



Some Analytic Approximations for Backward Stochastic Differential Equations

Jasmina Đorđević^a

^aFaculty of Sciences and Mathematics, University of Niš,
Višegradska 33, 18000 Niš, Serbia

Abstract. We consider an analytic iterative method to approximate the solution of the backward stochastic differential equation of general type. More precisely, we define a sequence of approximate equations and give sufficient conditions under which the approximate solutions converge with probability one and in p th moment sense, $p \geq 2$, to the solution of the initial equation under Lipschitz condition. The Z-algorithm for this iterative method is introduced and some examples are presented to illustrate the theory.

1. Introduction

The theory of backward stochastic differential equations (shorter BSDEs) was developed in the early 1990s, by Pardoux and Peng. They established some results on the existence and uniqueness of the adapted solutions in their founder paper [20]. Since then, BSDEs have been intensively developed both theoretically and in various applications. In papers of Pardoux and Peng [20] and Pardoux [21], they gave a probabilistic representation for the solutions of some semilinear and quasilinear parabolic partial differential equations in terms of solutions of BSDEs by obtaining a generalization of the well-known Feynman-Kac formula. Likewise, BSDEs are encountered in many fields of applied mathematics such as finance [14, 15], stochastic games and optimal control [6, 17], as well as partial differential equations and homogenization, for instance. There exists an extensive literature on BSDEs. We mention here the collected papers [16] edited by El Karoui and Mazliak, which contain a useful introduction into the theory of BSDEs and their applications.

Author by herself have dealt with different problems related to several type of backward differential equations. Problem of additive perturbations was considered by Janković, Jovanović and Đorđević in their paper [12] for nonhomogeneous BSDEs, and later on by Đorđević and Janković in [2] for Volterra BSDEs. Đorđević [5] proved the closeness result for the general type of perturbations for reflected BSDEs. For backward doubly stochastic differential equations Đorđević et al. [3, 4, 11] proved existence result for nonhomogeneous class of equations, L^p stability and obtained a generalization of the well-known Feynman-Kac formula for those equations (all those problems are proved under several different conditions).

The topic of the present paper is an analytic method named the Z-algorithm that is used for solving backward stochastic differential equations. The essentials of the problem have their origin in papers [23, 24]

2010 *Mathematics Subject Classification.* Primary 60H35; Secondary 93E10, 93E25

Keywords. Backward stochastic differential equations, iterative method, convergence with probability one and in the p -th moment, Z-algorithm.

Received: 27 June 2019; Revised: 27 January 2020; Accepted: 28 September 2020

Communicated by Miljana Jovanović

Supported by Grant No 174007 of MNTRS.

Email address: nina19@pmf.ni.ac.rs, djordjevichristina@gmail.com (Jasmina Đorđević)

by Zuber, who treated one of the general analytic iterative methods for solving the Cauchy problem for an ordinary differential equation $x' = f(t, x)$, $x(t_0) = x_0$. With this equations he analyzed solutions to the equations $x'_{n+1} = f_n(t, x_{n+1})$, $x_{n+1}(t_0) = x_0, n \in \mathbb{N}_0$, all defined on the interval $[t_0 - a, t_0 + a]$. In [23] Zuber showed that if $\sum_{n=1}^{\infty} \sup_{|t-t_0|<a} |f(t, x_n(t)) - f_n(t, x_n(t))| < \infty$, then there exists a constant $h, h \in (0, a]$, so that the sequence of the solutions $\{x_n, n \in \mathbb{N}\}$ uniformly converges to the solution x of the initial equation on the interval $[t_0 - h, t_0 + h]$. If the functions f_n are chosen well so that the approximate equations can be effectively solved, then an ε -approximation of the solution x can be effectively found in the sense that there exists natural number $n(\varepsilon)$ such that $\sup_{[t_0-h, t_0+h]} |x(t) - x_n(t)| < \varepsilon$ for all $n \geq n(\varepsilon)$.

The Z-algorithm in [23], represents a general algorithm for solving ordinary differential equations because many well-known, historically and practically important analytic and numerical methods are its special cases: the Picard method of successive approximations, Chaplygin methods of chords and tangents, Newton Kantorovich method and some interpolation methods, as Euler one, among other things. The specific fact is that the function f_n determines the solution x_{n+1} and depend on certain sense on x_n , the previous solution. For this reason the sequence $\{f_n, n \in \mathbb{N}\}$ is called the determining sequence for the Z-algorithm.

Zuber’s idea for iteration inspired authors to observe similar analytic method for solving other types of equations. Janković [7, 8] presented an analogous analytic method for the forward stochastic differential equation of the Itô type, while Janković and Jovanović [9] extended Zuber’s approach to various classes of stochastic hereditary differential equations. Recently, Janković, Vasilova and Krstić [10] adapted Z-algorithm for neutral stochastic functional differential equations.

It should be noted that in all those papers the choice of the determining sequence for the Z-algorithm has the most important role, because it provides the adequate approximation of the solution of given equation. Even though there has been several papers on numerical methods for BSDEs before (see[22]), all those papers observed a type of BSDE which has function g in diffusion coefficient equals to zero (see Eq.(1)). Comparing to this, in our paper a general type of BSDE (so called "nonhomogeneous") with function g different from zero is observed. Even more, this iterative method is more general compering to all that were proven until now (see more in the section Remarks and Conclusion).

Before stating the main problems and results to be explained and proved, we briefly reproduce only the essential notations and definitions which are necessary in our investigation. The initial assumption is that all random variables and processes are defined on a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq 0}, P)$ with a filtration $\{\mathcal{F}_t\}_{t \geq 0}$ generated by an m -dimensional Brownian motion $w = \{w(t)\}_{t \geq 0}$, i.e., $\mathcal{F}_t = \sigma\{w(s), 0 \leq s \leq t\}$. As usual, let $|\cdot|$ denote the Euclidean norm in R^n . The trace norm of a matrix B is denoted by $|B| = \sqrt{\text{trace}[B^T B]}$, where B^T is the transpose of a matrix or vector. Likewise, $L^r_{\mathcal{F}_T}(\Omega; R^d)$, $r > 0$, is the family of \mathcal{F}_T -measurable R^d -valued random variables X such that $E|X|^r < \infty$, while $\mathcal{M}^r([0, T]; R^d)$, $r > 0$, is the family of R^d -valued \mathcal{F}_t -adapted processes $\{\varphi(t)\}_{0 \leq t \leq T}$ such that $E \int_0^T |\varphi(t)|^r dt < \infty$.

In this paper, we consider the following nonhomogeneous BSDE

$$x(t) = \xi - \int_t^T f(x(s), y(s), s)ds - \int_t^T [g(x(s), s) + y(s)]dw(s), \quad t \in [0, T], \tag{1}$$

with a terminal condition $\xi \in L^2_{\mathcal{F}_T}(\Omega; R^d)$. The mappings $f : R^d \times R^{d \times m} \times [0, T] \times \Omega \rightarrow R^d$ and $g : R^d \times [0, T] \times \Omega \rightarrow R^{d \times m}$ are assumed to be $\mathcal{B}_d \otimes \mathcal{B}_{d \times m} \otimes \mathcal{P}$ -measurable and $\mathcal{B}_d \otimes \mathcal{P}$ -measurable, respectively, where \mathcal{P} is a σ -algebra of \mathcal{F}_t -progressively measurable subsets of $[0, T] \times \Omega$.

A pair of stochastic processes

$$\{x(t), y(t)\}_{0 \leq t \leq T} \in \mathcal{M}^2([0, T]; R^d) \times \mathcal{M}^2([0, T]; R^{d \times m})$$

is said to be a solution of Eq. (1) if $f(x(\cdot), y(\cdot), \cdot) \in \mathcal{M}^2([0, T]; R^d)$, $g(x(\cdot), \cdot) \in \mathcal{M}^2([0, T]; R^{d \times m})$ and Eq. (1) holds a.s. for every $t \in [0, T]$. A solution $\{x(t), y(t)\}_{0 \leq t \leq T}$ is said to be unique if for any other solution $\{\bar{x}(t), \bar{y}(t)\}_{0 \leq t \leq T}$ we have $P\{x(t) = \bar{x}(t), 0 \leq t \leq T\} = 1$.

In the field of control, $y(\cdot)$ is regarded as an adapted control and $x(\cdot)$ as the state of the system, in the sense that the choice of a control $y(\cdot)$ drives the state $x(\cdot)$ to the given target $x(T)$ at terminal time $t = T$.

As it was already mentioned, BSDEs are applied in finance where the pricing of a European claim is equivalent to solving the linear BSDE

$$x(t) = \xi - \int_t^T [r(s)x(s) + \theta(s)y(s)] ds - \int_t^T y(s) dw(s), \quad t \in [0, T],$$

for $x(T)$ representing the contingent claim, $r(t)$ the interest rate, $\theta(t)$ the risk premium and T the maturity time. However, the pricing of a European claim can be dependent on real parameters and on random excitations, and it can be modeled by the nonlinear BSDEs of the nonhomogeneous form. Because of this application also, it is important to estimate some approximation of the solution of BSDEs.

In [19, 20] the fundamental conditions of the existence and uniqueness of the solution to Eq. (1) are given.

Proposition 1.1. *If $f(0, 0, \cdot) \in \mathcal{M}^2([0, T]; R^d)$, $g(0, \cdot) \in \mathcal{M}^2([0, T]; R^{d \times m})$ and if f and g satisfy the uniform Lipschitz condition, that is, if there exists a positive constant $L > 0$ such that*

$$|f(x, y, t) - f(x', y', t)|^2 \leq L(|x - x'|^2 + |y - y'|^2) \quad a.s. \tag{2}$$

$$|g(x, t) - g(x', t)|^2 \leq L|x - x'|^2 \quad a.s., \tag{3}$$

for all $x, x' \in R^d$, $y, y' \in R^{d \times m}$ and $t \in [0, T]$, then there exists a unique solution $\{x(t), y(t)\}_{0 \leq t \leq T} \in \mathcal{M}^2([0, T]; R^d) \times \mathcal{M}_2([0, T]; R^{d \times m})$ to Eq. (1).

Moreover, if the following holds for final condition $x(T) \in L^p_{\mathcal{F}_T}(\Omega; R^d)$ and functions $f(0, 0, \cdot) \in \mathcal{M}^p([0, T]; R^d)$, $g(0, \cdot) \in \mathcal{M}^p([0, T]; R^{d \times m})$ for some $p \geq 2$, then $E|x(t)|^p < \infty$ for all $t \in [0, T]$ and $E \int_0^T |x(t)|^{p-2}|y(t)|^2 dt < \infty$.

Recall that the existence and uniqueness problems have been investigated under some other conditions, under non-Lipschitz conditions [18], for instance.

It is now well known that this Proposition 1.1 can be proved by the Picard iterative method which is based on the following. Let set $x_0(t) = y_0(t) \equiv 0, t \in [0, T]$, and for every $n \geq 1$ consider a pair $(x_n(t), y_n(t)) \in \mathcal{M}^2([0, T]; R^d) \times \mathcal{M}_2([0, T]; R^{d \times m})$ defined recursively defined by

$$x_{n+1}(t) = x(T) - \int_t^T f(x_n(s), y_n(s), s) ds - \int_t^T [g(x_n(s), s) + y_{n+1}(s)] dw(s), \quad t \in [0, T]. \tag{4}$$

Next, it is proved that this sequence of processes $(x_n, y_n)_{n \geq 1}$ converges to the process solution of (1).

This paper is devoted to generalize this method by some general approximation method which permit us to derive as special case a Z-algorithm. The paper is organized in the following way: In Section 2, we first formulate the problem, that is, we define a sequence of equations whose solutions represent approximations of the solution to the initial Eq. (1). Then, we present our main results—sufficient conditions under which the approximate solutions converge with probability one to the solution of Eq. (1). We also obtain straightforwardly an auxiliary result, i.e. the sequence of the approximate solutions converges in second order sense to the solution of the initial Eq. (1). In the same section, by similar methodology, but under some additional condition, the convergence of the approximate solutions in p th moment sense, $p \geq 2$, to the solution of the initial Eq. (1) is proven. In Section 3, we give some comments about algorithm and conclusions by which we introduce the notion of the Z-algorithm for Eq. (1). Section 4 is dedicated to remarks and examples which illustrate the previous theoretical considerations. At the end, we point out that the Picard method of iterations (4) is a special Z-algorithm.

2. Main results and their proofs

Together with Eq. (1) we consider the sequence of equations

$$\begin{aligned}
 x_{n+1}(t) &= \xi_{n+1} - \int_t^T f_n(x_{n+1}(s), y_{n+1}(s), s) ds \\
 &\quad - \int_t^T [g_n(x_{n+1}(s), s) + y_{n+1}(s)] dw(s), \quad t \in [0, T], \quad n \in \mathbb{N}_0,
 \end{aligned}
 \tag{5}$$

with final condition $\xi_{n+1} = x_{n+1}(T)$. We assume that the terminal conditions $\xi, \xi_{n+1} \in L^p_{\mathcal{F}_T}(\Omega; R^d)$ and that the functions $f_n, g_n, n \in \mathbb{N}_0$ are defined as f, g , respectively. We assume, with no special emphasis on conditions, that there exist unique solutions $\{x(t), y(t)\}_{0 \leq t \leq T}$ and $\{x_{n+1}(t), y_{n+1}(t)\}_{0 \leq t \leq T}$ to Eq. (1) and Eq. (5), respectively, satisfying $E \sup_{t \in [0, T]} |x(t)|^p < \infty$, $E \left(\int_0^T |y(t)|^2 dt \right)^{p/2} < \infty$ and $E \sup_{t \in [0, T]} |x_{n+1}(t)|^p < \infty$, $E \left(\int_0^T |y_{n+1}(t)|^2 dt \right)^{p/2} < \infty$, and that all the Lebesgue and Itô integrals employed further are well defined.

Obviously, it is quite natural to expect that if ξ_n, f_n, g_n are in some sense close to ξ, f, g respectively, then the sequence of the solutions $\{(x_{n+1}(t), y_{n+1}(t)), t \in [0, T], n \in \mathbb{N}_0\}$ to the Eqs. (5) will tend in some sense to the solution $\{(x(t), y(t)), t \in [0, T]\}$ of Eq. (1). In addition to the requirement that $\xi_n \rightarrow \xi, f_n(x, y, t) \rightarrow f(x, y, t), g_n(x, t) \rightarrow g(x, t)$ as $n \rightarrow +\infty$ uniformly in $[0, T] \times R^d \times R^{d \times m}$, and in accordance with [10], we seek that

$$\sum_{n=0}^{+\infty} E|\xi - \xi_{n+1}|^p < +\infty,
 \tag{6}$$

$$\sum_{n=0}^{+\infty} \sup_{(x,y,t) \in [0,T] \times R^d \times R^{d \times m}} |f(x, y, t) - f_n(x, y, t)|^p < +\infty,
 \tag{7}$$

$$\sum_{n=0}^{+\infty} \sup_{(x,t) \in [0,T] \times R^d} |g(x, t) - g_n(x, t)|^p < +\infty.
 \tag{8}$$

These conditions are essentially used to prove the main assertion in the mentioned paper.

Let us now introduce essential assumptions for our assertion:

\mathcal{A}_1 . Let $\xi, \xi_{n+1} \in L^p_{\mathcal{F}_T}(\Omega; R^d)$ and $\gamma_n := E|\xi - \xi_{n+1}|^p$, then

$$\sum_{n=0}^{+\infty} E|\xi - \xi_{n+1}|^p < +\infty \Leftrightarrow \sum_{n=0}^{+\infty} \gamma_n < +\infty.$$

\mathcal{A}_2 . For functions f, f_n let us define

$$\alpha_n := E \sup_{t \in [0, T]} |f(x_n(t), y_n(t), t) - f_n(x_n(t), y_n(t), t)|^p,$$

then

$$\sum_{n=0}^{+\infty} E \sup_{t \in [0, T]} |f(x_n(t), y_n(t), t) - f_n(x_n(t), y_n(t), t)|^p < +\infty \Leftrightarrow \sum_{n=0}^{+\infty} \alpha_n < +\infty.$$

\mathcal{A}_3 . For functions g, g_n let us define

$\beta_n := E \sup_{t \in [0, T]} |g(x_n(t), t) - g_n(x_n(t), t)|^p$, then

$$\sum_{n=0}^{+\infty} E \sup_{t \in [0, T]} |g(x_n(t), t) - g_n(x_n(t), t)|^p < +\infty \Leftrightarrow \sum_{n=0}^{+\infty} \beta_n < +\infty.$$

In the sequel, in order to obtain simpler notation, we will for $p \geq 2$ use notation

$$\begin{aligned} \gamma_n &:= E|\xi - \xi_{n+1}|^p \Rightarrow \sum_{n=0}^{+\infty} \gamma_n < +\infty, \\ \alpha_n &:= E \sup_{t \in [0, T]} |f(x, y, t) - f_n(x, y, t)|^p \Rightarrow \sum_{n=0}^{+\infty} \alpha_n < +\infty, \\ \beta_n &:= E \sup_{t \in [0, T]} |g(x, t) - g_n(x, t)|^p \Rightarrow \sum_{n=0}^{+\infty} \beta_n < +\infty. \end{aligned}$$

If we subtract Eq. (5) from Eq. (1), we find for $t \in [0, T]$ that

$$\begin{aligned} x(t) - x_{n+1}(t) &= \xi - \xi_{n+1} \\ &\quad - \int_t^T [f(x(s), y(s), s) - f_n(x_{n+1}(s), y_{n+1}(s), s)] ds \\ &\quad - \int_t^T [g(x(s), s) + y(s) - g_n(x_{n+1}(s), s) - y_{n+1}(s)] dw(s), \end{aligned}$$

and let us denote that

$$\begin{aligned} R_{n+1}(t) &= x(t) - x_{n+1}(t), \quad P_{n+1}(t) = y(t) - y_{n+1}(t), \\ \Delta f_n(x_n, y_n, s) &= |f(x_n(s), y_n(s), s) - f_n(x_n(s), y_n(s), s)|, \\ \Delta g_n(x_n, s) &= |g(x_n(s), s) - g_n(x_n(s), s)|. \end{aligned}$$

Theorem 2.1. Let $\xi, \xi_n \in L^2_{\mathcal{F}_T}(\Omega; \mathbb{R}^d)$, $n \in \mathbb{N}$, let also the functions f, f_n, g, g_n , $n \in \mathbb{N}_0$ satisfy the Lipschitz conditions (2) and (3) with constant $L > 0$, and assumptions $\mathcal{A}_1 - \mathcal{A}_3$ be satisfied for $p = 2$. Then, the sequence of the solutions $\{(x_{n+1}(t), y_{n+1}(t)), t \in [0, T], n \in \mathbb{N}_0\}$ to the Eqs. (5) satisfies that $x_{n+1} \xrightarrow{S^2} y, y_{n+1} \xrightarrow{M^2} z$ as $n \rightarrow \infty$, where $\{(x(t), y(t)), t \in [0, T]\}$ is a solution of Eq. (1).

Proof. If we apply the Itô formula to $|R_{n+1}(t)|^2$, we find for $t \in [0, T]$ that

$$\begin{aligned} |R_{n+1}(t)|^2 &= |\xi - \xi_{n+1}|^2 \\ &\quad - 2 \int_t^T R_{n+1}^T(s) \times [f(x(s), y(s), s) - f_n(x_{n+1}(s), y_{n+1}(s), s)] ds \\ &\quad - \int_t^T |g(x(s), s) + y(s) - g_n(x_{n+1}(s), s) - y_{n+1}(s)|^2 ds \\ &\quad - 2 \int_t^T R_{n+1}^T(s) \times [g(x(s), s) + y(s) - g_n(x_{n+1}(s), s) - y_{n+1}(s)] dw(s). \end{aligned}$$

This implies that

$$\begin{aligned} &|R_{n+1}(t)|^2 + \int_t^T |y(s) - y_{n+1}(s)|^2 ds \\ &\leq |\xi - \xi_{n+1}|^2 - 2 \int_t^T R_{n+1}^T(s) \times [f(x(s), y(s), s) - f_n(x_{n+1}(s), y_{n+1}(s), s)] ds \\ &\quad - 2 \int_t^T \text{trace} [[g(x(s), s) - g_n(x_{n+1}(s), s)]^T (y(s) - y_{n+1}(s))] ds \\ &\quad - 2 \int_t^T R_{n+1}^T(s) \times [g(x(s), s) + y(s) - g_n(x_{n+1}(s), s) - y_{n+1}(s)] dw(s) \\ &= |\xi - \xi_{n+1}|^2 + S_1(t) + S_2(t) + S_3(t). \end{aligned} \tag{9}$$

We will separately estimate each term on the right-hand side of the last inequality. In order to obtain estimates, we will use several times elementary inequalities:

1. $-2ab \leq a^2/\lambda_i + \lambda_i b^2$, where $\lambda_i > 0$ is a constant;
2. $(\sum_{i=1}^n a_i)^p \leq n^{p-1} \sum_{i=1}^n a_i^p$, $a_i \geq 0$, $p \in \mathbb{N}$.

From

$$\begin{aligned} & -2R_{n+1}^T(s)[f(x(s), y(s), s) - f_n(x_{n+1}(s), y_{n+1}(s), s)] \\ & \leq \lambda_1 |R_{n+1}(s)|^2 + \frac{1}{\lambda_1} |f(x(s), y(s), s) - f_n(x_{n+1}(s), y_{n+1}(s), s)|^2 \\ & \leq \left(\lambda_1 + \frac{4L}{\lambda_1}\right) |R_{n+1}(s)|^2 + \frac{8L}{\lambda_1} |P_n(s)|^2 \\ & \quad + \frac{8L}{\lambda_1} |R_n(s)|^2 + \frac{4L}{\lambda_1} |P_{n+1}(s)|^2 + \frac{4}{\lambda_1} \Delta f_n^2(x_n, y_n, s), \end{aligned}$$

we have that

$$\begin{aligned} S_1(t) & \leq \left(\lambda_1 + \frac{4L}{\lambda_1}\right) \int_t^T |R_{n+1}(s)|^2 ds + \frac{8L}{\lambda_1} \int_t^T |P_n(s)|^2 ds + \frac{8L}{\lambda_1} \int_t^T |R_n(s)|^2 ds \\ & \quad + \frac{4L}{\lambda_1} \int_t^T |P_{n+1}(s)|^2 ds + \frac{4}{\lambda_1} \int_t^T \Delta(f_n(x_n, y_n, s))^2 ds. \end{aligned} \tag{10}$$

The second term in (9) can be estimated by repeating completely the previous procedure. It follows that

$$\begin{aligned} & -2\text{trace}\left[[g(x(s), s) - g_n(x_{n+1}(s), s)]^T (y(s) - y_{n+1}(s))\right] \\ & \leq \lambda_2 |g(x(s), s) - g_n(x_{n+1}(s), s)|^2 + \frac{1}{\lambda_2} |y_{n+1}(s) - y(s)|^2 \\ & \leq \frac{1}{\lambda_2} |P_{n+1}(s)|^2 + 8L\lambda_2 |R_n(s)|^2 + 4\lambda_2 |R_{n+1}(s)|^2 + 4\lambda_2 (\Delta g_n(x_n, s))^2. \end{aligned}$$

Then,

$$S_2(t) \leq 4\lambda_2 \int_t^T |R_{n+1}(s)|^2 ds + 8L\lambda_2 \int_t^T |R_n(s)|^2 ds + \frac{1}{\lambda_2} \int_t^T |P_{n+1}(s)|^2 ds + 4\lambda_2 \int_t^T (\Delta g_n(x_n))^2 ds. \tag{11}$$

From (9), (10) and (11) we have that

$$\begin{aligned} & |R_{n+1}(t)|^2 + \int_t^T |P_{n+1}(s)|^2 ds \\ & \leq |\xi - \xi_{n+1}|^2 \\ & \quad + \left(\lambda_1 + \frac{4L}{\lambda_1} + 4\lambda_2\right) \int_t^T |R_{n+1}(s)|^2 ds + \frac{8L}{\lambda_1} \int_t^T |P_n(s)|^2 ds \\ & \quad + \left(\frac{8L}{\lambda_1} + 8L\lambda_2\right) \int_t^T |R_n(s)|^2 ds + \left(\frac{4L}{\lambda_1} + \frac{1}{\lambda_2}\right) \int_t^T |P_{n+1}(s)|^2 ds \\ & \quad + \frac{4}{\lambda_1} \int_t^T (\Delta f_n(x_n, y_n, s))^2 ds + 4\lambda_2 \int_t^T (\Delta g_n(x_n, s))^2 ds + S_3(t). \end{aligned} \tag{12}$$

Let us for $t_0 \in [0, T]$ define

$$I_{n+1}(t_0) := E \sup_{t \in [t_0, T]} |R_{n+1}(t)|^2, \quad J_{n+1}(t_0) := E \int_{t_0}^T |P_{n+1}(s)|^2 ds.$$

Further, let us rearrange the $E \sup_{t \in [t_0, T]} S_3(t)$ in the following way,

$$\begin{aligned} S_3 &= E \sup_{t \in [t_0, T]} S_3(t) = E \sup_{t \in [t_0, T]} \left(- \int_t^T \dots dw(s) \right) \\ &= E \sup_{t \in [t_0, T]} \left(- \int_{t_0}^T \dots dw(s) + \int_{t_0}^t \dots dw(s) \right) \\ &= E \sup_{t \in [t_0, T]} \left(\int_{t_0}^t \dots dw(s) \right). \end{aligned}$$

The application of the Burkholder-Davis-Gundy inequality [13, 19] and the procedures used above, yield

$$\begin{aligned} S_3 &\leq 8E \left| \int_{t_0}^T |R_{n+1}(s)|^2 |g(x(s), s) - g_n(x_{n+1}(s), s) + y(s) - y_{n+1}(s)|^2 ds \right| \\ &\leq 8E \left(\sup_{s \in [t_0, T]} |R_{n+1}(s)|^2 \int_{t_0}^T |g(x(s), s) - g_n(x_{n+1}(s), s) + y(s) - y_{n+1}(s)|^2 ds \right)^{\frac{1}{2}} \\ &\leq \frac{1}{2} I_{n+1}(t_0) + 32E \int_{t_0}^T |g(x(s), s) - g_n(x_{n+1}(s), s) + y(s) - y_{n+1}(s)|^2 ds \\ &\leq \frac{1}{2} I_{n+1}(t_0) + 512L \int_{t_0}^T I_n(s) ds + 256L \int_{t_0}^T I_{n+1}(s) ds \\ &\quad + 256T\beta_n + 2 \int_{t_0}^T J_{n+1}(s) ds. \end{aligned} \tag{13}$$

If we substitute estimate (13) in (12), and take $E \sup_{t \in [t_0, T]}$ over whole inequality (12) we have

$$\begin{aligned} I_{n+1}(t_0) + J_{n+1}(t_0) &\leq \gamma_n + \frac{1}{2} I_{n+1}(t_0) \\ &\quad + \left(\lambda_1 + \frac{4L}{\lambda_1} + 4\lambda_2 + 256L \right) \int_{t_0}^T I_{n+1}(s) ds + \frac{8L}{\lambda_1} J_n(t_0) \\ &\quad + \left(\frac{8L}{\lambda_1} + 8L\lambda_2 + 512L \right) \int_{t_0}^T I_n(s) ds + \left(\frac{4L}{\lambda_1} + \frac{1}{\lambda_2} + 2 \right) J_{n+1}(t_0) \\ &\quad + \frac{4T}{\lambda_1} \alpha_n + (4T\lambda_2 + 256T)\beta_n. \end{aligned} \tag{14}$$

Having in mind that control processes y and y_n are from $\mathcal{M}^2([0, T]; R^{d \times m})$, there exists some $M > 0$ such that $E \int_{t_0}^T |P_{n+1}(s)|^2 ds < M$ (also $\int_{t_0}^T J_{n+1}(s) ds < M$, and further $\int_{t_0}^T J_{n+1}(t_0) ds = (T - t_0)J_{n+1}(t_0) < M$). For an arbitrary partition of segment $[t_0, T]$ we have

$$E \int_{t_0}^T |P_{n+1}(s)|^2 ds = E \lim_{m \rightarrow \infty} \sum_m |P_{n+1}(s_m)|^2 \Delta_m.$$

Then for each $m \in \mathbb{N}$

$$E \lim_{m \rightarrow \infty} \sum_m |P_{n+1}(s_m)|^2 \Delta_m \geq E |P_{n+1}(s_m)|^2 \Delta_m \frac{J_{n+1}(t_0)}{J_{n+1}(t_0)}, J_{n+1}(t_0) \neq 0$$

so there exists some constant $r > 0$ such that

$$rJ_{n+1}(t_0) \leq E \int_{t_0}^T |P_{n+1}(s)|^2 ds.$$

This could be proven in another way also. From the assumptions of the Theorem, for every $n \in \mathbb{N}_0$ we have $\text{thT}(x_n, y_n) \neq (x, y)$, so there exists $\inf_n |P_{n+1}(s_m)|^2$, and let $r_1 = E \inf_n |P_{n+1}(s_m)|^2$. Further

$$E \int_{t_0}^T |P_{n+1}(s)|^2 ds \geq r_1(T - t_0) \frac{J_{n+1}(t_0)}{J_{n+1}(t_0)}, \quad J_{n+1}(t_0) \neq 0$$

so there exists some constant $r > 0$ such that

$$rJ_{n+1}(t_0) \leq E \int_{t_0}^T |P_{n+1}(s)|^2 ds.$$

(If $J_{n+1}(t_0) = 0$ then we chose $r = 1$, and we can always add zero member to any expression.)

For some constants k_1, k'_1 we have

$$\begin{aligned} I_{n+1}(t_0) + J_{n+1}(t_0) &\leq \gamma_n + \frac{1}{2} I_{n+1}(t_0) \\ &+ \left(\lambda_1 + \frac{4L}{\lambda_1} + 4\lambda_2 + 256L \right) \int_{t_0}^T I_{n+1}(s) ds + \frac{8Lk_1}{\lambda_1} J_n(t_0) \\ &+ \left(\frac{8L}{\lambda_1} + 8L\lambda_2 + 512L \right) \int_{t_0}^T I_n(s) ds + k'_1 \left(\frac{4L}{\lambda_1} + \frac{1}{\lambda_2} + 2 \right) J_{n+1}(t_0) \\ &+ \frac{4T}{\lambda_1} \alpha_n + (4T\lambda_2 + 256T)\beta_n. \end{aligned} \tag{15}$$

If we define for every $t \in [0, T]$

$$U_{n+1}(t) := I_{n+1}(t) + J_{n+1}(t), \tag{16}$$

from (15) we obtain

$$\frac{1}{2} U_{n+1}(t_0) \leq \gamma_n + k_2 \int_{t_0}^T U_n(s) ds + k_3 \int_{t_0}^T U_{n+1}(s) ds + \frac{4T}{\lambda_1} \alpha_n + (4T\lambda_2 + 256T)\beta_n, \tag{17}$$

where

$$\begin{aligned} k_2 &= \max \left\{ \frac{8L}{\lambda_1} + 8L\lambda_2 + 512L, \frac{8Lk_1}{\lambda_1} \right\}, \\ k_3 &= \max \left\{ \lambda_1 + \frac{4L}{\lambda_1} + 4\lambda_2 + 256L, k'_1 \left(\frac{4L}{\lambda_1} + \frac{1}{\lambda_2} + 2 \right) \right\}. \end{aligned}$$

Applying well known Gronwall-Bellman's inequality ([1], Theorem 1.5) on (17): Let $u(t)$ be a continuous function in $[\alpha, \beta]$, $a(t)$ be Riemann integrable function in $[\alpha, \beta]$ and $c = \text{const} > 0$. If $u(t) \leq a(t) + c \int_t^\beta u(s) ds$, $t \in [\alpha, \beta]$, then $u(t) \leq a(t) + c \int_t^\beta a(s) e^{c(s-t)} ds$, $t \in [\alpha, \beta]$. It follows that

$$\begin{aligned} U_{n+1}(t_0) &\leq \frac{k_2}{2} \int_{t_0}^T U_n(s) ds + \frac{1}{2} \left(\gamma_n + \frac{4T}{\lambda_1} \alpha_n + (4T\lambda_2 + 256T)\beta_n \right) \\ &+ \frac{k_3}{2} \int_{t_0}^T \left[\frac{k_2}{2} \int_s^T U_n(r) dr + \frac{1}{2} \left(\gamma_n + \frac{4T}{\lambda_1} \alpha_n + (4T\lambda_2 + 256T)\beta_n \right) \right] e^{\frac{k_3}{2}(s-t_0)} ds. \end{aligned}$$

Since

$$\frac{k_2 k_3}{4} \int_{t_0}^T e^{\frac{k_3}{2}(s-t_0)} \int_s^T U_n(r) dr ds = \frac{k_2}{2} \int_{t_0}^T e^{\frac{k_3}{2}(s-t_0)} (U_n(s) - 1) ds,$$

we obtain

$$U_{n+1}(t_0) \leq \frac{1}{2} \left(\gamma_n + \frac{4T}{\lambda_1} \alpha_n + (4T\lambda_2 + 256T)\beta_n \right) + \frac{1}{k_3} \left(\gamma_n + \frac{4T}{\lambda_1} \alpha_n + (4T\lambda_2 + 256T)\beta_n \right) \left(e^{\frac{k_3}{2}T} - 1 \right) + \frac{k_2}{2} \int_{t_0}^T e^{\frac{k_3}{2}(s-t_0)} U_n(s) ds.$$

For $S_m(t_0) = \sum_{n=0}^m U_n(t_0)$, $m \geq 0$ we have that

$$S_m(t_0) - U_0(t_0) \leq S_{m+1}(t_0) - U_0(t_0) \leq a \sum_{n=0}^m \gamma_n + b \sum_{n=0}^m \epsilon_n + c \int_{t_0}^T S_m(s) ds,$$

for a, b, c generic constants, and $\epsilon_n = \frac{4T}{\lambda_1} \alpha_n + (4T\lambda_2 + 256T)\beta_n$. Then

$$S_m(t_0) \leq U_0(0) + a \sum_{n=0}^m \gamma_n + b \sum_{n=0}^m \epsilon_n + c \int_{t_0}^T S_m(s) ds.$$

If we apply Gronwall-Bellman's inequality ([1], Theorem 1.5) on last inequality, we obtain

$$S_m(t_0) \leq U_0(0) + a \sum_{n=0}^m \gamma_n + b \sum_{n=0}^m \epsilon_n + c \int_{t_0}^T \left(U_0(0) + a \sum_{n=0}^m \gamma_n + b \sum_{n=0}^m \epsilon_n \right) e^{c(s-t_0)} ds = \left(U_0(0) + a \sum_{n=0}^m \gamma_n + b \sum_{n=0}^m \epsilon_n \right) e^{c(T-t_0)}, \quad t_0 \in [0, T]. \tag{18}$$

If we take $t_0 = 0$, from (18) it follows

$$S_m(0) \leq \left(U_0(0) + a \sum_{n=0}^m \gamma_n + b \sum_{n=0}^m \epsilon_n \right) e^{cT}.$$

From assumptions $\mathcal{A}_1 - \mathcal{A}_3$ it follows that

$$\sum_{n=0}^{\infty} U_n(0) = \lim_{m \rightarrow +\infty} S_m(0) < +\infty. \tag{19}$$

Regarding the convergence, it follows that

$$\sum_{n=0}^{\infty} U_n(0) \Rightarrow \lim_{n \rightarrow \infty} U_n(0) = 0,$$

which is equivalent with

$$\begin{aligned} U_{n+1}(0) &= 0 \\ \Leftrightarrow I_{n+1}(0) + J_{n+1}(0) &= 0 \\ \Leftrightarrow E \sup_{t \in [0, T]} |x(t) - x_{n+1}(t)|^2, E \int_{t_0}^T |y(t) - y_{n+1}(t)|^2 ds, \end{aligned}$$

which completes the proof. \square

The convergence of the defined sequence can be proven in a higher order sense, but with additional condition for the closenesses of the control processes of the solutions of the Eqs. (5) and Eq. (1). The following theorem illustrates this result.

Theorem 2.2. Let $\xi, \xi_n \in L^2_{\mathcal{F}_T}(\Omega; \mathbb{R}^d)$, $n \in \mathbb{N}$, and let the functions f, f_n, g, g_n , $n \in \mathbb{N}_0$ satisfy the Lipschitz conditions (2) and (3) with constant $L > 0$. Let also assumptions $\mathcal{A}_1 - \mathcal{A}_3$ be satisfied for $p \geq 2$. We assume that there exist unique solutions $\{(x_{n+1}(t), y_{n+1}(t)), t \in [0, T], n \in \mathbb{N}_0\}$ and $\{(x(t), y(t)), t \in [0, T]\}$ to the Eqs. (5) and Eq. (1), respectively. Then,

$$E \sup_{t \in [0, T]} |x_n(s) - x(s)|^p \rightarrow 0, \quad n \rightarrow +\infty,$$

$$E \sup_{t \in [0, T]} |y_n(s) - y(s)|^p \rightarrow 0, \quad n \rightarrow +\infty.$$

Proof. Similarly as in Theorem 2.1, if we apply the Itô formula to $|R_n(t)|^p, p \geq 2$, we find for $t \in [0, T]$ that

$$\begin{aligned} |R_{n+1}(t)|^p &= |\xi - \xi_{n+1}|^p - p \int_t^T |R_{n+1}(s)|^{p-2} R_{n+1}^T(s) \\ &\quad \times [f(x(s), y(s), s) - f_n(x_{n+1}(s), y_{n+1}(s), s)] ds \\ &\quad - \frac{p}{2} \int_t^T |R_{n+1}(s)|^{p-2} |g(x(s), s) + y(s) - g_n(x_{n+1}(s), s) - y_{n+1}(s)|^2 ds \\ &\quad - \frac{p(p-2)}{2} \int_t^T |R_{n+1}(s)|^{p-4} \times |R_{n+1}^T(s) [g(x(s), s) + y(s) - g_n(x_{n+1}(s), s) - y_{n+1}(s)]|^2 ds \\ &\quad - p \int_t^T |R_{n+1}(s)|^{p-2} R_{n+1}^T(s) \times [g(x(s), s) + y(s) - g_n(x_{n+1}(s), s) - y_{n+1}(s)] dw(s). \end{aligned}$$

This implies that

$$\begin{aligned} &|R_{n+1}(t)|^p + \frac{p}{2} \int_t^T |R_{n+1}(s)|^{p-2} |y(s) - y_{n+1}(s)|^2 ds \\ &\leq |\xi - \xi_{n+1}|^p - p \int_t^T |R_{n+1}(s)|^{p-2} R_{n+1}^T(s) \times [f(x(s), y(s), s) - f_n(x_{n+1}(s), y_{n+1}(s), s)] ds \\ &\quad - p \int_t^T |R_{n+1}(s)|^{p-2} \times \text{trace} [[g(x(s), s) - g_n(x_{n+1}(s), s))]^T (y(s) - y_{n+1}(s))] ds \\ &\quad - p \int_t^T |R_{n+1}(s)|^{p-2} R_{n+1}^T(s) \times [g(x(s), s) + y(s) - g_n(x_{n+1}(s), s) - y_{n+1}(s)] dw(s) \\ &= |\xi - \xi_{n+1}|^p + T_1(t) + T_2(t) + T_3(t). \end{aligned} \tag{20}$$

We will separately estimate each term on the right-hand side of the (20) using familiar inequalities similar as in Theorem 2.1.

From

$$\begin{aligned} &-R_{n+1}^T(s) [f(x(s), y(s), s) - f_n(x_{n+1}(s), y_{n+1}(s), s)] \\ &\leq \frac{\lambda_1}{2} |R_{n+1}(s)|^2 + \frac{1}{2\lambda_1} |f(x(s), y(s), s) - f_n(x_{n+1}(s), y_{n+1}(s), s)|^2 \\ &\leq \left(\frac{\lambda_1}{2} + \frac{2L}{\lambda_1} \right) |R_{n+1}(s)|^2 + \frac{4L}{\lambda_1} |P_n(s)|^2 \\ &\quad + \frac{4L}{\lambda_1} |R_n(s)|^2 + \frac{2L}{\lambda_1} |P_{n+1}(s)|^2 + \frac{2}{\lambda_1} (\Delta f_n(x_n, y_n, s))^2. \end{aligned}$$

Likewise, the inequality $a^{p-2}b^2 \leq \frac{p-2}{p} a^p + \frac{2}{p} b^p$, $a, b \geq 0$, $p \geq 2$, will be used to estimate the first term in (20), as well as, without special emphasis, more times in the sequel. Then,

$$\begin{aligned}
 T_1(t) &\leq \left(\frac{p\lambda_1}{2} + \frac{2pL}{\lambda_1} + \frac{8L(p-2)}{\lambda_1} + \frac{2(p-2)}{\lambda_1} \right) \int_t^T |R_{n+1}(s)|^p ds \\
 &\quad + \frac{8L}{\lambda_1} \int_t^T |P_n(s)|^p ds + \frac{8L}{\lambda_1} \int_t^T |R_n(s)|^p ds \\
 &\quad + \frac{2pL}{\lambda_1} \int_t^T |R_{n+1}(s)|^{p-2} |P_{n+1}(s)|^2 ds + \frac{4}{\lambda_1} \int_t^T \Delta(f_n(x_n, y_n, s))^p ds.
 \end{aligned} \tag{21}$$

The second term in (20) can be estimated by repeating completely the previous procedure. By applying the above elementary inequality 1 from Theorem 1 (for some constant λ_2) and by using (3), one can see that

$$\begin{aligned}
 &-\text{trace} \left[[g(x(s), s) - g_n(x_{n+1}(s), s)]^T (y(s) - y_{n+1}(s)) \right] \\
 &\leq \frac{\lambda_2}{2} |g(x(s), s) - g_n(x_{n+1}(s), s)|^2 + \frac{1}{2\lambda_2} |y_{n+1}(s) - y(s)|^2 \\
 &\leq \frac{1}{2\lambda_2} |P_{n+1}(s)|^2 + 4L\lambda_2 |R_n(s)|^2 + 2L\lambda_2 |R_{n+1}(s)|^2 + 2\lambda_2 (\Delta g_n(x_n, s))^2.
 \end{aligned}$$

Then,

$$\begin{aligned}
 T_2(t) &\leq (4(p-2)L\lambda_2 + 2L\lambda_2 p + 2(p-2)\lambda_2) \int_t^T |R_{n+1}(s)|^p + 8L\lambda_2 \int_t^T |R_n(s)|^p ds \\
 &\quad + \frac{p}{2\lambda_2} \int_t^T |R_{n+1}(s)|^{p-2} |P_{n+1}(s)|^2 ds + 4\lambda_2 \int_t^T (\Delta g_n(x_n, s))^p ds.
 \end{aligned} \tag{22}$$

From (20), (21) and (22) one can see that

$$\begin{aligned}
 &|R_{n+1}(t)|^p + \frac{p}{2} \int_t^T |R_{n+1}(s)|^{p-2} |P_{n+1}(s)|^2 ds \\
 &\leq |\xi - \xi_{n+1}|^p + \left(\frac{p\lambda_1}{2} + \frac{2pL}{\lambda_1} + \frac{2(4L+1)(p-2)}{\lambda_1} + 2(2L+1)(p-2)\lambda_2 + 2L\lambda_2 p \right) \int_t^T |R_{n+1}(s)|^p ds \\
 &\quad + \left(\frac{8L}{\lambda_1} + 8L\lambda_2 \right) \int_t^T |R_n(s)|^p ds \\
 &\quad + \left(\frac{2pL}{\lambda_1} + \frac{p}{2\lambda_2} \right) \int_t^T |R_{n+1}(s)|^{p-2} |P_{n+1}(s)|^2 ds \\
 &\quad + \frac{4}{\lambda_1} \int_t^T (\Delta f_n(x_n, y_n, s))^p ds + 4\lambda_2 \int_t^T (\Delta g_n(x_n, s))^p ds + T_3(t).
 \end{aligned} \tag{23}$$

Let us for $t_0 \in [0, T]$ define

$$I_{n+1}(t_0) := E \sup_{t \in [t_0, T]} |R_{n+1}(t)|^p, \quad J_{n+1}(t_0) := E \sup_{t \in [t_0, T]} |P_{n+1}(t)|^p.$$

Similar as in Theorem 2.1, the application of the Burkholder-Davis-Gundy inequality [13, 19], yield

$$\begin{aligned}
 T_3 &= \sup_{t \in [t_0, T]} T_3(t) \leq 4\sqrt{2}p E \left(\int_{t_0}^T |R_{n+1}(s)|^{2p-2} \times |g(x(s), s) - g_n(x_{n+1}(s), s) + y(s) - y_{n+1}(s)|^2 ds \right)^{\frac{1}{2}} \\
 &\leq 4\sqrt{2}p E \left(\sup_{s \in [t_0, T]} |R_{n+1}(s)|^p \int_{t_0}^T |R_{n+1}(s)|^{p-2} \times |g(x(s), s) - g_n(x_{n+1}(s), s) + y(s) - y_{n+1}(s)|^2 ds \right)^{\frac{1}{2}} \\
 &\leq \frac{1}{2} I_{n+1}(t_0) + (16p(p-2) + 256p) \int_{t_0}^T I_{n+1}(s) ds + 256pE \int_{t_0}^T (\Delta g_n(x_n, s))^p ds \\
 &\quad + 512p \int_{t_0}^T I_n(s) ds + 64pE \int_{t_0}^T J_{n+1}(s) ds.
 \end{aligned} \tag{24}$$

From (23) and (24) it follows

$$\begin{aligned}
 I_{n+1}(t_0) &+ \frac{p}{2} E \sup_{[t_0, T]} \int_t^T |R_{n+1}(s)|^{p-2} |P_{n+1}(s)|^2 ds \\
 &\leq E|\xi - \xi_{n+1}|^p + \left(\frac{p\lambda_1}{2} + \frac{2pL}{\lambda_1} \right. \\
 &\quad \left. + \frac{2(4L+1)(p-2)}{\lambda_1} + 2(2L+1)(p-2)\lambda_2 + 2L\lambda_2 p + 256p \right) \int_{t_0}^T I_{n+1}(s) ds \\
 &\quad + \frac{8L}{\lambda_1} \int_{t_0}^T J_n(s) ds + 64p \int_{t_0}^T J_{n+1}(s) ds + \left(\frac{8L}{\lambda_1} + 8L\lambda_2 + 512p \right) \int_{t_0}^T I_n(s) ds \\
 &\quad + \left(\frac{2pL}{\lambda_1} + \frac{p}{2\lambda_2} \right) E \sup_{[t_0, T]} \int_t^T |R_{n+1}(s)|^{p-2} |P_{n+1}(s)|^2 ds \\
 &\quad + \frac{4}{\lambda_1} \int_{t_0}^T (\Delta f_n(x_n, y_n, s))^p ds + (4\lambda_2 + 256p) \int_{t_0}^T (\Delta g_n(x_n, s))^p ds + \frac{1}{2} I_{n+1}(t_0).
 \end{aligned}$$

We choose λ_1, λ_2 such that $\frac{2pL}{\lambda_1} + \frac{p}{2\lambda_2} < \frac{p}{2}$, per example $\lambda_1 = 16L$ and $\lambda_2 = 4$, from last inequality we have

$$\begin{aligned}
 \frac{1}{2} I_{n+1}(t_0) &\leq E|\xi - \xi_{n+1}|^p + K_1(p) \int_{t_0}^T I_{n+1}(s) ds \\
 &\quad + \frac{1}{2} \int_{t_0}^T J_n(s) ds + 64p \int_{t_0}^T J_{n+1}(s) ds + K_2(p) \int_{t_0}^T I_n(s) ds \\
 &\quad + \frac{1}{4L} \int_{t_0}^T E \sup_{[t_0, T]} (\Delta f_n(x_n, y_n, s))^p ds + K_3(p) \int_{t_0}^T E \sup_{[t_0, T]} (\Delta g_n(x_n, s))^p ds.
 \end{aligned} \tag{25}$$

for

$$\begin{aligned}
 K_1(p) &= 16pL + \frac{2049p}{8} + (p-2) \left(\frac{4L+1}{8L} + 8(2L+1) \right) + 256p, \\
 K_2(p) &= \frac{1}{2} + 32L + 512p \\
 K_3(p) &= 16(1 + 16p).
 \end{aligned}$$

So from (25) we have that

$$\begin{aligned} & \frac{1}{2}I_{n+1}(t_0) + \frac{1}{2}J_{n+1}(t_0) \\ & \leq E|\xi - \xi_{n+1}|^p + K_1(p) \int_{t_0}^T I_{n+1}(s) ds + K_2(p) \int_{t_0}^T I_n(s) ds \\ & \quad + \frac{1}{2} \int_{t_0}^T J_n(s) ds + (64p + \frac{1}{2}) \int_{t_0}^T J_{n+1}(s) ds \\ & \quad + \frac{T}{4L} E \sup_{[t_0, T]} \Delta(f_n(x_n, y_n, s))^p ds + TK_3(p) E \sup_{[t_0, T]} (\Delta g_n(x_n, s))^p. \end{aligned}$$

For $U_n(t) = I_n(t) + J_n(t)$ we have

$$\frac{1}{2}U_{n+1}(t_0) \leq E|R_{n+1}(T)|^p + \widetilde{K}_1(p) \int_{t_0}^T U_{n+1}(s) ds + \widetilde{K}_2(p) \int_{t_0}^T U_n(s) ds + T\widetilde{K}_3(p)\varepsilon_n,$$

where

$$\begin{aligned} \widetilde{K}_1 &= \min\{0.5, \bar{r}\}, \quad \widetilde{K}_1(p) = \max\{K_1(p), (64p + 0.5)\}, \quad \widetilde{K}_2(p) = \max\{0.5, K_2(p)\} = K_2(p), \\ \widetilde{K}_3(p) &= \max\{1/4L, K_3(p)\}, \quad \varepsilon_n = E \sup_{[t_0, T]} [(\Delta f_n(x_n, y_n, s))^p + (\Delta g_n(x_n, s))^p]. \end{aligned}$$

By applying the well-known Gronwall-Bellman inequality, we obtain

$$\begin{aligned} U_{n+1}(t_0) & \leq 2 \left(\gamma_n + \widetilde{K}_3(p) \int_{t_0}^T U_n(s) ds + 2T\widetilde{K}_4(p)\varepsilon_n \right) \\ & \quad + 2\widetilde{K}_2(p) \int_{t_0}^T \left(2\gamma_n + 2\widetilde{K}_3(p) \int_s^T U_n(r) dr + 2T\widetilde{K}_4(p)\varepsilon_n \right) e^{2\widetilde{K}_2(p)(s-t_0)} ds. \end{aligned}$$

Since

$$4\widetilde{K}_2(p)\widetilde{K}_3(p) \int_{t_0}^T e^{2\widetilde{K}_2(p)(s-t_0)} \int_s^T U_n(r) dr ds = 2\widetilde{K}_3(p) \int_{t_0}^T U_n(s) \left(e^{2\widetilde{K}_2(p)(s-t_0)} - 1 \right) ds,$$

than

$$\begin{aligned} U_{n+1}(t_0) & \leq 2 \left(\gamma_n + \widetilde{K}_3(p) \int_{t_0}^T U_n(s) ds + 2T\widetilde{K}_4(p)\varepsilon_n \right) \\ & \quad + 2 \left(\gamma_n + 2T\widetilde{K}_4(p)\varepsilon_n \right) \left(e^{2\widetilde{K}_2(p)(T-t_0)} - 1 \right) \\ & \quad + 2\widetilde{K}_3(p) \int_{t_0}^T U_n(s) \left(e^{2\widetilde{K}_2(p)(s-t_0)} - 1 \right) ds. \end{aligned}$$

This is further

$$\begin{aligned} U_{n+1}(t_0) & \leq 2 \left(\gamma_n + 2T\widetilde{K}_4(p)\varepsilon_n \right) \\ & \quad + 2 \left(\gamma_n + 2T\widetilde{K}_4(p)\varepsilon_n \right) \left(e^{2\widetilde{K}_2(p)T} - 1 \right) + 2\widetilde{K}_3(p) \int_{t_0}^T U_n(s) e^{\frac{\widetilde{K}_2(p)}{\widetilde{K}_1}(s-t_0)} ds. \end{aligned} \tag{26}$$

For $S_m(t_0) = \sum_{n=0}^m U_n(t_0)$, $m \geq 0$ we have that

$$S_m(t_0) - U_0(t_0) \leq S_{m+1}(t_0) - U_0(t_0) \leq \bar{a} \sum_{n=0}^m \gamma_n + \bar{b} \sum_{n=0}^m \varepsilon_n + \bar{c} \int_{t_0}^T S_m(s) ds,$$

where $\bar{a}, \bar{b}, \bar{c}$ are generic constants, Then

$$S_m(t_0) \leq U_0(0) + \bar{a} \sum_{n=0}^m \gamma_n + \bar{b} \sum_{n=0}^m \varepsilon_n + \bar{c} \int_{t_0}^T S_m(s) ds,$$

which is

$$S_m(t_0) \leq U_0(0) + \bar{a} \sum_{n=0}^m \gamma_n + \bar{b} \sum_{n=0}^m \varepsilon_n + \bar{c} \int_{t_0}^T S_m(s) ds.$$

By applying Gronwall-Bellman inequality once again, the conclusion follows in the same way as in Theorem 2.1, than for $t_0 = 0$

$$S_m(0) \leq \left(U_0(0) + \bar{a} \sum_{n=0}^m \gamma_n + \bar{b} \sum_{n=0}^m \varepsilon_n \right) e^{\bar{c}T}.$$

From assumptions $\mathcal{A}_1 - \mathcal{A}_3$ it follows that

$$\sum_{n=0}^{\infty} U_n(0) = \lim_{m \rightarrow +\infty} S_m(0) < +\infty.$$

In accordance with Markov inequality we find for an arbitrary $\epsilon > 0$ that

$$\sum_{n=0}^{\infty} P \left\{ \sup_{t \in [0, T]} [|x_n(s) - x(s)| + |y_n(s) - y(s)|] > \epsilon \right\} \leq \frac{1}{\epsilon^p} \sum_{n=0}^{\infty} (U_0(0) + \Phi_n) < +\infty. \tag{27}$$

(Φ_n is finite by assumptions $\mathcal{A}_1 - \mathcal{A}_3$). This enables us to obtain straightforwardly the following auxiliary result; that the sequence of the solutions $\{(x_n(t), y_n(t)), t \in [0, T], n \in \mathbb{N}_0\}$ to the Eq. (5) converges in p th moment sense to the solution $\{(x(t), y(t)), t \in [0, T]\}$ of Eq. (1).

The proof follows straightforwardly from (27) since for generic constants Ψ_n , following holds

$$E \left(\sup_{t \in [0, T]} [|x_n(s) - x(s)|^p + |y_n(s) - y(s)|^p] \right) \leq \Psi_n \rightarrow 0, \quad n \rightarrow +\infty.$$

□

Theorem 2.1 and Theorem 2.2 use different techniques for the conclusion of the proof, that is why they are totally separated and both give strong results.

3. Comments and conclusions

- Regarding that it is not known how to estimate $U_{n+1}(t_0)$, i.e. how to solve explicitly the integral inequality (26), which is in fact a difference inequality, the sum $S_m(t_0)$ introduced and with the help of it, a conclusion about the convergence with probability one and in p th moment sense of the approximate solutions is derived.
- As it can be seen above, Theorem 2.1 and Theorem 2.2 remain to be valid with conditions $\mathcal{A}_1 - \mathcal{A}_3$ instead of more strict conditions (6) -(8). However, it should be noted, that in general, it is more difficult to verify $\mathcal{A}_1 - \mathcal{A}_3$ since all the iterations must be known.
- If the approximations (5) have the same final condition as observed BSDEs, ie if $\xi = \xi_{n+1}$, $n \in \mathbb{N}_0$, all proven statements still hold with simpler expressions in conclusions.

- In order to determine an ε -approximation of the solution $\{(x(t), y(t)), t \in [0, T]\}$ to the Eq. (1), our goal is to determine the pair of processes $\{(x_n(t), y_n(t)), t \in [0, T], n \in \mathbb{N}_0\}$ so that following holds:

$$P \left\{ \sup_{t \in [0, T]} [|x_n(s) - x(s)| + |y_n(s) - y(s)|] < \varepsilon \right\} < \delta,$$

for arbitrary small ε, δ and for n large enough. Theoretically, a sequence of approximations can be defined in the following way:

Let a zero approximation be $x_0(t) \equiv \xi, y_0(t) \equiv 0, t \in [0, T]$ where $E|\xi|^p < +\infty$. Further, let $\{\xi_n, n \in \mathbb{N}_0\}$ be a sequence of random variables defined as ξ such that $\sum_{n=0}^{+\infty} E|\xi - \xi_n|^p < +\infty$. Next, we choose functionals f_0, g_0 that are defined as f, g respectively, such that the Lipschitz condition holds with constants $L > 0$, and for some constants a_0, b_0 following holds $\sup_{(x,y,t)} |f(x, y, t) - f_0(x, y, t)|^p < a_0, \sup_{(x,t)} |g(x, t) - g_0(x, t)|^p < b_0$. In next step, we find the solution $\{(x_1(t), y_1(t)), t \in [0, T]\}$ to the equation

$$x_1(t) = \xi_1 - \int_t^T f_0(x_1(s), y_1(s), s) ds - \int_t^T [g_0(x_1(s), s) + y_1(s)] dw(s), \quad t \in [0, T],$$

where $x_1(T) = \xi_1$. If we know $\{(x_n(t), y_n(t)), t \in [0, T], n \in \mathbb{N}_0\}$ we choose functionals f_n, g_n in the same way as f_0, g_0 respectively, and that for some constants a_n, b_n which are n th terms of any convergent series, following conditions are satisfied $\sup_{(x,y,t)} |f(x, y, t) - f_n(x, y, t)|^p < a_n, \sup_{(x,t)} |g(x, t) - g_n(x, t)|^p < b_n$. The $(n + 1)$ -th approximation can be found as a solution of the equation

$$x_{n+1}(t) = \xi_{n+1} - \int_t^T f_n(x_{n+1}(s), y_{n+1}(s), s) ds - \int_t^T [g_n(x_{n+1}(s), s) + y_{n+1}(s)] dw(s), \quad t \in [0, T],$$

where $x_{n+1}(T) = \xi_{n+1}$. It follows that $\sum_{n=0}^{\infty} a_n \leq \sum_{n=0}^{\infty} a_n, \sum_{n=0}^{\infty} b_n \leq \sum_{n=0}^{\infty} b_n$, so assumptions $\mathcal{A}_1 - \mathcal{A}_3$ are satisfied, which provides that Theorem 2.1 holds with this sequence of iterations. Since the $(n + 1)$ -th approximation $\{(x_{n+1}(t), y_{n+1}(t)), t \in [0, T]\}$ of the solution to Eq. (5) is determined by the choice of ξ_n, f_n, g_n the previous iterative method is logically called the Z-algorithm, while the set

$$\{(\xi_n, f_n, g_n), n \in \mathbb{N}_0\},$$

is the determining sequence of the Z-algorithm, analogously to the paper [10] and Zuber’s paper [23].

From the theoretical point of view, the choice of the determining sequence unable us to investigate the solution of Eq. (5) and, in the best case, to solve Eq. (5). Regarding that general non homogeneous BDSDs are usually not effectively solved, the last requirement is extremely strong and it is almost impossible to form such an algorithm for BSDEs. However, it could be convenient to consider some analytic or numerical iterative procedures and conclude which of them could be treated as the Z-algorithm. In such cases, very complex proofs of the convergence of iterations with probability one, as well as in p th moment sense, could be exceeded. For instance, we will prove in the sequel that the Picard method of iterations (5) is a special Z-algorithm.

- It should be noted that results from this paper can be generalized by introducing non lipschitz conditions for the coefficient of equation. The proofs of the theorems would be proven using Bihari’s inequality. Also, for future work some other types of backward equations can be considered, backward doubly stochastic differential equations, backward stochastic Volterra integral equations etc. Z-algorithm for those equations can be also observed under several conditions for the coefficients (Lipschits, nonlipschits).

4. Remarks and examples

It should be noted that the previous procedures can be applied in order to study the following limit problem: together with Eq. (1) we consider the sequence of equations

$$x_{n+1}(t) = \xi_{n+1} - \int_t^T f_n(x_{n+1}(s), y_{n+1}(s), s) ds - \int_t^T [g_n(x_{n+1}(s), s) + y_{n+1}(s)] dw(s), \quad t \in [0, T], \quad n \in \mathbb{N}_0,$$

with final condition $\xi_n = x_n(T)$. After applying the same steps as in previous theorems, after first application of Gronwall-Bellman's inequality, it follows that all theorems hold (in this case, there is no need for introducing process S).

In the sequel we will illustrate few examples of a given Z-algorithm.

Example 1. Let $\{\xi_n, n \in \mathbb{N}_0\}$ be such that (6) holds, and let f_n, g_n which determine $\{(x_{n+1}(t), y_{n+1}(t)), t \in [0, T], n \in \mathbb{N}_0\}$, be defined iteratively in the following way: For $n \in \mathbb{N}_0, (x, y) \in R^d \times R^{d \times m}$, and fixed $t \in [0, T]$

$$\begin{aligned} f_n(x, y, t) &= \lambda_n(x - x_n(t), y - y_n(t)) + f(x_n(t), y_n(t), t), \\ g_n(x, t) &= \bar{\lambda}_n(x - x_n(t)) + g(x_n(t), t), \end{aligned} \tag{28}$$

where functions λ_n is defined in the same way as function f , while function $\bar{\lambda}_n$ is defined as function g , $\lambda_n(0, 0, t) \equiv 0, \bar{\lambda}_n(0, t) \equiv 0$, and introduced functions satisfy Lipschitz condition with some positive constant L . Since

$$f_n(x_n(t), y_n(t), t) - f(x_n(t), y_n(t), t) \equiv 0, \quad g_n(x_n(t), t) - g(x_n(t), t) \equiv 0,$$

assumptions $\mathcal{A}_1 - \mathcal{A}_3$ are satisfied. Therefore Theorem 2.2 yields that the sequence of iterations $\{(x_n(t), y_n(t)), t \in [0, T], n \in \mathbb{N}_0\}$ to the Eq. (5) converges in p-th moment sense to the solution $\{(x(t), y(t)), t \in [0, T]\}$ of Eq. (1). Obviously, this iterative procedure describes the Z-algorithm with the determining sequence

$$\{(\xi_n, f_n, g_n), n \in \mathbb{N}_0\}$$

given by (28). ★

Example 2. In particular, we would linearize the approximate coefficients (28) by taking, for $n \in \mathbb{N}_0$,

$$\begin{aligned} f_n(x, y, t) &= \lambda_n(x - x_n(t)) + \tilde{\lambda}_n(y - y_n(t)) + f(x_n(t), y_n(t), t), \\ g_n(x, t) &= \bar{\lambda}_n(x - x_n(t)) + g(x_n(t), t), \end{aligned} \tag{29}$$

where $\bar{\lambda}_n = (\bar{\lambda}_{1n}, \bar{\lambda}_{2n}, \dots, \bar{\lambda}_{mn})$, and $\lambda_n, \tilde{\lambda}_n, \bar{\lambda}_n$ are scalar sequences. ★

Example 3. This is the most important example. If in (29) we take that $\xi_n = \xi, \lambda_n = \tilde{\lambda}_n = \bar{\lambda}_n = 0, n \in \mathbb{N}_0$, we obtain the Picard iterations (4). All conditions of Theorem 2.2 are satisfied, and it yields that the sequence of iteration $\{(x_n(t), y_n(t)), t \in [0, T], n \in \mathbb{N}_0\}$ converges with probability one to the solution $\{(x(t), y(t)), t \in [0, T]\}$ of Eq. (1). Therefore, the Picard method of iterations is a special Z-algorithm. ★

References

- [1] D. Bainov and P. Simeonov, Integral Inequalities and Applications, Kluwer Academic Publishers, Dordrecht, Netherlands, 1992.
- [2] J. Đorđević, S. Janković, Backward stochastic Volterra integral equations with additive perturbations, Applied Mathematics and Computation, 265 (2015), 903–910.
- [3] J. Đorđević, On a class of backward doubly stochastic differential equations with continuous coefficients, IMA Journal of Applied Mathematics, 81 (2016), 26–41.
- [4] J. Đorđević, L^p -stability of solutions of backward doubly stochastic differential equations, Filomat 31:8(2017) 2356–2379, DOI 10.2298/FIL1708365D.

- [5] J. Đorđević, S. Janković, Reflected backward stochastic differential equations with perturbations, *Discrete and Continuous Dynamical System*, doi:10.3934/dcds.2018075.
- [6] S. Hamadène, J.P. Lepeltier, Reflected BSDEs and mixed game problem, *Stochastic Processes and Their Appl.* 85 (2) (2000) 177–188.
- [7] S. Janković, Iterative procedure for solving stochastic differential equations, *Mathematica Balkanica, New Series* 1 (1) (1987) 64–71.
- [8] S. Janković, Some special iterative procedures for solving stochastic differential equations of Ito type, *Mathematica Balkanica, New Series* (1) (1989) 44–50.
- [9] S. Janković, M. Jovanović, Some analytic iterative methods for solving various classes of stochastic hereditary differential equations, *Facta Universitatis, Series Mechanics, Automatic Control and Robotics* 4 (16) (2004) 11–31.
- [10] S. Janković, M. Vasilova, M. Krstić, Some analytic approximations for neutral stochastic functional differential equations, *Applied Mathematics and Computation*, 217 (2010) 3615–3623.
- [11] S. Janković, J. Đorđević, M. Jovanović, On a class of backward doubly stochastic differential equations, *Applied Mathematics and Computation*, 217 (2011), 8754–8764, Corrigendum to On a class of backward doubly stochastic differential equations, *Appl. Math. Comput.* 218 (2012) 9033–9034.
- [12] S. Janković, M. Jovanović, J. Djordjevic, Perturbed backward stochastic differential equations, 55 (2012), 1734–1745.
- [13] I. Karatzas and S. E. Shreve, *Brownian Motion and Stochastic Calculus*, Springer-Verlag, Berlin, 1991.
- [14] N. El-Karoui, E. Pardoux, M.C. Quenez, Reflected backwards SDEs and American options, *Numerical Methods in Finance, Publ. Newton Inst.*, Cambridge University Press, Cambridge (1997) 215–231.
- [15] N. El-Karoui, S. Peng, M.C. Quenez, Backward stochastic differential equations in finance, *Math. Finance* 7 (1) (1997) 1–71.
- [16] N. El-Karoui, L. Mazliak (Eds.), *Backward Stochastic Differential Equations*, Pitman Res. Notes Math., vol. 364, Longman, Harlow, 1997.
- [17] N. El-Karoui, S. Hamadène, BSDEs and risk-sensitive control, zero-sum and nonzero sum game problems of stochastic functional differential equations, *Stochastic Processes and Their Appl.* 107 (1) (2003) 145–169.
- [18] X. Mao, Adapted solutions of backward stochastic differential equations with non-Lipschitz coefficients, *Stochastic Processes and Their Applications* 58 (1995) 281–292.
- [19] X. Mao, *Stochastic Differential Equations and Applications*, Horwood, Chichester, UK (Second Edition), 2008.
- [20] E. Pardoux, S.G. Peng, Adapted solution of a backward stochastic differential equation, *Systems & Control Letters* 14 (1) (1990) 55–61.
- [21] E. Pardoux, BSDEs, weak convergence and homogenization of semilinear parabolic PDEs, in: *Nonlinear Analysis, Differential Equations and Control*, (Montreal, QC, 1998) (F.H. Clarke, R.J. Stern, and G.Sabidussi, eds.), NATO Sci.Ser.C. Math.Phys.Sci, vol. 528, Kluwer Acad. Publ., Dordrecht (1999) 503–549.
- [22] Jianfeng Zhang, A numerics scheme for BSDEs, *Ann. Appl. Probab.* 14 (1) (2004) 459–488.
- [23] R. Zuber, About one algorithm for solving first order differential equations (I), *Zastosow. Math.* 8 (4) (1966) 351–363 (in Polish).
- [24] R. Zuber, About one algorithm for solving first order differential equations (II), *Zastosow. Math.* 11 (1) (1966) 85–97 (in Polish).
- [25] E. Pardoux, S.G. Peng, Backward stochastic differential equations and quasilinear parabolic partial differential equations, in: *Stochastic Partial Differential Equations and Their Applications*, (Charlotte, NC, 1991) (B. Rozowski and R. Sowers, eds.), *Lecture Notes in Control and Information Systems*, Vol. 176, Springer, Berlin (1992) 200–217.
- [26] M.I. Friedlin, A.D. Wentzell, *Random Perturbations of Dynamical Systems*, Springer, Berlin, 1984.
- [27] S. Janković, M. Jovanović, Perturbed stochastic hereditary differential equations with integral contractors, *Computers & Math. with Appl.* 42 (2001) 871–881.
- [28] S. Janković, M. Jovanović, Generalized stochastic perturbation-dependent differential equations, *Stochastic Analysis and Appl.* 20 (6) (2002) 1281–1307.
- [29] S. Janković, M. Jovanović, *Analytic approximations of solutions to stochastic differential equations* (monograph), Faculty of Science and Mathematics, University of Niš, 2008.
- [30] M. Jovanović, S. Janković, On perturbed nonlinear Itô type stochastic integrodifferential equations, *J. Math. Analysis and Appl.* 269 (2002) 301–316.
- [31] M. Jovanović, S. Janković, Functionally perturbed stochastic differential equations, *Mathematische Nachrichten* 279 (16) (2006) 1808–1822.
- [32] M. Jovanović, S. Janković, Neutral stochastic functional differential equations with additive perturbations, *Applied Math. and Computations*, doi:10.1016/j.amc.2009.03.031
- [33] R. Khasminskii, On stochastic processes defined by differential equations with a small parameter. *Theory Prob. Appl.* 11 (1966) 211–228.
- [34] B.J. Ma, P. Protter, J.S. Martin, S. Torres, Numerical method for backward stochastic differential equations, *The Annals of Applied Prob.* 12 (1) (2002) 302–316.
- [35] J. Stoyanov, D. Botev, Quantitative results for perturbed stochastic differential equations, *J. Appl. Math. and Stoch. Analysis* 9 (3) (1996) 255–261.
- [36] J. Stoyanov, Regularly perturbed stochastic differential systems with an internal random noise. In: *Proc. 2nd World Congress Nonlin. Anal.*, Athens, July 1996.