}\mathbb{E}\left[|\widehat{Y}_i|^{2p}\right] \leq C \quad and \quad \sup_{0 \leq i \leq N-1}\mathbb{E}\left[|\widehat{Z}_i|^{2p}\right] \leq C.$$

Consequently, we also have the estimate $\mathbb{E}[\sum_{i=0}^{N-1}(|\widehat{Z}_i|^2 h)^p] \leq C.$

PROOF. We start by estimating $\widehat{Z}_i$. Using the Cauchy–Schwarz inequality, $1 - \theta' \le 1$, the fact that $\mathbb{E}_i[|H_{i+1}|^2] = \frac{\Lambda d}{h}$ and $(a+b)^2 \le 2(a^2+b^2)$,

$$|\widehat{Z}_i| \le 2^{\frac{1}{2}} \big(\mathbb{E}_i\big[|Y_{t_{i+1}} - \mathbb{E}_i[Y_{t_{i+1}}]|^2\big] + \mathbb{E}_i\big[|f^h(t_i, Y_{t_{i+1}}, 0)|^2 h^2\big]\big)^{\frac{1}{2}} \Big(\frac{\Lambda d}{h}\Big)^{\frac{1}{2}}.$$

So for all $p \ge 2$, since $(a+b)^q \le 2^{q-1}(a^q + b^q)$,

$$|\widehat{Z}_i|^p \le 2^{p-1}\big(\mathbb{E}_i\big[|Y_{t_{i+1}} - \mathbb{E}_i[Y_{t_{i+1}}]|^2\big]^{\frac{p}{2}} + \mathbb{E}_i\big[|f^h(t_i, Y_{t_{i+1}}, 0)|^2 h^2\big]^{\frac{p}{2}}\big)\Big(\frac{\Lambda d}{h}\Big)^{\frac{p}{2}}.$$

We now take the expectation,

$$\mathbb{E}[|\widehat{Z}_i|^p] \le 2^{p-1}\big(\mathbb{E}\big[\mathbb{E}_i\big[|Y_{t_{i+1}} - \mathbb{E}_i[Y_{t_{i+1}}]|^2\big]^{\frac{p}{2}}\big]$$

$$+ \mathbb{E}\big[\mathbb{E}_i\big[|f^h(t_i, Y_{t_{i+1}}, 0)|^2 h^2\big]^{\frac{p}{2}}\big]\big)\Big(\frac{\Lambda d}{h}\Big)^{\frac{p}{2}}.$$

Now, let us first observe that, using $|f^h| \le f$, **(Growth)**, Jensen's inequality, the tower property of expectations and using Theorem A.2, we have

$$\mathbb{E}\big[\mathbb{E}_i\big[|f^h(t_i, Y_{t_{i+1}}, 0)|^2 h^2\big]^{\frac{p}{2}}\big] = 2^{p-1}\big((K_t)^{2\frac{p}{2}} + (K_y)^{2\frac{p}{2}}\mathbb{E}[|Y_{t_{i+1}}|^{pm}]\big)h^p \le Ch^p.$$

We now address the main difficulty, the term $\mathbb{E}[\mathbb{E}_i[|Y_{t_{i+1}} - \mathbb{E}_i[Y_{t_{i+1}}]|^2]^{\frac{p}{2}}]$. Using the dynamics of the BSDE over $[t_i, t_{i+1}]$ we easily find the identity

$$Y_{t_{i+1}} - \mathbb{E}_i[Y_{t_{i+1}}] = \mathbb{E}_i\bigg[\int_{t_i}^{t_{i+1}} f(u, Y_u, Z_u)\,\mathrm{d}u\bigg]$$

$$- \int_{t_i}^{t_{i+1}} f(u, Y_u, Z_u)\,\mathrm{d}u + \int_{t_i}^{t_{i+1}} Z_u\,\mathrm{d}W_u.$$

Taking the square, using $(\sum_{i=1}^n a_i)^2 \le n\sum_{i=1}^n a_i^2$ and Jensen and/or Cauchy–Schwarz inequalities we obtain

$$|Y_{t_{i+1}} - \mathbb{E}_i[Y_{t_{i+1}}]|^2 \le 3\mathbb{E}_i\bigg[h\int_{t_i}^{t_{i+1}} |f(u, Y_u, Z_u)|^2\,\mathrm{d}u\bigg]$$

$$+ 3h\int_{t_i}^{t_{i+1}} |f(u, Y_u, Z_u)|^2\,\mathrm{d}u + 3\bigg|\int_{t_i}^{t_{i+1}} Z_u\,\mathrm{d}W_u\bigg|^2.$$

We then take the conditional expectation, use the Itô isometry and **(Growth)** to have

$$\mathbb{E}_i\big[|Y_{t_{i+1}} - \mathbb{E}_i[Y_{t_{i+1}}]|^2\big] \le 6h\mathbb{E}_i\bigg[\int_{t_i}^{t_{i+1}} |f(u, Y_u, Z_u)|^2\,\mathrm{d}u\bigg] + 3d\mathbb{E}_i\bigg[\int_{t_i}^{t_{i+1}} |Z_u|^2\,\mathrm{d}u\bigg]$$

$$\le 18(K_t)^2 h^2 + 18(K_y)^2 h\mathbb{E}_i\bigg[\int_{t_i}^{t_{i+1}} |Y_u|^{2m}\,\mathrm{d}u\bigg]$$

$$+ (18(K_z)^2 h + 3d)\mathbb{E}_i\bigg[\int_{t_i}^{t_{i+1}} |Z_u|^2\,\mathrm{d}u\bigg].$$

Consequently, taking the power $\frac{p}{2}$ and using $(\sum_{i=1}^n a_i)^q \le n^{q-1} \sum_{i=1}^n a_i^q$ we have

$$\mathbb{E}_i[|Y_{t_{i+1}} - \mathbb{E}_i[Y_{t_{i+1}}]|^2]^{\frac{p}{2}}$$
$$\le Ch^p + Ch^{\frac{p}{2}} \mathbb{E}_i\left[\int_{t_i}^{t_{i+1}} |Y_u|^{2m}\, \mathrm{d}u\right]^{\frac{p}{2}} + C\mathbb{E}_i\left[\int_{t_i}^{t_{i+1}} |Z_u|^2\, \mathrm{d}u\right]^{\frac{p}{2}}.$$

Here, the constant C depends on p, the growth constants of f and uses $h \le T$. We can then take the expectation, and further use the Jensen inequality (repeatedly) to obtain

$$\mathbb{E}[\mathbb{E}_i[|Y_{t_{i+1}} - \mathbb{E}_i[Y_{t_{i+1}}]|^2]^{\frac{p}{2}}] \le Ch^p + C\mathbb{E}\left[\left(\int_{t_i}^{t_{i+1}} |Z_u|^2\, \mathrm{d}u\right)^{\frac{p}{2}}\right].$$

Gathering the two estimates, we see that we have in the end

$$\mathbb{E}[|\widehat{Z}_i|^p] \le C\left(Ch^p + C\mathbb{E}\left[\left(\int_{t_i}^{t_{i+1}} |Z_u|^2\, \mathrm{d}u\right)^{\frac{p}{2}}\right] + Ch^p\right)\left(\frac{\Delta d}{h}\right)^{\frac{p}{2}}$$
$$\le Ch^{\frac{p}{2}} + Ch^{-\frac{p}{2}} \mathbb{E}\left[\left(\int_{t_i}^{t_{i+1}} |Z_u|^2\, \mathrm{d}u\right)^{\frac{p}{2}}\right].$$

We have now proven the pivotal estimate that will allow us to conclude. Using [19], Theorem 3.4, equation (3.20), that is, $|Z_t| \le C(1 + |X_t|)\, \mathrm{d}t \otimes \mathrm{d}\mathbb{P}$-a.s., to better control the last integral term we obtain, in combination with Theorem A.1, that

$$\mathbb{E}[|\widehat{Z}_i|^p] \le Ch^{\frac{p}{2}} + Ch^{-\frac{p}{2}} \mathbb{E}\left[\left(\int_{t_i}^{t_{i+1}} (1 + |X_u|^2)\, \mathrm{d}u\right)^{\frac{p}{2}}\right]$$
$$\le C\left(h^{\frac{p}{2}} + h^{-\frac{p}{2}} h^{\frac{p}{2}}\left(1 + \mathbb{E}\left[\sup_{0 \le u \le T} |X_u|^p\right]\right)\right) \le C(h^{\frac{p}{2}} + 1)$$

Since $h \le T$ there exists C, independent on h, such that $\sup_{0 \le i \le N-1} \mathbb{E}[|\widehat{Z}_i|^p] \le C$.

The second estimate for $(\widehat{Z}_i)_i$ can be obtained either directly from the above pivotal estimate or from the latest estimate: since $p \ge 1$,

$$\mathbb{E}\left[\sum_{i=0}^{N-1} (|\widehat{Z}_i|^2 h)^p\right] \le \sup_{i=0,\ldots,N-1} \mathbb{E}[|\widehat{Z}_i|^{2p}] T h^{p-1} \le C.$$

Following the arguments just used to prove the estimate for $(\widehat{Z}_i)_i$ it is rather straightforward to prove the remaining estimate for $\widehat{Y}_i$, and hence we omit it. $\square$

4. Convergence of the schemes.

In this section, we prove that scheme (2.1) converges to the solution of the BSDE, as claimed by Theorem 2.4. Following the outline of proof described in Section 2.3, we prove that the scheme is almost-stable and estimate the sum of the discretization errors.

4.1. *Stability estimate of the scheme.*   We consider one step of the scheme with inputs $Y_{t_{i+1}}$ and Y_{i+1}, for which the outputs are respectively $(\widehat{Y}_i, \widehat{Z}_i)$ and (Y_i, Z_i) [recall (2.7)–(2.8) and (2.1)]. We denote by δx the difference $\widehat{x} - x$ for a generic quantity x, $\widehat{x}$ being the counterpart of x when the input is $Y_{t_{i+1}}$. We consequently have

$$\begin{cases} \delta Y_i = \delta Y_{i+1} + \{f^h(t_i, Y_{t_{i+1}}, \widehat{Z}_i) - f^h(t_i, Y_{i+1}, Z_i)\}h - \delta \Delta M_{i+1}, \\ \delta Z_i = \mathbb{E}_i\big[\big(\delta Y_{i+1} + (1 - \theta')\{f^h(t_i, Y_{t_{i+1}}, 0) - f^h(t_i, Y_{i+1}, 0)\}\big)H_{i+1}^*\big]. \end{cases}$$

From Lemma 2.2, we recall that $\delta \Delta M_{i+1}$ has the decomposition $\delta \Delta M_{i+1} = \delta \zeta_i \Lambda^{-1} H_{i+1} h + \delta \Delta N_{i+1}$, and we have $\delta \zeta_i = \mathbb{E}_i[(\delta Y_{i+1} + \{f^h(t_i, Y_{t_{i+1}}, \widehat{Z}_i) - f^h(t_i, Y_{i+1}, Z_i)\}h)H_{i+1}^*]$.

LEMMA 4.1.   *Under* **(TMon)**, **(TReg)** *and* **(TRegY)**, *there exist* $c, C \geq 0$ *such that, without any restriction on the time-step (aside from* $0 < h \leq T$*), we have*

$$|\delta Y_i|^2 + \frac{1}{4}|\delta Z_i|^2 dh + \mathbb{E}_i\big[|\delta \Delta N_{i+1}|^2\big] \leq (1 + ch)\mathbb{E}_i\big[|\delta Y_{i+1}|^2\big] + C\mathcal{I}_i,$$

where

$$\mathcal{I}_i = \mathbb{E}_i\big[\mathcal{R}^{(\mathrm{mon})}(t_i, Y_{t_{i+1}}, Y_{i+1}, \widehat{Z}_i)\big]h$$
$$+ \mathbb{E}_i\big[\mathcal{R}^{(\mathrm{regY})}(t_i, Y_{t_{i+1}}, Y_{i+1}, \widehat{Z}_i)^2\big]h^2 + \mathbb{E}_i\big[\mathcal{R}^{(\mathrm{regY})}(t_i, Y_{t_{i+1}}, Y_{i+1}, 0)^2\big]h^2.$$

PROOF.   The proof presented here works with the extra assumption that $L_z^h > 0$, while **(TReg)** assumes $L_z^h \geq 0$. The case of $L_z^h = 0$ is easy to derive and does not require the constant η below. Overall, the estimations here are very similar to those for the one-step size estimate (see Proposition 3.2). To shorten the notation, we set $\mathcal{I}_i = A_i h + B_i h^2 + B_i^0 h^2$ with A_i, B_i, B_i^0 the 1st, 2nd and 3rd terms in the sum defining $\mathcal{I}_i$ above. Squaring and taking conditional expectation, we have

$$|\delta Y_i|^2 + \mathbb{E}_i\big[|\delta \Delta M_{i+1}|^2\big] = \mathbb{E}_i\big[|\delta Y_{i+1}|^2$$
$$+ 2\langle \delta Y_{i+1}, \{f^h(t_i, Y_{t_{i+1}}, \widehat{Z}_i) - f^h(t_i, Y_{i+1}, Z_i)\}h\rangle$$
$$+ |f^h(t_i, Y_{t_{i+1}}, \widehat{Z}_i) - f^h(t_i, Y_{i+1}, Z_i)|^2 h^2\big].$$

As usual, we add-and-subtract $f^h(t_i, Y_{i+1}, \widehat{Z}_i)$ and then make use of **(TMon)** and **(TReg)** to estimate the 2nd term on the RHS with Young's inequality with parameter η and obtain

$$|\delta Y_i|^2 + \mathbb{E}_i\big[|\delta \Delta M_{i+1}|^2\big]$$
$$\leq (1 + 2(M_y^h + \eta)h)\mathbb{E}_i\big[|\delta Y_{i+1}|^2\big] + A_i h$$
$$+ \frac{(L_z^h)^2}{2\eta}|\delta Z_i|^2 h + \mathbb{E}_i\big[|f^h(t_i, Y_{t_{i+1}}, \widehat{Z}_i) - f^h(t_i, Y_{i+1}, Z_i)|^2\big]h^2.$$

Due to the orthogonality of $H_{i+1}h$ and $\delta\Delta N_{i+1}$ and using a Young inequality, we have

$$
\begin{aligned}
\mathbb{E}_i\big[|\delta\Delta M_{i+1}|^2\big] &= |\delta\zeta_i|^2\Lambda^{-2}\mathbb{E}_i\big[|H_{i+1}h|^2\big] + \mathbb{E}_i\big[|\delta\Delta N_{i+1}|^2\big] \\
&= |\delta Z_i - \delta D_i|^2\Lambda^{-1}dh + \mathbb{E}_i\big[|\delta\Delta N_{i+1}|^2\big] \\
&\geq \frac{1}{2}|\delta Z_i|^2\Lambda^{-1}dh - |\delta D_i|^2\Lambda^{-1}dh + \mathbb{E}_i\big[|\delta\Delta N_{i+1}|^2\big].
\end{aligned}
$$

So, the inequality for $|\delta Y_i|^2 + \mathbb{E}_i[|\delta\Delta M_{i+1}|^2]$ becomes, since $1 \leq \Lambda^{-1}$,

$$
\begin{aligned}
|\delta Y_i|^2 + &\Big(\frac{1}{2} - \frac{(L_z^h)^2}{2\eta d}\Big)|\delta Z_i|^2 dh + \mathbb{E}_i\big[|\delta\Delta N_{i+1}|^2\big] \\
&\leq \big(1 + 2(M_y^h + \eta)h\big)\mathbb{E}_i\big[|\delta Y_{i+1}|^2\big] + A_i h \\
&\quad + |\delta D_i|^2\Lambda^{-1}dh + \mathbb{E}_i\big[|f^h(t_i, Y_{t_{i+1}}, \widehat{Z}_i) - f^h(t_i, Y_{i+1}, Z_i)|^2\big]h^2.
\end{aligned}
$$

We now focus on δD_i:

$$
\begin{aligned}
-\delta D_i &= \delta\zeta_i - \delta Z_i \\
&= \mathbb{E}_i\big[\big(\{f^h(t_i, Y_{t_{i+1}}, \widehat{Z}_i) - f^h(t_i, Y_{i+1}, Z_i)\} \\
&\quad - (1-\theta')\{f^h(t_i, Y_{t_{i+1}}, 0) - f^h(t_i, Y_{i+1}, 0)\}\big)hH_{i+1}^*\big].
\end{aligned}
$$

Hence, using the Cauchy–Schwarz inequality, $\mathbb{E}_i[|H_{i+1}|^2] = \Lambda dh^{-1}$ and $\Lambda \leq 1$, we have

$$
\begin{aligned}
|\delta D_i|^2\Lambda^{-1}dh &\leq \Lambda^{-1}dh\mathbb{E}_i\big[|H_{i+1}|^2\big] \times \mathbb{E}_i\big[|\big(f^h(t_i, Y_{t_{i+1}}, \widehat{Z}_i) - f^h(t_i, Y_{i+1}, Z_i)\big) \\
&\quad - (1-\theta')\big(f^h(t_i, Y_{t_{i+1}}, 0) - f^h(t_i, Y_{i+1}, 0)\big)|^2\big]h^2 \\
&\leq 2d^2\mathbb{E}_i\big[|f^h(t_i, Y_{t_{i+1}}, \widehat{Z}_i) - f^h(t_i, Y_{i+1}, Z_i)|^2\big]h^2 \\
&\quad + 2d^2(1-\theta')^2\mathbb{E}_i\big[|f^h(t_i, Y_{t_{i+1}}, 0) - f^h(t_i, Y_{i+1}, 0)|^2\big]h^2.
\end{aligned}
$$

Reinjecting this in the previous estimate leads to

$$
\begin{aligned}
|\delta Y_i|^2 + &\Big(\frac{1}{2} - \frac{(L_z^h)^2}{2\eta d}\Big)|\delta Z_i|^2 dh + \mathbb{E}_i\big[|\delta\Delta N_{i+1}|^2\big] \\
&\leq \big(1 + 2(M_y^h + \eta)h\big)\mathbb{E}_i\big[|\delta Y_{i+1}|^2\big] + A_i h \\
&\quad + (1 + 2d^2)\mathbb{E}_i\big[|f^h(t_i, Y_{t_{i+1}}, \widehat{Z}_i) - f^h(t_i, Y_{i+1}, Z_i)|^2\big]h^2 \\
&\quad + 2d^2(1-\theta')^2\mathbb{E}_i\big[|f^h(t_i, Y_{t_{i+1}}, 0) - f^h(t_i, Y_{i+1}, 0)|^2\big]h^2.
\end{aligned}
$$

We now use **(TRegY)** and **(TReg)** to estimate

$$\left|f^h(t_i, Y_{t_{i+1}}, \widehat{Z}_i) - f^h(t_i, Y_{i+1}, Z_i)\right|^2$$
$$\leq 4(L_y^h)^2 |\delta Y_{i+1}|^2 + 4\mathcal{R}^{(\mathrm{regY})}(t_i, Y_{t_{i+1}}, Y_{i+1}, \widehat{Z}_i)^2 + 2(L_z^h)^2 |\delta Z_i|^2,$$
$$\left|f^h(t_i, Y_{t_{i+1}}, 0) - f^h(t_i, Y_{i+1}, 0)\right|^2$$
$$\leq 2(L_y^h)^2 |\delta Y_{i+1}|^2 + 2\mathcal{R}^{(\mathrm{regY})}(t_i, Y_{t_{i+1}}, Y_{i+1}, 0)^2.$$

This leads to

$$|\delta Y_i|^2 + \left(\frac{1}{2} - \frac{(L_z^h)^2}{2\eta d}\right)|\delta Z_i|^2 dh + \mathbb{E}_i\left[|\delta \Delta N_{i+1}|^2\right]$$
$$\leq \left(1 + [2(M_y^h + \eta) + 4(1 + 2d^2)(L_y^h)^2 h\right.$$
$$\left. + 4d^2(1 - \theta')^2(L_y^h)^2 h]h\right)\mathbb{E}_i\left[|\delta Y_{i+1}|^2\right]$$
$$+ A_i h + 4(1 + 2d^2)B_i h^2 + 4d^2(1 - \theta')^2 B_i^0 h^2$$
$$+ 2(1 + 2d^2)(L_z^h)^2 \mathbb{E}_i\left[|\delta Z_i|^2 h\right]h.$$

All that remains is to estimate the term with $|\delta Z_i|^2 h$ on the RHS. We have, since $\Lambda \leq 1$,

$$|\delta Z_i|^2 h = \mathbb{E}_i\left[\left(\delta Y_{i+1} + (1 - \theta')\{f^h(t_i, Y_{t_{i+1}}, 0) - f^h(t_i, Y_{i+1}, 0)\}h\right)H_{i+1}^*\right]^2 h$$
$$\leq h^{-1}\mathbb{E}_i\left[|H_{i+1}h|^2\right]$$
$$\times \mathbb{E}_i\left[|\delta Y_{i+1} + (1 - \theta')\{f^h(t_i, Y_{t_{i+1}}, 0) - f^h(t_i, Y_{i+1}, 0)\}h|^2\right]$$
$$\leq 2d\left(1 + 2(1 - \theta')^2(L_y^h)^2 h^2\right)\mathbb{E}_i\left[\delta Y_{i+1}^2\right] + 4d(1 - \theta')^2 B_i^0 h^2.$$

Define the quantities

$$c^h := 2(M_y^h + \eta) + 4(1 + 2d^2)(L_y^h)^2 h + 4d^2(1 - \theta')^2(L_y^h)^2 h$$
$$+ 4(1 + 2d^2)(L_z^h)^2 d\left(1 + 2(1 - \theta')^2(L_y^h)^2 h^2\right) \quad \text{and}$$
$$C^h := 1 + 4(1 + 2d^2) + 4d(1 - \theta')^2\left[d + 2(1 + 2d^2)(L_z^h)^2 h\right].$$

Then, by plugging in the estimates for $|\delta Z_i|^2 h$ and reorganizing the terms we have

$$|\delta Y_i|^2 + \left(\frac{1}{2} - \frac{(L_z^h)^2}{2\eta d}\right)|\delta Z_i|^2 dh + \mathbb{E}_i\left[|\delta \Delta N_{i+1}|^2\right]$$
$$\leq (1 + c^h h)\mathbb{E}_i\left[|\delta Y_{i+1}|^2\right] + C^h\left(A_i h + B_i h^2 + B_i^0 h^2\right).$$

To complete the proof, first, we choose $\eta = \frac{2(L_z^h)^2}{d}$. Second, we note that with this choice of η, and the assumptions on the constants made in **(TRegY)**, **(TReg)** and

(TMon), there exist $c, C \geq 0$ such that $c^h \leq c$ and $C^h \leq C$ for all $h > 0$ (provided $h \leq T$). $\quad\square$

Remark 3.3 applies to the constants c^h and C^h. We can now prove that the scheme is almost-stable. We recall that μ is defined in Section 2.3.

PROPOSITION 4.2. *Under* **(Aξ^N)**, **(TGrowth)**, **(TMonGr)**, **(TMon)**, **(TReg)** *and* **(TRegY)**, *there exist* $c, C \geq 0$ *such that for all* $N \geq 1$ *we have*

$$\mathbb{E}\big[|\delta Y_i|^2\big] + \frac{1}{4}\mathbb{E}\big[|\delta Z_i|^2 dh\big] \leq (1 + ch)\mathbb{E}\big[|\delta Y_{i+1}|^2\big] + Ch^{\mu+1}.$$

PROOF. The first and easy step is to take expectation in the estimate from Lemma 4.1. One must then estimate the imperfection terms in the $\mathbb{E}[\mathcal{I}_i]$. We treat each case separately and we recall that C is a constant whose value can change from line to line.

Case 1. We assume for convenience that $\mathcal{R}^{(\mathrm{regY})}$ and $\mathcal{R}^{(\mathrm{mon})}$ satisfy **(TCvg)**.1 with the same constants. In practice, for the multiplicative tamings (see Appendix B), we find $q = 2m$ and $p = 2$ for $\mathcal{R}^{(\mathrm{regY})}$, and $q = 4m$ and $p = 2$ for $\mathcal{R}^{(\mathrm{mon})}$. But q does not affect the rates. Alternatively, we could argue that $|y|^q \leq 1 + |y|^{2q}$ so both cases fit with $q = 4m$. Let us start with $\mathcal{R}^{(\mathrm{regY})}$.

$$\begin{aligned}
B_i h^2 &= \mathbb{E}\big[|\mathcal{R}^{(\mathrm{regY})}(t_i, Y_{t_{i+1}}, Y_{i+1}, \widehat{Z}_i)|^2\big]h^2 \\
&\leq C\mathbb{E}\big[\big(1 + |Y_{t_{i+1}}|^q + |Y_{i+1}|^q + |\widehat{Z}_i|^p\big)^2\big]h^{2\alpha}h^2 \\
&\leq C\big(1 + \mathbb{E}[|Y_{t_{i+1}}|^{2q}] + \mathbb{E}[|Y_{i+1}|^{2q}] + \mathbb{E}[|\widehat{Z}_i|^{2p}]\big)h^{2\alpha+2} \leq Ch^{2\alpha+2}.
\end{aligned}$$

Here, we have used moment bounds from Theorem A.2 and Propositions 3.5 and 3.6. The case of $B_i^0 h^2$ is similar. Define consequently $\mu_1 = -1 + 2\alpha + 2 = 2\alpha + 1$.

Let us now handle $\mathcal{R}^{(\mathrm{mon})}$.

$$\begin{aligned}
A_i h &= \mathbb{E}\big[\mathcal{R}^{(\mathrm{mon})}(t_i, Y_{t_{i+1}}, Y_{i+1}, \widehat{Z}_i)\big]h \\
&\leq C\mathbb{E}\big[1 + |Y_{t_{i+1}}|^q + |Y_{i+1}|^q + |\widehat{Z}_i|^p\big]h^{\alpha}h \\
&= C\big(1 + \mathbb{E}[|Y_{t_{i+1}}|^q] + \mathbb{E}[|Y_{i+1}|^q] + \mathbb{E}[|\widehat{Z}_i|^p]\big)h^{\alpha+1} \\
&\leq Ch^{\alpha+1}.
\end{aligned}$$

Define consequently $\mu_2 = -1 + \alpha + 1 = \alpha$. The proof is completed by taking $\mu = \min(\mu_1, \mu_2) = \alpha$.

Case 2. We assume again that $\mathcal{R}^{(\mathrm{regY})}$ and $\mathcal{R}^{(\mathrm{mon})}$ satisfy **(TCvg)**.2 with the same constants. In practice, for the outer taming by projection, we find $q = 2m$ and $p = 0$ for $\mathcal{R}^{(\mathrm{regY})}$, and $q = 2m$ and $p = 2$ for $\mathcal{R}^{(\mathrm{mon})}$. Let us start with $\mathcal{R}^{(\mathrm{regY})}$.

Using $\mathbb{1}_{A \cup B} \leq \mathbb{1}_A + \mathbb{1}_B$ and the Cauchy–Schwarz inequality,

$$\mathbb{E}\big[|\mathcal{R}^{(\mathrm{regY})}(t_i, Y_{t_{i+1}}, Y_{i+1}, \widehat{Z}_i)|^2\big]h^2$$

$$\leq C\mathbb{E}\big[(1 + |Y_{t_{i+1}}|^q + |Y_{i+1}|^q + |\widehat{Z}_i|^p)^2$$

$$\times (\mathbb{1}_{\{|f(t_i, Y_{t_{i+1}}, \widehat{Z}_i)| > r(h)\}} + \mathbb{1}_{\{|f(t_i, Y_i, \widehat{Z}_i)| > r(h)\}})^2\big]h^2$$

$$\leq C\mathbb{E}\big[1 + |Y_{t_{i+1}}|^{4q} + |Y_{i+1}|^{4q} + |\widehat{Z}_i|^{4p}\big]^{\frac{1}{2}}$$

$$\times \big(\mathbb{E}[\mathbb{1}_{\{|f(t_i, Y_{t_{i+1}}, \widehat{Z}_i)| > r(h)\}}] + \mathbb{E}[\mathbb{1}_{\{|f(t_i, Y_{i+1}, \widehat{Z}_i)| > r(h)\}}]\big)^{\frac{1}{2}}h^2.$$

Now, we systematically use the moment bounds from Theorem A.2 and Propositions 3.5 and 3.6, as well as the Markov inequality with a power $l \geq 1$ yet to be determined, to have

$$\mathbb{E}\big[|\mathcal{R}^{(\mathrm{regY})}(t_i, Y_{t_{i+1}}, Y_{i+1}, \widehat{Z}_i)|^2\big]h^2$$

$$\leq C\big(\mathbb{E}\big[|f(t_i, Y_{t_{i+1}}, \widehat{Z}_i)|^l\big] + \mathbb{E}\big[|f(t_i, Y_{i+1}, \widehat{Z}_i)|^l\big]\big)^{\frac{1}{2}}r(h)^{-\frac{l}{2}}h^2$$

$$\leq C\big(1 + \mathbb{E}\big[|Y_{t_{i+1}}|^{lm}\big] + \mathbb{E}\big[|Y_{i+1}|^{lm}\big] + \mathbb{E}\big[|\widehat{Z}_i|^l\big]\big)^{\frac{1}{2}}h^{\frac{\beta l}{2}}h^2 \leq Ch^{\frac{\beta l}{2}+2}.$$

Define $\mu_1 = -1 + \frac{\beta l}{2} + 2 = \frac{\beta l}{2} + 1$.

Let us now handle $\mathcal{R}^{(\mathrm{mon})}$. We use the Cauchy–Schwarz inequality, the inequality $(\sum_{i=1}^n a_i)^k \leq n^{k-1}\sum_{i=1}^n a_i^k$, the Markov inequality with a power $l \geq 1$ yet to be determined, and the moment bounds from Theorem A.2 and Propositions 3.5 and 3.6:

$$\mathbb{E}\big[|\mathcal{R}^{(\mathrm{mon})}(t_i, Y_{t_{i+1}}, Y_{i+1}, \widehat{Z}_i)|\big]h$$

$$\leq C\mathbb{E}\big[1 + |Y_{t_{i+1}}|^{2q} + |Y_{i+1}|^{2q} + |\widehat{Z}_i|^{2p}\big]^{\frac{1}{2}}$$

$$\times \mathbb{E}\big[\mathbb{1}_{\{|f(t_i, Y_{t_{i+1}}, \widehat{Z}_i)| > r(h)\}} + \mathbb{1}_{\{|f(t_i, Y_{i+1}, \widehat{Z}_i)| > r(h)\}}\big]^{\frac{1}{2}}h$$

$$\leq C\big(\mathbb{E}\big[|f(t_i, Y_{t_{i+1}}, \widehat{Z}_i)|^l\big] + \mathbb{E}\big[|f(t_i, Y_{i+1}, \widehat{Z}_i)|^l\big]\big)^{\frac{1}{2}}r(h)^{-\frac{l}{2}}h.$$

Using **(Growth)**, we therefore have

$$\mathbb{E}\big[|\mathcal{R}^{(\mathrm{mon})}(t_i, Y_{t_{i+1}}, Y_{i+1}, \widehat{Z}_i)|\big]h$$

$$\leq C\big(1 + \mathbb{E}\big[|Y_{t_{i+1}}|^{lm}\big] + \mathbb{E}\big[|Y_{i+1}|^{lm}\big] + \mathbb{E}\big[|\widehat{Z}_i|^l\big]\big)^{\frac{1}{2}}h^{\frac{\beta l}{2}+1}$$

$$\leq Ch^{\frac{\beta l}{2}+1}.$$

Define $\mu_2 = -1 + \frac{\beta l}{2} + 1 = \frac{\beta l}{2}$. We have the desired result with $\mu = \min(\mu_1, \mu_2) = \frac{\beta l}{2}$. Note that since $\beta > 0$, by taking l big enough one can make the exponent of h

be as big as wanted. Naturally, the constant C depends on the powers q, p and l eventually chosen.

Case 3. We assume again that $\mathcal{R}^{(\mathrm{regY})}$ and $\mathcal{R}^{(\mathrm{mon})}$ satisfy **(TCvg)**.3 with the same constants. In practice, for the inner taming by projection, we find $\mathcal{R}^{(\mathrm{regY})} = 0$, and $q = m$ and $p = 0$ for $\mathcal{R}^{(\mathrm{mon})}$. Let us start with $\mathcal{R}^{(\mathrm{regY})}$. Using $\mathbb{1}_{A \cup B} \leq \mathbb{1}_A + \mathbb{1}_B$ and the Cauchy–Schwarz inequality,

$$
\begin{aligned}
\mathbb{E}\big[&|\mathcal{R}^{(\mathrm{regY})}(t_i, Y_{t_{i+1}}, Y_{i+1}, \widehat{Z}_i)|^2\big]h^2 \\
&\leq C\mathbb{E}\big[\big(1 + |Y_{t_{i+1}}|^q + |Y_{i+1}|^q + |\widehat{Z}_i|^p\big)^2\big(\mathbb{1}_{\{|Y_{t_{i+1}}|>r(h)\}} + \mathbb{1}_{\{|Y_i|>r(h)\}}\big)^2\big]h^2 \\
&\leq C\mathbb{E}\big[1 + |Y_{t_{i+1}}|^{4q} + |Y_{i+1}|^{4q} + |\widehat{Z}_i|^{4p}\big]^{\frac{1}{2}} \\
&\quad \times \big(\mathbb{E}[\mathbb{1}_{\{|Y_{t_{i+1}}|>r(h)\}}] + \mathbb{E}[\mathbb{1}_{\{|Y_{i+1}|>r(h)\}}]\big)^{\frac{1}{2}}h^2.
\end{aligned}
$$

Now, we systematically use the moment bounds from Theorem A.2 and Propositions 3.5 and 3.6, as well as the Markov inequality with a power $l \geq 1$ yet to be determined, to have

$$
\begin{aligned}
\mathbb{E}\big[&|\mathcal{R}^{(\mathrm{regY})}(t_i, Y_{t_{i+1}}, Y_{i+1}, \widehat{Z}_i)|^2\big]h^2 \\
&\leq C\big(\mathbb{E}[|Y_{t_{i+1}}|^l] + \mathbb{E}[|Y_{i+1}|^l]\big)^{\frac{1}{2}}r(h)^{-\frac{l}{2}}h^2 \leq Ch^{\frac{\gamma l}{2}}h^2.
\end{aligned}
$$

Define $\mu_1 = -1 + \frac{\gamma l}{2} + 2 = \frac{\gamma l}{2} + 1$.

Let us now handle $\mathcal{R}^{(\mathrm{mon})}$. We use the Cauchy–Schwarz inequality, the inequality $(\sum_{i=1}^n a_i)^k \leq n^{k-1} \sum_{i=1}^n a_i^k$, the Markov inequality with a power $l \geq 1$ yet to be determined and the moment bounds from Theorem A.2 and Propositions 3.5 and 3.6:

$$
\begin{aligned}
\mathbb{E}\big[&\mathcal{R}^{(\mathrm{mon})}(t_i, Y_{t_{i+1}}, Y_{i+1}, \widehat{Z}_i)\big]h \\
&\leq C\mathbb{E}\big[\big(1 + |Y_{t_{i+1}}|^q + |Y_{i+1}|^q + |\widehat{Z}_i|^p\big)\big(\mathbb{1}_{\{|Y_{t_{i+1}}|>r(h)\}} + \mathbb{1}_{\{|Y_{i+1}|>r(h)\}}\big)\big]h \\
&\leq C\mathbb{E}\big[1 + |Y_{t_{i+1}}|^{2q} + |Y_{i+1}|^{2q} + |\widehat{Z}_i|^{2p}\big]^{\frac{1}{2}}\mathbb{E}\big[\mathbb{1}_{\{|Y_{t_{i+1}}|>r(h)\}} + \mathbb{1}_{\{|Y_{i+1}|>r(h)\}}\big]^{\frac{1}{2}}h \\
&\leq C\big(\mathbb{E}[|Y_{t_{i+1}}|^l] + \mathbb{E}[|Y_{i+1}|^l]\big)^{\frac{1}{2}}r(h)^{-\frac{l}{2}}h \\
&\leq Ch^{\frac{\gamma l}{2}+1}.
\end{aligned}
$$

Define $\mu_2 = -1 + \frac{\gamma l}{2} + 1 = \frac{\gamma l}{2}$. We have the desired result with $\mu = \min(\mu_1, \mu_2) = \frac{\gamma l}{2}$. Note that since $\gamma > 0$, by taking l big enough one can make the exponent of h be as big as wanted. Naturally, the constant C depends on the power l eventually chosen. $\quad\square$

4.2. *Time-discretization errors.* We now turn to the estimation of the local errors, that is to say the error between the BSDE dynamics and the time-discretization scheme (2.1) introduced over one time-step, and in particular their total sum. As before, the assumptions of Section 2.1 are in force.

PROPOSITION 4.3. *Assume* **(TReg)**, **(TCvg)** *and* **(AH)**. *There exists a constant $C \geq 0$ such that, for all $N \geq 1$,*

$$\sum_{i=0}^{N-1} \mathbb{E}[|\overline{Z}_{t_i} - \widehat{Z}_i|^2]h \leq Ch \quad and \quad \sum_{i=0}^{N-1} \mathbb{E}[|Y_{t_i} - \widehat{Y}_i|^2] \leq Ch^2.$$

The proof of this estimate is split in two parts. The first is the estimations for the Z-component, while the second those for the Y-component.

PROOF OF THE ESTIMATE FOR THE Z-COMPONENT IN PROPOSITION 4.3.
First, recall that from the martingale increment property of H_{i+1} we have

$$\mathbb{E}_i\left[\int_{t_i}^{t_{i+1}} Z_u \, dW_u H_{i+1}^*\right] = \mathbb{E}_i\left[\left(Y_{t_{i+1}} + \int_{t_i}^{t_{i+1}} f(u, Y_u, Z_u) \, du\right) H_{i+1}^*\right].$$

We write

$$\overline{Z}_{t_i} - \widehat{Z}_i = \mathbb{E}_i\left[\int_{t_i}^{t_{i+1}} Z_u \, dW_u \frac{\Delta W_{t_{i+1}}^*}{h}\right] - \mathbb{E}_i\left[\int_{t_i}^{t_{i+1}} Z_u \, dW_u H_{i+1}^*\right]$$

$$+ \mathbb{E}_i\left[\left(Y_{t_{i+1}} + \int_{t_i}^{t_{i+1}} f(u, Y_u, Z_u) \, du\right) H_{i+1}^*\right]$$

$$- \mathbb{E}_i\left[(Y_{t_{i+1}} + (1 - \theta') f^h(t_i, Y_{t_{i+1}}, 0)h) H_{i+1}^*\right].$$

Regrouping the terms yields

$$\overline{Z}_{t_i} - \widehat{Z}_i = \mathbb{E}_i\left[\int_{t_i}^{t_{i+1}} Z_u \, dW_u \left(\frac{\Delta W_{t_{i+1}}}{h} - H_{i+1}\right)^*\right]$$

$$+ \mathbb{E}_i\left[\theta' \int_{t_i}^{t_{i+1}} f(u, Y_u, Z_u) \, du H_{i+1}^*\right]$$

$$+ \mathbb{E}_i\left[(1 - \theta') \int_{t_i}^{t_{i+1}} f(u, Y_u, Z_u) - f^h(t_i, Y_{t_{i+1}}, 0) \, du H_{i+1}^*\right].$$

Further decomposing the last term leads finally to

$$\overline{Z}_{t_i} - \widehat{Z}_i = \mathbb{E}_i\left[\int_{t_i}^{t_{i+1}} Z_u \, dW_u \left(\frac{\Delta W_{t_{i+1}}}{h} - H_{i+1}\right)^*\right]$$

$$+ \mathbb{E}_i\left[\theta' \int_{t_i}^{t_{i+1}} f(u, Y_u, Z_u) \, du H_{i+1}^*\right]$$

$$+ (1 - \theta') \mathbb{E}_i \left[\int_{t_i}^{t_{i+1}} f(u, Y_u, Z_u) - f(u, Y_{t_{i+1}}, Z_u) \, du \, H_{i+1}^* \right]$$

$$+ (1 - \theta') \mathbb{E}_i \left[\int_{t_i}^{t_{i+1}} f(u, Y_{t_{i+1}}, Z_u) - f(u, Y_{t_{i+1}}, 0) \, du \, H_{i+1}^* \right]$$

$$+ (1 - \theta') \mathbb{E}_i \left[\int_{t_i}^{t_{i+1}} f(u, Y_{t_{i+1}}, 0) - f(t_i, Y_{t_{i+1}}, 0) \, du \, H_{i+1}^* \right]$$

$$+ (1 - \theta') \mathbb{E}_i \left[\int_{t_i}^{t_{i+1}} f(t_i, Y_{t_{i+1}}, 0) - f^h(t_i, Y_{t_{i+1}}, 0) \, du \, H_{i+1}^* \right]$$

$$= \mathcal{E}_H + \mathcal{E}_{\theta'} + \mathcal{E}_{\text{PR:Y}} + \mathcal{E}_Z + \mathcal{E}_t + \mathcal{E}_{\text{tamed.f}}.$$

We now want to estimate (the expected square of) each of these terms $\mathcal{E}_.$.

Estimation of $\mathcal{E}_H$. Using the Cauchy–Schwarz inequality and the Itô isometry, we have

$$|\mathcal{E}_H|^2 \le d \mathbb{E}_i \left[\int_{t_i}^{t_{i+1}} |Z_u|^2 \, du \right] \mathbb{E} \left[\left| \frac{\Delta W_{t_{i+1}}}{h} - H_{i+1} \right|^2 \right],$$

since $\Delta W_{t_{i+1}} / h - H_{i+1}$ is independent from $\mathcal{F}_i$. Hence, taking expectations,

$$\mathbb{E}[|\mathcal{E}_H|^2] \le d \mathbb{E} \left[\int_{t_i}^{t_{i+1}} |Z_u|^2 \, du \right] \mathbb{E} \left[\left| \frac{\Delta W_{t_{i+1}}}{h} - H_{i+1} \right|^2 \right].$$

Estimation of $\mathcal{E}_{\theta'}$. Using the Cauchy–Schwarz inequality and **(AH)**, then Jentzen we have $|\mathcal{E}_{\theta'}|^2 \le \theta'^2 \mathbb{E}_i [h \int_{t_i}^{t_{i+1}} |f(u, Y_u, Z_u)|^2 \, du] \Lambda d / h$. Taking expectations, since $\Lambda \le 1$, yields $\mathbb{E}[|\mathcal{E}_{\theta'}|^2] \le \theta'^2 d \mathbb{E}[\int_{t_i}^{t_{i+1}} |f(u, Y_u, Z_u)|^2 \, du]$.

Estimation of $\mathcal{E}_{\text{PR:Y}}$. Using the Cauchy–Schwarz inequality and **(AH)** as above, and then the Y-regularity **(RegY)**,

$$|\mathcal{E}_{\text{PR:Y}}|^2 \le (1 - \theta')^2 \mathbb{E}_i \left[h \int_{t_i}^{t_{i+1}} |f(u, Y_u, Z_u) - f(u, Y_{t_{i+1}}, Z_u)|^2 \, du \right] \frac{\Lambda d}{h}$$

$$\le (1 - \theta')^2 \Lambda d L_y^2 \mathbb{E}_i \left[\int_{t_i}^{t_{i+1}} (1 + |Y_u|^{m-1} + |Y_{t_{i+1}}|^{m-1})^2 |Y_u - Y_{t_{i+1}}|^2 \, du \right].$$

When taking the expectation we obtain, using the Cauchy–Schwarz inequality, $\Lambda \le 1$, and that $Y \in \mathcal{S}^p$ for any $p \ge 2$

$$\mathbb{E}[|\mathcal{E}_{\text{PR:Y}}|^2] \le 3^{3/2} (1 - \theta')^2 d L_y^2 (1 + 2\|Y\|_{\mathcal{S}^{4(m-1)}}^{4(m-1)})^{\frac{1}{2}}$$

(4.1)
$$\times \int_{t_i}^{t_{i+1}} \mathbb{E}[|Y_u - Y_{t_{i+1}}|^4]^{\frac{1}{2}} \, du$$

$$\le 3^{3/2} (1 - \theta')^2 d L_y^2 C_Y h (\text{REG}_{Y,4}(h))^{\frac{1}{2}}.$$

With the term $\text{REG}_{Y,4}(h)$ following from the path-regularity Theorem A.2 and from Theorem A.1, it holds that $1 + 2\|Y\|_{\mathcal{S}^{4(m-1)}}^{4(m-1)} \le C_Y$ for some constant $C_Y > 0$.

Estimation of $\mathcal{E}_Z$. Using the Cauchy–Schwarz inequality and **(AH)** as above, and then the Z-regularity **(Reg)**, and $\Lambda \leq 1$,

$$|\mathcal{E}_Z|^2 \leq (1-\theta')^2 \mathbb{E}_i \left[h \int_{t_i}^{t_{i+1}} |f(u, Y_{t_{i+1}}, Z_u) - f(u, Y_{t_{i+1}}, 0)|^2 \, du \right] \frac{\Lambda d}{h}$$

$$\leq (1-\theta')^2 d\Lambda L_z^2 \mathbb{E}_i \left[\int_{t_i}^{t_{i+1}} |Z_u|^2 \, du \right]$$

$$\Rightarrow \quad \mathbb{E}[|\mathcal{E}_Z|^2] \leq (1-\theta')^2 dL_z^2 \mathbb{E} \left[\int_{t_i}^{t_{i+1}} |Z_u|^2 \, du \right].$$

Estimation of $\mathcal{E}_t$. Using the Cauchy–Schwarz inequality and **(AH)** as above, and then the t-regularity **(Reg)**, and $\Lambda \leq 1$,

$$|\mathcal{E}_t^2| \leq (1-\theta')^2 \mathbb{E}_i \left[h \int_{t_i}^{t_{i+1}} |f(u, Y_{t_{i+1}}, 0) - f(t_i, Y_{t_{i+1}}, 0)|^2 \, du \right] \frac{\Lambda d}{h}$$

$$\leq (1-\theta')^2 dL_t \frac{h^2}{2} \leq (1-\theta')^2 dL_t h^2.$$

Estimation of $\mathcal{E}_{\text{tamed.f}}$. Using the Cauchy–Schwarz inequality and **(AH)** as above,

$$|\mathcal{E}_{\text{tamed.f}}|^2 \leq (1-\theta')^2 \mathbb{E}_i \left[h \int_{t_i}^{t_{i+1}} |f(t_i, Y_{t_{i+1}}, 0) - f^h(t_i, Y_{t_{i+1}}, 0)|^2 \, du \right] \frac{\Lambda d}{h}$$

$$\leq (1-\theta')^2 \Lambda dh \mathbb{E}_i [|(f - f^h)(Y_{t_{i+1}}, 0)|^2].$$

Hence, in expectation, $\mathbb{E}[|\mathcal{E}_{\text{tamed.f}}|^2] \leq (1-\theta')^2 dh \mathbb{E}[|(f - f^h)(t_i, Y_{t_{i+1}}, 0)|^2]$.

Gathering the estimates. We finally obtain

$$\sum_{i=0}^{N-1} \mathbb{E}[|\overline{Z}_{t_i} - \widehat{Z}_i|^2] h \leq 6dh \mathbb{E}\left[\int_0^T |Z_u|^2 \, du \right] \max_{i=0,\ldots,N-1} \mathbb{E}\left[\left| \frac{\Delta W_{t_{i+1}}}{h} - H_{i+1} \right|^2 \right]$$

$$+ 6\theta'^2 dh \mathbb{E}\left[\int_0^T |f(u, Y_u, Z_u)|^2 \, du \right]$$

$$+ 6 \cdot 3^{3/2} (1-\theta')^2 dh L_y^2 C_Y (\text{REG}_{Y,4}(h))^{\frac{1}{2}}$$

$$+ 6(1-\theta')^2 dh L_z^2 \mathbb{E}\left[\int_0^T |Z_u|^2 \, du \right]$$

$$+ 6(1-\theta')^2 dL_t h^2$$

$$+ 6(1-\theta')^2 dh^2 \sum_{i=0}^{N-1} \mathbb{E}[|(f - f^h)(t_i, Y_{t_{i+1}}, 0)|^2].$$

Here, $\text{REG}_{Y,4}(h) = \sup_{|s-t| \leq h} \mathbb{E}[|Y_s - Y_t|^4]$. From the path-regularity Theorem A.2, there exists C_{PR} such that $\text{REG}_{Y,4}(h)^{\frac{1}{2}} \leq C_{\text{PR}} h$. Consequently, there

exists a constant C (independent of N) such that

$$\sum_{i=0}^{N-1} \mathbb{E}[|\overline{Z}_{t_i} - \widehat{Z}_i|^2]h \leq Ch \max_{i=0,\ldots,N-1} \mathbb{E}\left[\left|\frac{\Delta W_{t_{i+1}}}{h} - H_{i+1}\right|^2\right] + C\theta'^2 h + CL_y^2 h^2$$

$$+ C(1-\theta')^2 L_z^2 h + C(1-\theta')^2 L_t h^2$$

$$+ Ch^2 \sum_{i=0}^{N-1} \mathbb{E}[|(f - f^h)(t_i, Y_{t_{i+1}}, 0)|^2].$$

The result then follows from the **(AH)**.3 and Lemma 4.4 below. $\square$

LEMMA 4.4.   *Under* **(TCvg)** *there exists a constant* $C \geq 0$ *such that, for all* N,

$$\sum_{i=0}^{N-1} \mathbb{E}[|(f - f^h)(t_i, Y_{t_{i+1}}, 0)|^2] \leq C$$

and

$$\sum_{i=0}^{N-1} \mathbb{E}[|(f - f^h)(t_i, Y_{t_{i+1}}, \overline{Z}_{t_i})|^2] \leq C.$$

We postpone the proof of the lemma above to Appendix A.5 and proceed with the second part of the proof of Proposition 4.3.

PROOF OF THE ESTIMATE FOR THE Y-COMPONENT OF PROPOSITION 4.3. We first decompose

$$Y_{t_i} - \widehat{Y}_i = \mathbb{E}_i\left[\int_{t_i}^{t_{i+1}} f(u, Y_u, Z_u)\,du - f^h(t_i, Y_{t_{i+1}}, \widehat{Z}_i)h\right]$$

$$= \mathbb{E}_i\left[\int_{t_i}^{t_{i+1}} f(u, Y_u, Z_u) - f(u, Y_{t_{i+1}}, Z_u)\,du\right]$$

$$+ \mathbb{E}_i\left[\int_{t_i}^{t_{i+1}} f(u, Y_{t_{i+1}}, Z_u) - f(u, Y_{t_{i+1}}, \overline{Z}_{t_i})\,du\right]$$

$$+ \mathbb{E}_i\left[\int_{t_i}^{t_{i+1}} f(u, Y_{t_{i+1}}, \overline{Z}_{t_i}) - f(t_i, Y_{t_{i+1}}, \overline{Z}_{t_i})\,du\right]$$

$$+ \mathbb{E}_i[f(t_i, Y_{t_{i+1}}, \overline{Z}_{t_i}) - f^h(t_i, Y_{t_{i+1}}, \overline{Z}_{t_i})]h$$

$$+ \mathbb{E}_i[f^h(t_i, Y_{t_{i+1}}, \overline{Z}_{t_i}) - f^h(t_i, Y_{t_{i+1}}, \widehat{Z}_i)]h$$

$$= \mathcal{E}_{\text{PR:Y}} + \mathcal{E}_{\text{PR:Z}} + \mathcal{E}_t + \mathcal{E}_{\text{tamed.f}} + \mathcal{E}_{\tau_i(Z)}.$$

We now estimate (the expected square of) each of these terms.

Estimation of $\mathcal{E}_{\mathrm{PR:Y}}$. Using the Cauchy–Schwarz inequality and **(RegY)**, we have

$$|\mathcal{E}_{\mathrm{PR:Y}}|^2 \leq h L_y^2 \mathbb{E}_i \left[\int_{t_i}^{t_{i+1}} \left(1 + |Y_u|^{m-1} + |Y_{t_{i+1}}|^{m-1}\right)^2 |Y_u - Y_{t_{i+1}}|^2 \, \mathrm{d}u \right]$$

$$\Rightarrow \quad \mathbb{E}\left[|\mathcal{E}_{\mathrm{PR:Y}}|^2\right] \leq h L_y^2 \times C_Y h \left(\mathrm{REG}_{Y,4}(h)\right)^{\frac{1}{2}},$$

where the last line follows by taking expectations and arguing as in (4.1).

Estimation of $\mathcal{E}_{\mathrm{PR:Z}}$. Using Cauchy–Schwarz's inequality and **(Reg)**, we have

$$|\mathcal{E}_{\mathrm{PR:Z}}|^2 \leq \mathbb{E}_i \left[h \int_{t_i}^{t_{i+1}} \left| f(u, Y_{t_{i+1}}, Z_u) - f(u, Y_{t_{i+1}}, \overline{Z}_{t_i}) \right|^2 \, \mathrm{d}u \right]$$

$$\leq h L_z^2 \mathbb{E}_i \left[\int_{t_i}^{t_{i+1}} |Z_u - \overline{Z}_{t_i}|^2 \, \mathrm{d}u \right].$$

Hence, taking expectations, we obtain $\mathbb{E}[|\mathcal{E}_{\mathrm{PR:Z}}|^2] \leq h L_z^2 \mathbb{E}[\int_{t_i}^{t_{i+1}} |Z_u - \overline{Z}_{t_i}|^2 \, \mathrm{d}u]$.

Estimation of $\mathcal{E}_t$. Using Cauchy–Schwarz's inequality and **(Reg)**, we have

$$|\mathcal{E}_t|^2 \leq \mathbb{E}_i \left[h \int_{t_i}^{t_{i+1}} \left| f(u, Y_{t_{i+1}}, \overline{Z}_{t_i}) - f(t_i, Y_{t_{i+1}}, \overline{Z}_{t_i}) \right|^2 \, \mathrm{d}u \right] \leq L_t^2 \frac{h^3}{2}.$$

Estimation of $\mathcal{E}_{\mathrm{tamed.f}}$. Taking squares, using Jensen and taking expectations and using the Cauchy–Schwarz inequality, $\mathbb{E}[|\mathcal{E}_{\mathrm{tamed.f}}|^2] \leq h^2 \mathbb{E}[|(f - f^h) \times (t_i, Y_{t_{i+1}}, \overline{Z}_{t_i})|^2]$.

Estimation of $\mathcal{E}_{\tau_i(Z)}$. Taking the square, using the Cauchy–Schwarz inequality and using the Z-regularity **(TReg)**, we have

$$|\mathcal{E}_{\tau_i(Z)}|^2 \leq h^2 \mathbb{E}_i \left[\left| f^h(t_i, Y_{t_{i+1}}, \overline{Z}_{t_i}) - f^h(t_i, Y_{t_{i+1}}, \widehat{Z}_i) \right|^2 \right]$$

$$\leq h^2 (L_z^h)^2 \mathbb{E}_i \left[|\overline{Z}_{t_i} - \widehat{Z}_i|^2 \right].$$

Hence, taking expectations, we obtain $\mathbb{E}[|\mathcal{E}_{\tau_i(Z)}|^2] \leq h^2 (L_z^h)^2 \mathbb{E}[|\overline{Z}_{t_i} - \widehat{Z}_i|^2]$.

Gathering the estimates. We finally obtain

$$\sum_{i=0}^{N-1} \mathbb{E}\left[|Y_{t_i} - \widehat{Y}_i|^2\right] \leq 5 \cdot 3^{3/2} h^2 L_y^2 C_Y \left(\mathrm{REG}_{Y,4}(h)\right)^{\frac{1}{2}} \cdot N$$

$$+ 5h L_z^2 \cdot \mathrm{REG}_{Z,2}(h) + 5h^3 L_t^2 \cdot N$$

$$+ 5h^2 \sum_{i=0}^{N-1} \mathbb{E}\left[|(f - f^h)(t_i, Y_{t_{i+1}}, \overline{Z}_{t_i})|^2\right]$$

$$+ 5h (L_z^h)^2 \sum_{i=0}^{N-1} \mathbb{E}_i\left[|\overline{Z}_{t_i} - \widehat{Z}_i|^2\right] h.$$

Here, we used again the notation set in the path-regularity Theorem A.2 for $\text{REG}_{Y,4}(h)$ and $\text{REG}_{Z,2}(h)$. From the said result, we have $\text{REG}_{Y,4}(h)^{\frac{1}{2}} \leq Ch$ and $\text{REG}_{Z,2}(h) \leq Ch$, and hence, using the estimates on the size of the solution (see Theorem A.1) the first three terms on the RHS of the above estimate are bounded above by some Ch^2. The same goes for the fourth term by Lemma 4.4 and the first part of the proof guarantees that the same goes for the last term. This completes the proof of Proposition 4.3. $\square$

4.3. *Proof of convergence* (*Theorem* 2.4). We have established with Proposition 4.2 that the scheme is almost-stable, with the rate μ introduced in Section 2.3. By the fundamental Lemma 2.6, we therefore have the global error estimate

$$(\text{ERR}_N)^2 \leq C\mathbb{E}[|\xi - \xi^N|^2] + C\left(\sum_{i=0}^{N-1} \frac{\tau_i(Y)}{h} + \tau_i(Z)\right) + Ch^\mu.$$

By assumption $(\mathbf{A}\xi^N)$, the first term is bounded above by Ch for some C. Proposition 4.3 guarantees that the second term $\sum_{i=0}^{N-1} \frac{\tau_i(Y)}{h} + \tau_i(Z)$ is also bounded above by Ch. Therefore, we have proven $(\text{ERR}_N)^2 \leq Ch + Ch^\mu$, as claimed in Theorem 2.4.

APPENDIX A: AUXILIARY RESULTS

A.1. Background results on monotone BSDE with polynomial growth. The results stated in this section hold under the assumption listed in Section 2.1. They can be found in [19], Sections 2 and 3, and are slightly adapted to suit the framework in the main body of the present work.

THEOREM A.1 (Existence and uniqueness). *The FBSDE* (1.1)–(1.2) *has a unique solution* $(X, Y, Z) \in \mathcal{S}^p \times \mathcal{S}^p \times \mathcal{H}^p$ *for any* $p \geq 2$. *Moreover, it holds that*

$$\|Y\|_{\mathcal{S}^p}^p + \|Z\|_{\mathcal{H}^p}^p \leq C_p\{\|g(X_T)\|_{L^p}^p + \|f(\cdot, X_\cdot, 0, 0)\|_{\mathcal{H}^p}^p\}.$$

Take the uniform partition $\pi = (t_i)_{i=0,\ldots,N}$ of $[0, T]$ as defined in Section 2.2 and with mesh size $|\pi| = h = T/N$. Define the random variables $(\overline{Z}_{t_i})_{i=0,\ldots,N-1}$ by

$$\overline{Z}_{t_i} := \frac{1}{h}\mathbb{E}_i\left[\int_{t_i}^{t_{i+1}} Z_u \, du\right].$$

We have then the following result ([19], Theorem 3.5 and Corollary 3.6)

THEOREM A.2 (Integrability and path regularity). *For any* $p \geq 2$, *there exists a positive constant* C *independent of* h *such that*

$$\sup_{t_i \in \pi} \mathbb{E}[|Y_{t_i}|^p] \leq C \quad and \quad \text{REG}_{Y,p}(h) := \sup_{|s-t| \leq h; t, s \in [0,T]} \mathbb{E}[|Y_s - Y_t|^p] \leq Ch^{\frac{p}{2}},$$

moreover, for the control component Z we have

$$\sum_{i=0}^{N-1} \mathbb{E}\big[(|\overline{Z}_{t_i}|^2 h)^{\frac{p}{2}}\big] \leq C \quad \text{and} \quad \sup_{t_i \in \pi \cap [0,T)} \mathbb{E}\big[|\overline{Z}_{t_i}|^p\big] \leq C$$

and the respective path regularity result

$$\mathrm{REG}_{Z,2}(h) := \mathbb{E}\left[\sum_{i=0}^{N-1} \int_{t_i}^{t_{i+1}} |Z_t - \overline{Z}_{t_i}|^2 \, dt\right] \leq Ch.$$

A.2. Proof of Lemma 2.2 (discrete-time martingale representation).

PROOF OF LEMMA 2.2. Equation (2.2) clearly defines ΔM_{i+1} as being $\mathcal{Y}_{i+1} + f^h(\mathcal{Y}_{i+1}, \mathcal{Z}_i)h - \mathbb{E}_i[\mathcal{Y}_{i+1} + f^h(\mathcal{Y}_{i+1}, \mathcal{Z}_i)h]$. Conversely, defining ΔM_{i+1} clearly yields a martingale increment: $\mathbb{E}_i[\Delta M_{i+1}] = 0$. So we now want to prove the existence and uniqueness of the decomposition $(\zeta_i, \Delta N_{i+1})$ of the martingale increment ΔM_{i+1}.

Uniqueness. Let $(\zeta_i, \Delta N_{i+1})$ be such a decomposition. That is, we have ζ_i a $\mathcal{F}_i$-measurable r.v. and ΔN_{i+1} a martingale increment orthogonal to $H_{i+1}h$ (in other words $\mathbb{E}_i[\Delta N_{i+1}] = 0$ and $\mathbb{E}_i[\Delta N_{i+1}(H_{i+1}h)^*] = 0$), satisfying the decomposition (2.3). Multiplying the equation (2.3) by H_{i+1}^* on the right, taking conditional expectation, using the orthogonality between H_{i+1}^* and ΔN_{i+1}, and recalling **(AH)**.2 implies

$$\mathbb{E}_i\big[\Delta M_{i+1} H_{i+1}^*\big] = \mathbb{E}_i\big[\zeta_i \Lambda^{-1}(H_{i+1}h)(H_{i+1}^*)\big] + \mathbb{E}_i\big[\Delta N_{i+1} H_{i+1}^*\big]$$
$$= \zeta_i \Lambda^{-1} h^{-1} \mathbb{E}_i\big[(H_{i+1}h)(H_{i+1}h)^*\big] + 0 = \zeta_i,$$

which yields the uniqueness of ζ_i. On the other hand, (2.3) directly implies $\Delta N_{i+1} = \Delta M_{i+1} - \zeta_i \Lambda^{-1} H_{i+1}h$, which yields the uniqueness of ΔN_{i+1}.

Existence. Define $\zeta_i = \mathbb{E}_i[\Delta M_{i+1} H_{i+1}^*]$ and $\Delta N_{i+1} = \Delta M_{i+1} - \zeta_i \Lambda^{-1} H_{i+1}h$. It is then obvious that we have $\Delta M_{i+1} = \zeta_i \Lambda^{-1} H_{i+1}h + \Delta N_{i+1}$, and that ζ_i is $\mathcal{F}_i$-measurable. It remains to check that ΔN_{i+1} is a martingale increment and that it is orthogonal to $H_{i+1}h$. The first point follows easily from the fact that ΔM_{i+1} and H_{i+1} are martingale increments and that ζ_i is $\mathcal{F}_i$-measurable:

$$\mathbb{E}_i[\Delta N_{i+1}] = \mathbb{E}_i\big[\Delta M_{i+1} - \zeta_i \Lambda^{-1} H_{i+1}h\big] = \mathbb{E}_i[\Delta M_{i+1}] - \zeta_i \Lambda^{-1} \mathbb{E}_i[H_{i+1}]h = 0.$$

The second point follows by computing, using the definition of ζ_i,

$$\mathbb{E}_i\big[\Delta N_{i+1}(H_{i+1}h)^*\big] = \mathbb{E}_i\big[(\Delta M_{i+1} - \zeta_i \Lambda^{-1} H_{i+1}h)(H_{i+1}h)^*\big]$$
$$= \mathbb{E}_i\big[\Delta M_{i+1} H_{i+1}^*\big]h - \zeta_i \Lambda^{-1} \Lambda h I_d$$
$$= \mathbb{E}_i\big[\Delta M_{i+1} H_{i+1}^*\big]h - \zeta_i h = 0.$$

Therefore, the pair $(\zeta_i, \Delta N_{i+1})$ so-defined is a solution, which proves existence.

To conclude, we just rewrite

$$
\begin{aligned}
\zeta_i &= \mathbb{E}_i\big[\Delta M_{i+1} H_{i+1}^*\big] \\
&= \mathbb{E}_i\big[(\mathcal{Y}_{i+1} + f^h(\mathcal{Y}_{i+1}, \mathcal{Z}_i)h - \mathbb{E}_i[\mathcal{Y}_{i+1} + f^h(\mathcal{Y}_{i+1}, \mathcal{Z}_i)h])H_{i+1}^*\big] \\
&= \mathbb{E}_i\big[(\mathcal{Y}_{i+1} + f^h(\mathcal{Y}_{i+1}, \mathcal{Z}_i)h)H_{i+1}^*\big]. \qquad \square
\end{aligned}
$$

A.3. Iteration and the fundamental lemma. We first state a particularly useful "iteration lemma."

LEMMA A.3. *Let* (a_i), (b_i), (c_i), $i \in \{0, \ldots, N\}$, *be sequences of positive numbers. Assume that there exist constants* $c \geq 0$ *and* $h > 0$, *such that, for all* $i \in \{0, \ldots, N-1\}$,

$$
(A.1) \qquad\qquad a_i + b_i \leq (1 + ch)a_{i+1} + c_i.
$$

Then, for all i,

$$
a_i + \sum_{j=i}^{N-1} b_j \leq e^{c(N-i)h} a_N + \sum_{j=i}^{N-1} e^{c(j-i)h} c_j \leq e^{c(N-i)h} a_N + e^{c(N-1-i)h} \sum_{j=i}^{N-1} c_j.
$$

PROOF. The first estimate is clearly true for $i = N - 1$ (even for $i = N$ in fact), since $1 + ch \leq e^{ch}$. Then, for any given $i \leq N - 2$, if it is true for $i + 1$, by multiplying both sides by e^{ch} we find that

$$
e^{ch} a_{i+1} + e^{ch} \sum_{j=i+1}^{N-1} b_j \leq e^{c(N-i)h} a_N + \sum_{j=i+1}^{N-1} e^{c(j-i)h} c_j.
$$

Summing this inequality with (A.1) and noting that $\sum_{j=i+1}^{N-1} b_j \leq e^{ch} \sum_{j=i+1}^{N-1} b_j$ due to the positivity of the b_j terms gives the first estimate for the given i. The second follows from the fact that $\sum_{j=i}^{N-1} e^{c(j-i)h} c_j \leq e^{c(N-1-i)h} \sum_{j=i}^{N-1} c_j$ since the c_i's are positive. $\square$

We now prove the fundamental lemma.

PROOF OF LEMMA 2.6. We write

$$
\begin{aligned}
\mathbb{E}\big[|Y_{t_i} - Y_i|^2\big] &= \mathbb{E}\big[|Y_{t_i} - \widehat{Y}_i + \widehat{Y}_i - Y_i|^2\big] \\
&\leq \left(1 + \frac{1}{h}\right)\mathbb{E}\big[|Y_{t_i} - \widehat{Y}_i|^2\big] + (1 + h)\mathbb{E}\big[|\widehat{Y}_i - Y_i|^2\big],
\end{aligned}
$$

$$
\mathbb{E}\big[|\overline{Z}_{t_i} - Z_i|^2\big] = \mathbb{E}\big[|\overline{Z}_{t_i} - \widehat{Z}_i + \widehat{Z}_i - Z_i|^2\big] \leq 2\mathbb{E}\big[|\overline{Z}_{t_i} - \widehat{Z}_i|^2\big] + 2\mathbb{E}\big[|\widehat{Z}_i - Z_i|^2\big].
$$

Since the scheme is almost-stable, we have

$$
\mathbb{E}\big[|Y_{t_i} - Y_i|^2\big] + \frac{1}{8}\mathbb{E}\big[|\overline{Z}_{t_i} - Z_i|^2\big]h
$$

$$
\leq \Big(1 + \frac{1}{h}\Big)\mathbb{E}\big[|Y_{t_i} - \widehat{Y}_i|^2\big] + (1+h)\mathbb{E}\big[|\widehat{Y}_i - Y_i|^2\big]
$$

$$
\quad + \frac{1}{4}\mathbb{E}\big[|\overline{Z}_{t_i} - \widehat{Z}_i|^2\big]h + \frac{1}{4}\mathbb{E}\big[|\widehat{Z}_i - Z_i|^2\big]h
$$

$$
\leq (1+h)\Big(\mathbb{E}\big[|\widehat{Y}_i - Y_i|^2\big] + \frac{1}{4}\mathbb{E}\big[|\widehat{Z}_i - Z_i|^2\big]h\Big) + \Big(1+\frac{1}{h}\Big)\tau_i(Y) + \frac{1}{4}\tau_i(Z)
$$

$$
\leq (1+ch)\mathbb{E}\big[|Y_{t_{i+1}} - Y_{i+1}|^2\big] + C\Big(\frac{\tau_i(Y)}{h} + \tau_i(Z)\Big) + Ch^{\mu+1},
$$

where the constants c and C changed on the last line and we used $h \leq T$. Applying Lemma A.3 then gives, since $t_i = ih$ and $h = \frac{T}{N}$,

$$
\mathbb{E}\big[|Y_{t_i} - Y_i|^2\big] + \frac{1}{8}\sum_{j=i}^{N-1}\mathbb{E}\big[|\overline{Z}_{t_j} - Z_j|^2\big]h
$$

$$
\leq e^{c(T-t_i)}\Big(\mathbb{E}\big[|\xi - \xi^N|^2\big] + \sum_{j=i}^{N-1}C\Big(\frac{\tau_i(Y)}{h} + \tau_i(Z)\Big) + \sum_{j=i}^{N-1}Ch^{\mu+1}\Big)
$$

$$
\leq C\mathbb{E}\big[|\xi - \xi^N|^2\big] + C\Big(\sum_{j=i}^{N-1}\frac{\tau_i(Y)}{h} + \tau_i(Z)\Big) + Ch^{\mu}. \qquad \square
$$

A.4. Proof of Proposition 3.5 (moment bounds for the scheme).

PROOF OF PROPOSITION 3.5. Taking the power $p \geq 1$ in the estimate of Proposition 3.4, using $(a+b)^p \leq 2^{p-1}(a^p + b^p)$, we have

$$
|Y_i|^{2p} \leq \big(e^{cT}\mathbb{E}_i\big[|\xi^N|^2\big] + e^{cT}CT\big)^p \leq 2^{p-1}e^{pcT}\mathbb{E}_i\big[|\xi^N|^2\big]^p + 2^{p-1}e^{pcT}(CT)^p.
$$

Using the Jensen inequality and taking the expectation we therefore have

$$
\mathbb{E}\big[|Y_i|^{2p}\big] \leq 2^{p-1}e^{pcT}\mathbb{E}\big[|\xi^N|^{2p}\big] + 2^{p-1}e^{pcT}(CT)^p.
$$

Given the moment assumption in $(\mathbf{A}\xi^N)$, this proves the first estimate. For the second, we come back to the one-step estimate of Proposition 3.2, take the power $p \geq 1$, use $a^p + b^p \leq (a+b)^p$ and use the Jensen inequality to write

$$
|Y_i|^{2p} + \Big(\frac{d}{4}\Big)^p(|Z_i|^2h)^p
$$

$$
\leq \big(e^{ch}\mathbb{E}_i\big[|Y_{i+1}|^2\big] + Ch\big)^p
$$

$$= e^{pch}\mathbb{E}_i[|Y_{i+1}|^2]^p + \sum_{k=1}^{p} \binom{p}{k} (e^{ch}\mathbb{E}_i[|Y_{i+1}|^2])^{p-k}(Ch)^k$$

$$\leq e^{pch}\mathbb{E}_i[|Y_{i+1}|^{2p}] + \sum_{k=1}^{p} \binom{p}{k} (e^{ch}\mathbb{E}_i[|Y_{i+1}|^2])^{p-k}(Ch)^k.$$

Iterating this estimate (see Lemma A.3), we obtain

$$\left(\frac{d}{4}\right)^p \mathbb{E}_i\left[\sum_{j=i}^{N-1}(|Z_j|^2 h)^p\right]$$

$$\leq e^{pc(N-i)h}\mathbb{E}_i[|Y_N|^{2p}]$$

$$+ \sum_{j=i}^{N-1} e^{pc(j-i)h}\mathbb{E}_i\left[\sum_{k=1}^{p} \binom{p}{k} (e^{ch}\mathbb{E}_j[|Y_{j+1}|^2])^{p-k}(Ch)^k\right]$$

$$\leq e^{pc(T-t_i)}\mathbb{E}_i[|\xi^N|^{2p}]$$

$$+ e^{pc(N-1-i)h} \sum_{j=i}^{N-1}\sum_{k=1}^{p} \binom{p}{k} e^{c(p-k)h}\mathbb{E}_i[\mathbb{E}_j[|Y_{j+1}|^2]^{p-k}](Ch)^k.$$

One can then use the Hölder inequality and the Jensen inequality to further obtain

$$\left(\frac{d}{4}\right)^p \mathbb{E}_i\left[\sum_{j=i}^{N-1}(|Z_j|^2 h)^p\right]$$

$$\leq e^{pc(T-t_i)}\mathbb{E}_i[|\xi^N|^{2p}] + e^{pc(T-t_i)} \sum_{j=i}^{N-1}\sum_{k=1}^{p} \binom{p}{k} \mathbb{E}_i[\mathbb{E}_j[|Y_{j+1}|^{2p}]]^{\frac{p-k}{p}}(Ch)^k$$

$$= e^{pc(T-t_i)}\mathbb{E}_i[|\xi^N|^{2p}] + e^{pc(T-t_i)} \sum_{k=1}^{p} \binom{p}{k} \sum_{j=i}^{N-1} \mathbb{E}_i[|Y_{j+1}|^{2p}]^{\frac{p-k}{p}}(Ch)^k.$$

In particular, for $i = 0$, we obtain in the end

$$\left(\frac{d}{4}\right)^p \mathbb{E}\left[\sum_{j=0}^{N-1}(|Z_j|^2 h)^p\right]$$

$$\leq e^{pcT}\mathbb{E}[|\xi^N|^{2p}] + e^{pcT} \sum_{k=1}^{p} \binom{p}{k} \left(\sup_{0\leq j\leq N-1} \mathbb{E}[|Y_{j+1}|^{2p}]\right)^{\frac{p-k}{p}} C^k h^{k-1} T,$$

which, in view of the moment estimate just proved for (Y_i), yields the desired result. $\quad\square$

A.5. Proof of Lemma 4.4 (vanishing effect of the tamed drivers).

PROOF OF LEMMA 4.4.   We verify the claim of the lemma in each of the three cases covered by our assumption (**TCvg**). We only prove the second estimate, as it is clear from the proof how the first follows. Recall that $R^h = f - f^h$, and that C denotes a constant whose value may change from line to line.

Case 1. We use the inequality $(\sum_{i=1}^n a_i)^2 \leq n \sum_{i=1}^n a_i^2$ to write

$$\mathbb{E}[|R^h(Y_{t_{i+1}}, \overline{Z}_{t_i})|^2] \leq C\mathbb{E}[(1 + |Y_{t_{i+1}}|^q + |\overline{Z}_{t_i}|^p)^2]h^{2\alpha} \leq Ch^{2\alpha},$$

where we used the moment bounds from Theorem A.2. In turn, we have

$$\sum_{i=0}^{N-1} \mathbb{E}[|R^h(Y_{t_{i+1}}, \overline{Z}_{t_i})|^2] \leq Ch^{2\alpha-1} \leq C,$$

since $\alpha \geq \frac{1}{2}$ by assumption. This proves the desired result for Case 1.

Case 2. Using $(\sum_{i=1}^n a_i)^q \leq C \sum_{i=1}^n a_i^q$, the Cauchy–Schwarz inequality, the Markov inequality with a power $l \geq 1$ yet to be determined, and (**Growth**), we have

$$\mathbb{E}[|R^h(Y_{t_{i+1}}, \overline{Z}_{t_i})|^2] \leq C\mathbb{E}[(1 + |Y_{t_{i+1}}|^{2q} + |\overline{Z}_{t_i}|^{2q})\mathbb{1}_{\{|f(Y_{t_{i+1}}, \overline{Z}_{t_i}) > r(h)|\}}]$$

$$\leq C\mathbb{E}[1 + |Y_{t_{i+1}}|^{4q} + |\overline{Z}_{t_i}|^{4p}]^{\frac{1}{2}}(\mathbb{E}[|f(Y_{t_{i+1}}, \overline{Z}_{t_i})|^l]r(h)^{-l})^{\frac{1}{2}}$$

$$\leq C\mathbb{E}[1 + |Y_{t_{i+1}}|^{4q} + |\overline{Z}_{t_i}|^{4q}]^{\frac{1}{2}}\mathbb{E}[1 + |Y_{t_{i+1}}|^{lm} + |\overline{Z}_{t_i}|^l]^{\frac{1}{2}}h^{\frac{\beta l}{2}}.$$

We have then $\mathbb{E}[|R^h(Y_{t_{i+1}}, \overline{Z}_{t_i})|^2] \leq Ch^{\frac{\beta l}{2}}$ by systematically using the moment bounds from Theorem A.2. The desired result then follows since, given that $\beta > 0$, we can take $l = \frac{2}{\beta}$, so that $\frac{\beta l}{2} \geq 1$.

Case 3. Using $(\sum_{i=1}^n a_i)^q \leq C \sum_{i=1}^n a_i^q$, the Cauchy–Schwarz inequality, and the Markov inequality with a power $l \geq 1$ yet to be determined, we have

$$\mathbb{E}[|R^h(Y_{t_{i+1}}, \overline{Z}_{t_i})|^2] \leq C\mathbb{E}[(1 + |Y_{t_{i+1}}|^{2q} + |\overline{Z}_{t_i}|^{2p})\mathbb{1}_{\{|Y_{t_{i+1}}| > r(h)\}}]$$

$$\leq C\mathbb{E}[1 + |Y_{t_{i+1}}|^{4q} + |\overline{Z}_{t_i}|^{4p}]^{\frac{1}{2}}\mathbb{E}[\mathbb{1}_{\{|Y_{t_{i+1}}| > r(h)\}}]^{\frac{1}{2}}$$

$$\leq C\mathbb{E}[1 + |Y_{t_{i+1}}|^{4q} + |\overline{Z}_{t_i}|^{4p}]^{\frac{1}{2}}\mathbb{E}[|Y_{t_{i+1}}|^l]^{\frac{1}{2}}h^{\frac{\gamma l}{2}}.$$

Using systematically the moment bounds from Theorem A.2, we then have

$$\mathbb{E}[|R^h(Y_{t_{i+1}}, \overline{Z}_{t_i})|^2] \leq Ch^{\frac{\gamma l}{2}}.$$

The desired result then follows since, given that $\gamma > 0$, we can take l, so that $\frac{\gamma l}{2} \geq 1$. $\square$

APPENDIX B: VERIFICATIONS FOR SOME STANDARD WAYS TO TAME THE DRIVER

Some verifications can be found in the preprint version of this work: arXiv: 1607.06733. The material presented in Appendix B there was also subject to peer review, but removed from the final version of this manuscript due to size constraints.

REFERENCES

[1] BENDER, C. and DENK, R. (2007). A forward scheme for backward SDEs. *Stochastic Process. Appl.* **117** 1793–1812. MR2437729

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

[3] BRIAND, P., DELYON, B. and MÉMIN, J. (2002). On the robustness of backward stochastic differential equations. *Stochastic Process. Appl.* **97** 229–253. MR1875334

[4] BRIAND, P. and LABART, C. (2014). Simulation of BSDEs by Wiener chaos expansion. *Ann. Appl. Probab.* **24** 1129–1171.

[5] CHASSAGNEUX, J.-F. An introduction to the numerical approximation of BSDEs. In 2*nd Summer School of the Euro-Mediterranean Research Center for Mathematics and Its Applications* (*EMRCMA*).

[6] CHASSAGNEUX, J.-F., JACQUIER, A. and MIHAYLOV, I. (2016). An explicit Euler scheme with strong rate of convergence for financial SDEs with non-Lipschitz coefficients. *SIAM J. Financial Math.* **7** 993–1021. MR3582409

[7] CHASSAGNEUX, J.-F. and RICHOU, A. (2015). Numerical stability analysis of the Euler scheme for BSDEs. *SIAM J. Numer. Anal.* **53** 1172–1193. Available at arXiv:1407.0887.

[8] CHASSAGNEUX, J.-F. and RICHOU, A. (2016). Numerical simulation of quadratic BSDEs. *Ann. Appl. Probab.* **26** 262–304. MR3449318

[9] CHERIDITO, P. and STADJE, M. (2013). BSΔEs and BSDEs with non-Lipschitz drivers: Comparison, convergence and robustness. *Bernoulli* **19** 1047–1085. Available at arXiv:1002.0175v1.

[10] CRISAN, D. and MANOLARAKIS, K. (2014). Second order discretization of backward SDEs and simulation with the cubature method *Ann. Appl. Probab.* **24** 652–678. MR3178494

[11] EL KAROUI, N., PENG, S. and QUENEZ, M. C. (1997). Backward stochastic differential equations in finance. *Math. Finance* **7** 1–71.

[12] GOBET, E. and TURKEDJIEV, P. (2016). Approximation of backward stochastic differential equations using Malliavin weights and least-squares regression. *Bernoulli* **22** 530–562.

[13] HENRY, D. (1981). *Geometric Theory of Semilinear Parabolic Equations. Lecture Notes in Math.* **840**. Springer, Berlin.

[14] HUTZENTHALER, M. and JENTZEN, A. (2014). On a perturbation theory and on strong convergence rates for stochastic ordinary and partial differential equations with non-globally monotone coefficients. Preprint. Available at arXiv:1401.0295.

[15] HUTZENTHALER, M. and JENTZEN, A. (2015). *Numerical Approximations of Stochastic Differential Equations with Non-globally Lipschitz Continuous Coefficients. Memoirs of the American Mathematical Society* **236**. Amer. Math. Soc., Providence, RI.

[16] HUTZENTHALER, M., JENTZEN, A. and KLOEDEN, P. E. (2011). Strong and weak divergence in finite time of Euler's method for stochastic differential equations with non-globally Lipschitz continuous coefficients. *Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.* **467** 1563–1576.

[17] HUTZENTHALER, M., JENTZEN, A. and KLOEDEN, P. E. (2012). Strong convergence of an explicit numerical method for SDEs with nonglobally Lipschitz continuous coefficients. *Ann. Appl. Probab.* **22** 1611–1641. MR2985171

[18] HUTZENTHALER, M., JENTZEN, A. and WANG, X. (2013). Exponential integrability properties of numerical approximation processes for nonlinear stochastic differential equations. Preprint. Available at arXiv:1309.7657.

[19] LIONNET, A., DOS REIS, G. and SZPRUCH, L. (2015). Time discretization of FBSDE with polynomial growth drivers and reaction–diffusion PDEs. *Ann. Appl. Probab.* **25** 2563–2625. MR3375884

[20] ROTHE, F. (1984). *Global Solutions of Reaction–Diffusion Systems. Lecture Notes in Math.* **1072**. Springer, Berlin.

[21] SZPRUCH, L. and ZHANG, X. (2013). *V*-integrability, asymptotic stability and comparison theorem of explicit numerical schemes for SDEs. Preprint. Available at arXiv:1310.0785.

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

A. LIONNET
INRIA PARIS—ECOLE NATIONALE DES
   PONTS ET CHAUSSÉES
2 RUE SIMONE IFF
75012 PARIS
FRANCE
E-MAIL: arnaud.lionnet@cmla.ens-cachan.fr

G. DOS REIS
L. SZPRUCH
SCHOOL OF MATHEMATICS
UNIVERSITY OF EDINBURGH
EDINBURGH, EH9 3JZ
UNITED KINGDOM
E-MAIL: g.dosreis@ed.ac.uk
            l.szpruch@ed.ac.uk