J. Math. Anal. Appl. 363 (2010) 128–137

Contents lists available at ScienceDirect

Journal of Mathematical Analysis and Applications

www.elsevier.com/locate/jmaa

An approximate method via Taylor series for stochastic functional differential equations

Marija Milošević, Miljana Jovanović [1], Svetlana Janković *

University of Niš, Faculty of Science and Mathematics, Višegradska 33, 18000 Niš, Serbia

ARTICLE INFO

Article history:
Received 17 February 2009
Available online 8 August 2009
Submitted by M. Ledoux

Keywords:
Stochastic functional differential equation
Fréchet derivative
Taylor approximation
L^p and almost sure convergence

ABSTRACT

The subject of this paper is an analytic approximate method for stochastic functional differential equations whose coefficients are functionals, sufficiently smooth in the sense of Fréchet derivatives. The approximate equations are defined on equidistant partitions of the time interval, and their coefficients are general Taylor expansions of the coefficients of the initial equation. It will be shown that the approximate solutions converge in the L^p-norm and with probability one to the solution of the initial equation, and also that the rate of convergence increases when degrees in Taylor expansions increase, analogously to real analysis.

1. Introduction and preliminary results

In many circumstances, some of the most frequent and most important stochastic models, when the future states of systems depend not only on present states, but also on their past history, are described by complex stochastic functional differential equations. The main interest in the field has often been directed to the existence, uniqueness and stability, as well as to the study of qualitative and quantitative properties of the solutions. We refer the reader to Mohamed [15] and to more papers and books by X. Mao [12–14], and the literature cited therein, among others. Moreover, because it is almost impossible to solve these equations explicitly, it is important to find some analytic and numerical approximations of the solutions. However, there is only a small amount of papers referring to such problems, Buckwar [3], Mao [14], for example. The present paper refers to an analytic method, which could lead to some constructions of appropriate numerical methods.

Before stating the main results to be proved, we will briefly reproduce only the essential notations and definitions, which are necessary in our investigation. Our initial assumption is that all random variables and processes are defined on a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geqslant 0}, P)$ with a filtration $\{\mathcal{F}_t\}_{t\geqslant 0}$ satisfying the usual conditions (that is, it is increasing and right-continuous, and $\mathcal{F}_0$ contains all P-null sets). Let $w(t) = (w_1(t), w_2(t), \dots, w_m(t))^T$, $t \geqslant 0$, be an m-dimensional standard Brownian motion, $\mathcal{F}_t$-adapted and independent of $\mathcal{F}_0$. Let the Euclidean norm be denoted by $|\cdot|$ and, for simplicity, $trace[B^T B] = |B|^2$ for a matrix B, where B^T is the transpose of a vector or a matrix.

For a given $\tau > 0$, let $C([-\tau, 0]; R^d)$ be the family of continuous functions φ from $[-\tau, 0]$ to R^d, equipped with the supremum norm $\|\varphi\| = \sup_{-\tau\leqslant\theta\leqslant 0} |\varphi(\theta)|$. Obviously, $(C([-\tau, 0]; R^d), \|\cdot\|)$ is a Banach space.

Taking into account the Gaussian white noise, the evolution of such a system can be described with a stochastic functional differential equation of the form

$$dx(t) = f(x_t, t)\,dt + g(x_t, t)\,dw(t), \quad t \in [t_0, T], \qquad x_{t_0} = \xi, \tag{1}$$

* Corresponding author.
E-mail addresses: marija.misa@sezampro.co.yu (M. Milošević), mima@pmf.ni.ac.yu (M. Jovanović), svjank@pmf.ni.ac.yu (S. Janković).

[1] Supported by Grant No. 44003 of MNTRS.

0022-247X/$ – see front matter
doi:10.1016/j.jmaa.2009.07.061

where the functionals

$$f : C\big([-\tau,0]; R^d\big) \times [t_0,T] \to R^d, \qquad g : C\big([-\tau,0]; R^d\big) \times [t_0,T] \to R^d \times R^m$$

are Borel measurable, $x(t)$ is a d-dimensional state process and $x_t = \{x(t+\theta),\ \theta \in [-\tau,0]\}$ is a $C([-\tau,0]; R^d)$-valued stochastic process which is regarded as the past history of the state. Because of the past dependence, the initial condition is defined on the entire interval $[t_0-\tau, t_0]$, that is

$$x_{t_0} = \xi = \big\{\xi(\theta),\ \theta \in [-\tau,0]\big\}, \tag{2}$$

where ξ is an $\mathcal{F}_{t_0}$-measurable and $C([-\tau,0]; R^d)$-valued random variable such that $E\|\xi\|^2 < \infty$.

A d-dimensional stochastic process $\{x(t),\ t \in [t_0-\tau,T]\}$ is said to be a solution to Eq. (1) if it is a.s. continuous and $\{x_t,\ t \in [t_0,T]\}$ is $\mathcal{F}_t$-adapted, $\int_{t_0}^T |f(x_t,t)|\,dt < \infty$ a.s., $\int_{t_0}^T |g(x_t,t)|^2\,dt < \infty$ a.s., $x_{t_0} = \xi$ a.s. and for every $t \in [t_0,T]$, the integral form of Eq. (1) holds a.s.

A solution $\{x(t),\ t \in [t_0-\tau,T]\}$ is said to be unique if any other solution $\{\tilde{x}(t),\ t \in [t_0-\tau,T]\}$ is indistinguishable from it, in the sense that $P\{x(t) = \tilde{x}(t),\ t \in [t_0-\tau,T]\} = 1$.

The basic existence-and-uniqueness theorem [12, Theorem 5.2.5, p. 153] guarantees that if f and g satisfy the uniform Lipschitz condition and the linear growth condition, that is, if there exists a constant $K > 0$ such that

$$\big|f(\varphi,t) - f(\psi,t)\big| \vee \big|g(\varphi,t) - g(\psi,t)\big| \leqslant K\|\varphi - \psi\|, \tag{3}$$

$$\big|f(\varphi,t)\big| \vee \big|g(\varphi,t)\big| \leqslant K\big(1 + \|\varphi\|\big) \tag{4}$$

for all $t \in [t_0,T]$ and $(\varphi,\psi) \in C([-\tau,0]; R^d)$, then there exists a unique a.s. continuous solution $x(t)$ to Eq. (1). Moreover, if $E\|\xi\|^p < \infty$ for any $p \geqslant 2$, then $E\sup_{t_0-\tau \leqslant t \leqslant T} |x(t)|^p < \infty$ [12, Theorem 5.4.1, p. 158].

Essentially, the fundamentals of the approximate method considered here go back to papers [1,2] by M.A. Atalla and [6,7] by S. Janković and D. Ilić. In [1] the solution $x = \{x(t),\ t \in [0,1]\}$ of an ordinary stochastic differential equation $dx(t) = a(x(t),t)\,dt + b(x(t),t)\,dw(t)$, $t \in [0,1]$, $x(0) = x_0$, is approximated by the processes $x_n = \{x_n(t),\ t \in [0,1]\}$, $n \in N$, by successively connecting the solutions $\{x_n(t),\ t \in [t_k,t_{k+1}]\}$, $k = 0,\dots,n-1$ of the equations $dx_n(t) = a(x_n(t_k),t)\,dt + b(x_n(t_k),t)\,dw(t)$, $x_n(0) = x_0$, $t \in [t_k,t_{k+1}]$ at division points t_k of an arbitrary partition $0 = t_0 < t_1 < \cdots < t_n = 1$ of the time interval. The rate of this approximation in the L^p-norm, $p \geqslant 2$, is found to be $O(\delta_n^{p/2})$ when $n \to \infty$ and $\delta_n = \max_{0 \leqslant k \leqslant n-1}(t_{k+1} - t_k) \to 0$.

We mention here the key paper [2] by Atalla in which he improved his own result from [1] by using a sequence of linear stochastic differential equations $dx_n(t) = [a(x_n(t_k),t) + a'_x(x_n(t_k),t)(x_n(t) - x_n(t_k))]\,dt + [b(x_n(t_k),t) + b'_x(x_n(t_k),t)) \times (x_n(t) - x_n(t_k))]\,dw(t)$, $x_n(0) = x_0$, $t \in [t_k,t_{k+1})$, $k = 0,\dots,n-1$, that is, the equations in which the drift and diffusion coefficients are Taylor approximations of $a(x,t)$ and $b(x,t)$ up to the first derivative in x. The rate of this approximation, in the L^p-norm, was $O(\delta_n^p)$ when $n \to \infty$ and $\delta_n \to 0$.

Having in mind that Taylor approximations, as polynomials, could be a useful tool to approximate analytically or numerically the coefficients of stochastic differential equations, Atalla's concept in [2] is appropriately extended in [6] in the sense that the drift and diffusion coefficients of approximate equations are taken to be Taylor polynomials in x of degrees m_1 and m_2 for $a(x,t)$ and $b(x,t)$, respectively. The rate of this approximation, in the L^p-norm, has been found to be $O(\delta_n^{(m+1)p/2})$ when $n \to \infty$ and $\delta_n \to 0$, where $m = \min\{m_1,m_2\}$. This result was extended in [7] to stochastic integrodifferential equations.

By following the concepts from papers [6,7], we want to approximate the solution to Eq. (1) with a sequence of solutions to stochastic functional differential equations the drift and diffusion coefficients of which are Taylor expansions of f and g, respectively, up to arbitrary derivatives in x. Having in mind that f and g are functionals, we will approximate them by using Fréchet derivatives and general Taylor formula. For this reason, in the remainder of this section we first give a brief survey about these notions referring, in general, to an arbitrary mapping (see [4,5], for example). In the next section we formulate the problem and present our main results. Note that the proofs of the assertions in the next section are completely different with respect to the ones from [6,7], since the present and past states of the solution appear in Eq. (1), and also because of the need to apply Fréchet derivatives and general Taylor formula.

Let V and W be vector spaces on the same scalar field. A mapping $T : V^k \to W$ is said to be multi-linear if it is linear in each argument, that is, if $T(v_1,\dots,v_{i-1},a_iv_i + u_i,v_{i+1},\dots,v_k) = a_iT(v_1,\dots,v_{i-1},v_i,v_{i+1},\dots,v_k) + T(v_1,\dots,v_{i-1},u_i,v_{i+1},\dots,v_k)$ for each $1 \leqslant i \leqslant k$ and for all scalars a_i and $v_i,u_i \in V$. Denote by $\mathcal{T}(V^k \to W)$ a set of all multi-linear mappings $T : V^k \to W$, which itself generates a vector space in a usual way.

Definition 1. Let V and W be Banach spaces and D an open subset of V. A mapping $T : D \to W$ is a Fréchet-differentiable in $x \in D$ if there exists a bounded linear operator $A \in \mathcal{L}(V \to W)$ satisfying

$$\lim_{\|h\|\to 0} \frac{\|T(x+h) - T(x) - A(h)\|}{\|h\|} = 0.$$

The operator A is called the Fréchet derivative of the mapping T at x and it will be denoted by $T'_{(x)}$. If the mapping $T: D \to W$ is Fréchet differentiable on D, then higher order Fréchet derivatives in x, denoted by $T^{(2)}_{(x)}, T^{(3)}_{(x)}, \dots, T^{(k)}_{(x)}$, are multi-linear mappings from V to W. These facts suggest a natural identification of the spaces $\mathcal{L}(V \to \mathcal{T}(V^k \to W))$ and $\mathcal{T}(V^{k+1} \to W)$, that is, $\mathcal{L}(V \to \mathcal{T}(V^k \to W) = \mathcal{T}(V^{k+1} \to W)$. The spaces $\mathcal{L}(V \to \mathcal{T}(V^k \to W))$ and $\mathcal{T}(V^{k+1} \to W)$ are isometric.

Let D be a convex subset of a Banach space V and let T be an arbitrary $n+1$ times Fréchet-differentiable mapping on D. If x and $x+h$ belong to D, then

$$T(x+h) = \sum_{k=0}^{n} \frac{1}{k!} T^{(k)}_{(x)}(\underbrace{h, \dots, h}_{k}) + W(x,h) \tag{5}$$

is a Taylor approximation of the mapping T in the neighborhood of x, where the Lagrange form of the residuum $W(x,h)$ is

$$W(x,h) = \frac{1}{(n+1)!} T^{(n+1)}_{(x+th)}(\underbrace{h, \dots, h}_{n+1}) \tag{6}$$

for some $t \in (0,1)$ and

$$\left\| W(x,h) \right\| \leqslant \frac{1}{(n+1)!} \sup_{t\in[0,1]} \left\| T^{(n+1)}_{(x+th)} \right\| \cdot \|h\|^{n+1}. \tag{7}$$

2. Main results

Let us first present Eq. (1) in its equivalent integral form,

$$x(t) = \xi(0) + \int_{t_0}^{t} f(x_s, s)\, ds + \int_{t_0}^{t} g(x_s, s)\, dw(s), \quad t \in [t_0, T] \tag{8}$$

with the initial condition (2). Let also

$$t_0 < t_1 < \cdots < t_n = T \tag{9}$$

be an equidistant partition of the interval $[t_0, T]$, that is, the partition points are

$$t_k = t_0 + \frac{k}{n}(T - t_0), \quad k = 0, 1, \dots, n,$$

with $\delta_n = (T - t_0)/n \in (0,1)$ for large enough integers $n \in N$.

As it has already been mentioned, the solution $x = \{x(t),\ t \in [t_0 - \tau, T]\}$ to Eq. (8) will be approximated on the partition (9) by the solutions $\{x^n(t),\ t \in [t_k, t_{k+1}]\}$, $k = 0, 1, \dots, n-1$ of the equations

$$\begin{aligned} x^n(t) = x^n(t_k) &+ \int_{t_k}^{t} \sum_{i=0}^{m_1} \frac{f^{(i)}_{(x^n_{t_k}, s)}(x^n_s - x^n_{t_k}, \dots, x^n_s - x^n_{t_k})}{i!}\, ds \\ &+ \int_{t_k}^{t} \sum_{i=0}^{m_2} \frac{g^{(i)}_{(x^n_{t_k}, s)}(x^n_s - x^n_{t_k}, \dots, x^n_s - x^n_{t_k})}{i!}\, dw(s), \quad t \in [t_k, t_{k+1}], \end{aligned} \tag{10}$$

satisfying the initial conditions $x_{t_0} = \xi$, $x^n_{t_k} = \{x^n(t_k + \theta),\ \theta \in [-\tau, 0]\}$, $k = 1, 2, \dots, n-1$, the drift and diffusion coefficients of which are Taylor expansions of f and g in the first argument in the neighborhood of the points $x^n_{t_k}$, up to the m_1-th and m_2-th Fréchet derivatives, respectively.

By following the concept presented in Section 1, the approximate solution $x^n = \{x^n(t),\ t \in [t_0 - \tau, T]\}$ will be obtained as an a.s. continuous process, by connecting successively the initial condition $\xi = \{\xi(\theta),\ \theta \in [-\tau, 0]\}$ and the processes $\{x^n(t),\ t \in [t_k, t_{k+1}]\}$ at the division points t_k whenever $k = 0, 1, \dots, n-1$.

Obviously, it must be required that f and g satisfy appropriate conditions, first of all, they must be sufficiently smooth. With no particular emphasis on conditions, we suppose the existence and uniqueness of the solutions to Eqs. (8) and (10), and we emphasize only the conditions explicitly used in our discussion. In spite of the assumptions that f and g satisfy the Lipschitz condition (3) and growth condition (4), we introduce the following assumptions:

$\mathcal{A}_1$: The functionals f and g have Taylor expansions in the argument x up to the m_1-th and m_2-th Fréchet derivatives, respectively.

$\mathcal{A}_2$: The functionals $f_{(x,t)}^{(m_1+1)}$ and $g_{(x,t)}^{(m_2+1)}$ are uniformly bounded, i.e., there exist positive constants $L_1, L_2 > 0$ such that

$$\sup_{C([-\tau,0];R^d)\times[t_0,T]} \left\| f_{(x,t)}^{(m_1+1)}(h,\dots,h)\right\| \leqslant L_1 \cdot \|h\|^{m_1+1},$$
$$\sup_{C([-\tau,0];R^d)\times[t_0,T]} \left\| g_{(x,t)}^{(m_2+1)}(h,\dots,h)\right\| \leqslant L_2 \cdot \|h\|^{m_2+1}.$$

$\mathcal{A}_3$: There exist unique, a.s. continuous solutions x and x^n to Eqs. (8) and (10), respectively, such that, for $p \geqslant 2$,

$$E \sup_{t\in[t_0-\tau,T]} \left|x(t)\right|^p < \infty, \qquad E \sup_{t\in[t_0-\tau,T]} \left|x^n(t)\right|^{(M+1)^2 p} \leqslant Q < \infty,$$

where $M = \max\{m_1, m_2\}$ and $Q > 0$ is a constant independent of n. Moreover, we suppose that $E\|\xi\|^{(M+1)^2 p} < \infty$ and that all the Lebesgue and Ito integrals employed further are also well defined.

$\mathcal{A}_4$: The initial condition (2) is uniformly Lipschitz continuous, i.e., there exists a constant $\beta > 0$ such that, for all $-\tau \leqslant \theta_1, \theta_2 \leqslant 0$,

$$\left|\xi(\theta_2) - \xi(\theta_1)\right| \leqslant \beta|\theta_2 - \theta_1|.$$

Furthermore, we will apply several times, without special emphasis, the elementary inequality $(\sum_{i=1}^m a_i)^q \leqslant m^{q-1}\sum_{i=1}^m a_i^q$, $a_i > 0$, $q \in N$, the usual Ito isometry, the Hölder inequality to Lebesgue integrals and the well-known Burkholder–Davis–Gundy inequality to Ito integrals [8,12].

In order to estimate the closeness between the solutions x and x^n, we first state some auxiliary results, which can be treated independently of the mentioned problem, but which are essentially used in the proof of the main result.

Proposition 1. *Let* $\{x^n(t),\ t \in [t_k - \tau, t_{k+1}]\}$, $k = 0, 1, \dots, n-1$, *be the solution to Eq.* (10) *and let the condition* (4) *and the assumptions* $\mathcal{A}_1$, $\mathcal{A}_2$, $\mathcal{A}_3$ *be satisfied. Then, for every* $2 \leqslant r \leqslant (M+1)p$,

$$E \sup_{s\in[t_k,t]} \left|x^n(s) - x^n(t_k)\right|^r \leqslant C \cdot n^{-r/2}, \quad t \in [t_k, t_{k+1}],\ k = 0, 1, \dots, n-1,$$

where C is a generic constant independent of n.

Proof. For reasons of notational simplicity, denote that

$$F\left(x_t^n, t; x_{t_k}^n\right) = \sum_{i=0}^{m_1} \frac{f_{(x_{t_k}^n,t)}^{(i)}(x_t^n - x_{t_k}^n, \dots, x_t^n - x_{t_k}^n)}{i!},$$
$$G\left(x_t^n, t, x_{t_k}^n\right) = \sum_{i=0}^{m_2} \frac{g_{(x_{t_k}^n,t)}^{(i)}(x_t^n - x_{t_k}^n, \dots, x_t^n - x_{t_k}^n)}{i!}.$$

Then, in view of $\mathcal{A}_1$, for $t \in [t_k, t_{k+1}]$, $k = 0, 1, \dots, n-1$ and $\bar\theta \in (0,1)$,

$$f\left(x_t^n, t\right) = F\left(x_t^n, t; x_{t_k}^n\right) + \frac{f_{(x_{t_k}^n+\bar\theta(x_t^n - x_{t_k}^n),t)}^{(m_1+1)}(x_t^n - x_{t_k}^n, \dots, x_t^n - x_{t_k}^n)}{(m_1+1)!},$$
$$g\left(x_t^n, t\right) = G\left(x_t^n, t; x_{t_k}^n\right) + \frac{g_{(x_{t_k}^n+\bar\theta(x_t^n - x_{t_k}^n),t)}^{(m_2+1)}(x_t^n - x_{t_k}^n, \dots, x_t^n - x_{t_k}^n)}{(m_2+1)!}. \tag{11}$$

In order to estimate $E\sup_{s\in[t_k,t]}|x^n(s) - x^n(t_k)|^r$, we apply the previously cited elementary inequality to Eq. (10), the Hölder inequality to the Lebesgue integral and the Burkholder–Davis–Gundy inequality to the Ito integral for $r > 2$, that is, Doob inequality for $r = 2$. So, we find, for all $t \in [t_k, t_{k+1}]$, $k = 0, 1, \dots, n-1$, that

$$\begin{aligned} E \sup_{s\in[t_k,t]} \left|x^n(s) - x^n(t_k)\right|^r &\leqslant 2^{r-1}(t-t_k)^{r-1}\int_{t_k}^t E\left|F\left(x_s^n, x_{t_k}^n, s\right)\right|^r ds + 2^{r-1}c_r(t-t_k)^{\frac{r}{2}-1}\int_{t_k}^t E\left|G\left(x_s^n, x_{t_k}^n, s\right)\right|^r ds \\ &\equiv 2^{r-1}(t-t_k)^{\frac{r}{2}-1}\left[(t-t_k)^{\frac{r}{2}} J_1(t) + c_r J_2(t)\right], \end{aligned} \tag{12}$$

where c_r is a universal constant from the Burkholder–Davis–Gundy inequality, while $J_1(t)$ and $J_2(t)$ are the appropriate integrals. On the basis of Taylor expansion (11), the growth condition (4) and the assumptions $\mathcal{A}_1$, $\mathcal{A}_2$ and $\mathcal{A}_3$, we conclude that $J_1(t)$ can be estimated in the following way,

$$
\begin{aligned}
J_1(t) &= \int_{t_k}^{t} E\big|f\big(x_s^n, s\big) - \big[f\big(x_s^n, s\big) - F\big(x_s^n, x_{t_k}^n, s\big)\big]\big|^r\,ds \\
&\equiv \int_{t_k}^{t} E\left|f\big(x_s^n, s\big) - \frac{f^{(m_1+1)}_{(x_{t_k}^n+\bar\theta(x_s^n - x_{t_k}^n),s)}(x_s^n - x_{t_k}^n, \dots, x_s^n - x_{t_k}^n)}{(m_1+1)!}\right|^r ds \\
&\leqslant 2^{r-1}\left[K^r\int_{t_k}^{t} E\big|f\big(x_s^n, s\big)\big|^r\,ds + \frac{L_1^r}{[(m_1+1)!]^r}\int_{t_k}^{t} E\big\|x_s^n - x_{t_k}^n\big\|^{(m_1+1)r}\,ds\right] \\
&\leqslant 2^{r-1}\left[K^r 2^{r-1}\int_{t_k}^{t}\big(1 + E\big\|x_s^n\big\|^r\big)\,ds + \frac{L_1^r}{[(m_1+1)!]^r}2^{(m_1+1)r-1}\int_{t_k}^{t}\big(E\big\|x_s^n\big\|^{(m_1+1)r} + E\big\|x_{t_k}^n\big\|^{(m_1+1)r}\big)\,ds\right] \\
&\leqslant 2^{2r-1}\left[K^r(2+Q)(t-t_k) + 2^{(m_1+1)r}\frac{(1+Q)L_1^r}{[(m_1+1)!]^r}(t-t_k)\right] \\
&\equiv C_1(t-t_k),
\end{aligned}
$$

where $\bar\theta \in (0,1)$ and $C_1 \equiv C_1(K, L_1, Q, r, m_1)$ is a generic constant.

Similarly, by repeating the previous procedure, we see that

$$J_2(t) \leqslant C_2(t-t_k),$$

where $C_2 \equiv C_2(K, L_2, Q, r, m_2)$ is a generic constant. Therefore, (12) yields

$$
\begin{aligned}
E\sup_{s\in[t_k,t]}\big|x^n(s) - x^n(t_k)\big|^r &\leqslant 2^{r-1}(t-t_k)^{\frac{r}{2}}\big[C_1(T-t_k)^{\frac{r}{2}} + c_r C_2\big] \\
&\leqslant C(t-t_k)^{\frac{r}{2}} \\
&\leqslant C\cdot n^{-r/2}, \quad t\in[t_k, t_{k+1}],
\end{aligned}
$$

where C is a generic constant independent of n. □

Let us recall that, for the proof of the next assertion, it is of great importance that the partition of the interval $[t_0, T]$ be equidistant. Then, if we shift for τ two arbitrary points $t_1, t_2 \in [t_k, t_{k+1}]$, the points $t_1 - \tau$ and $t_2 - \tau$ will also fall into the same interval $[t_j, t_{j+1}]$, $j = 0, 1, \dots, k$, or into $[t_0 - \tau, t_0]$.

Proposition 2. *Let the conditions of Proposition* 1 *and the assumption* $\mathcal{A}_4$ *be satisfied. Then, for all* $2 \leqslant r \leqslant (M+1)p$,

$$E\big\|x_t^n - x_{t_k}^n\big\|^r \leqslant B\cdot n^{-r/2}, \quad t\in[t_k, t_{k+1}],\ k = 0, 1, \dots, n-1,$$

where B is a generic constant independent of n.

Proof. In order to prove this statement, we must discuss three cases.

Case 1. Assume $t - \tau < t_0$. Then both processes $x_t^n = \{x^n(t+\theta),\ \theta\in[-\tau,0]\}$ and $x_{t_k}^n = \{x^n(t_k+\theta),\ \theta\in[-\tau,0]\}$ coincide with the initial condition for some $\theta\in[-\tau,0]$. In accordance with the norm in $C([-\tau,0]; R^d)$, one gets

$$
\begin{aligned}
E\big\|x_t^n - x_{t_k}^n\big\|^r &= E\sup_{\theta\in[-\tau,0]}\big|x^n(t+\theta) - x^n(t_k+\theta)\big|^r \\
&= E\sup_{u\in[t_k-\tau,t_k]}\big|x^n(u+t-t_k) - x^n(u)\big|^r \\
&\leqslant E\sup_{u\in[t_k-\tau,t_0+t_k-t]}\big|x^n(u+t-t_k) - x^n(u)\big|^r \\
&\quad + E\sup_{u\in[t_0+t_k-t,t_0]}\big|x^n(u+t-t_k) - x^n(u)\big|^r \\
&\quad + E\sup_{u\in[t_0,t_k]}\big|x^n(u+t-t_k) - x^n(u)\big|^r.
\end{aligned}
$$

The first term is determined for $u + t - t_k \leqslant t_0$ and $u \leqslant t_0$, the second one for $u + t - t_k \geqslant t_0$ and $u \leqslant t_0$, and the third one for $u + t - t_k \geqslant t_0$ and $u \geqslant t_0$. Thus, it follows that

$$\begin{aligned}E\|x_t^n - x_{t_k}^n\|^r &\leqslant E \sup_{u\in[t_k-\tau, t_0+t_k-t]} \left|\xi(u+t-t_k-t_0) - \xi(u-t_0)\right|^r \\ &\quad + E \sup_{u\in[t_0+t_k-t, t_0]} \left|x^n(u+t-t_k) - \xi(u-t_0)\right|^r \\ &\quad + E \sup_{u\in[t_0,t_k]} \left|x^n(u+t-t_k) - x^n(u)\right|^r \\ &\equiv (A_1 + A_2 + A_3). \end{aligned} \tag{13}$$

The estimate A_1 follows from the assumption $\mathcal{A}_4$, i.e., from the Lipschitz condition for the initial data,

$$A_1 \leqslant \beta^r (T-t_k)^{\frac{r}{2}} n^{-r/2}. \tag{14}$$

To estimate A_2, we use the fact that $u+t-t_k \in [t_0, t_1]$ whenever $u \leqslant t_0$, which is implied by the uniformness of the partition, and then Proposition 1 and the assumption $\mathcal{A}_4$. Finally,

$$\begin{aligned}A_2 &\leqslant 2^{r-1}\Big[E \sup_{u\in[t_0+t_k-t, t_0]} \left|x^n(u+t-t_k) - x^n(t_0)\right|^r + E \sup_{u\in[t_0+t_k-t,t_0]} \left|\xi(0) - \xi(u-t_0)\right|^r\Big] \\ &\leqslant 2^{r-1}\big[C + \beta^r (T-t_k)^{\frac{r}{2}}\big] n^{-r/2}. \end{aligned} \tag{15}$$

In order to estimate A_3, we can notice that there exist $i \in \{0, \dots, k\}$ and $v \in [t_i, t_{i+1}]$ so that $A_3 = E|x^n(v+t-t_k) - x^n(v)|^r$. Then, we differentiate two cases.

First, let $t_i \leqslant v \leqslant v+t-t_k \leqslant t_{i+1}$. In view of Proposition 1 we have

$$A_3 \leqslant 2^{r-1}\big[E\left|x^n(v+t-t_k) - x^n(t_i)\right|^r + E\left|x^n(v) - x^n(t_i)\right|^r\big] \leqslant 2^r C n^{-r/2}. \tag{16}$$

Second, let $t_i \leqslant v \leqslant t_{i+1} \leqslant v+t-t_k \leqslant t_{i+2}$. Then, Proposition 1 again yields

$$A_3 \leqslant 3^{r-1}\big[E\left|x^n(v+t-t_k) - x^n(t_{i+1})\right|^r + E\left|x^n(t_{i+1}) - x^n(t_i)\right|^r + E\left|x^n(v) - x^n(t_i)\right|^r\big] \leqslant 3^r C n^{-r/2}. \tag{17}$$

Now, (14), (15), (16) and (17) together with (13) gives

$$E\|x_t^n - x_{t_k}^n\|^r \leqslant B n^{-r/2},$$

where B is a generic constant independent of n.

Case 2. Let $t_k - \tau < t_0 \leqslant t - \tau$. Now, only the process $x_{t_k}^n = \{x^n(t_k+\theta),\ \theta \in [-\tau, 0]\}$ coincides with the initial condition for some $\theta \in [-\tau, 0]$. Similarly to Case 1, one gets

$$\begin{aligned}E\|x_t^n - x_{t_k}^n\|^r &\leqslant E \sup_{u\in[t_k-\tau,t_0]} \left|x^n(u+t-t_k) - x^n(u)\right|^r + E \sup_{u\in[t_0,t_k]} \left|x^n(u+t-t_k) - x^n(u)\right|^r \\ &\leqslant (A_2 + A_3) \\ &\leqslant B n^{-r/2}. \end{aligned}$$

Case 3. Let $t_k - \tau \geqslant t_0$. Then, in view of Case 1,

$$E\|x_t^n - x_{t_k}^n\|^r \leqslant E \sup_{u\in[t_0,t_k]} \left|x^n(u+t-t_k) - x^n(u)\right|^r = A_3 \leqslant B n^{-r/2}.$$

Thus, the proof becomes complete. □

Finally, by applying Proposition 2, one can prove the main result in this paper, that the sequence of approximate solutions $\{x^n,\ n \in N\}$ converges in the p-th moment sense to the solution x of Eq. (8). This conclusion follows from the next theorem, in which the rate of the closeness between x and x^n is given.

Theorem 1. *Let x be a solution to Eq.* (8) *and x^n be its approximate solution determined with Eq.* (10)*. Let also the conditions of Proposition* 2 *and the Lipschitz condition* (3) *be satisfied. Then, for $p \geqslant 2$,*

$$E \sup_{t\in[t_0-\tau,T]} \left|x(t) - x^n(t)\right|^p \leqslant H \cdot n^{-(m+1)p/2},$$

where $m = \min\{m_1, m_2\}$ and H is a generic constant independent of n.

Proof. For an arbitrary $t \in [t_0, T]$, by substituting Eqs. (8) and (10), it follows that

$$x(t) - x^n(t) = \int_{t_0}^{t} \sum_{k:\, t_k \leqslant t} \left[f(x_s, s) - F\left(x_s^n, s; x_{t_k}^n\right) \right] I_{[t_k, t_{k+1})}(s)\, ds + \int_{t_0}^{t} \sum_{k:\, t_k \leqslant t} \left[g(x_s, s) - G\left(x_s^n, s; x_{t_k}^n\right) \right] I_{[t_k, t_{k+1})}(s)\, dw(s). \tag{18}$$

Since x and x^n satisfy the same initial condition, one obtains

$$\begin{aligned}
E \sup_{s\in[t_0-\tau,t]} \left|x(s) - x^n(s)\right|^p &\leqslant E \sup_{s\in[t_0-\tau,t_0]} \left|x(s) - x^n(s)\right|^p + E \sup_{s\in[t_0,t]} \left|x(s) - x^n(s)\right|^p \\
&= E \sup_{s\in[t_0,t]} \left|x(s) - x^n(s)\right|^p \\
&\leqslant 2^{p-1} E \sup_{s\in[t_0,t]} \left| \int_{t_0}^{s} \sum_{k:\, t_k \leqslant s} \left[f(x_u, u) - F\left(x_u^n, u; , x_{t_k}^n\right) \right] I_{[t_k, t_{k+1})}(u)\, du \right|^p \\
&\quad + 2^{p-1} E \sup_{s\in[t_0,t]} \left| \int_{t_0}^{s} \sum_{k:\, t_k \leqslant s} \left[g(x_u, u) - G\left(x_u^n, u; x_{t_k}^n\right) \right] I_{[t_k, t_{k+1})}(u)\, dw_u \right|^p \\
&\leqslant 2^{p-1} (t - t_0)^{p-1} E \int_{t_0}^{t} \left| \sum_{k:\, t_k \leqslant t} \left[f(x_u, u) - F\left(x_u^n, u; x_{t_k}^n\right) \right] I_{[t_k, t_{k+1})}(u) \right|^p du \\
&\quad + c_p 2^{p-1} (t - t_0)^{\frac{p}{2}-1} E \int_{t_0}^{t} \left| \sum_{k:\, t_k \leqslant t} \left[g(x_u, u) - G\left(x_u^n, u; x_{t_k}^n\right) \right] I_{[t_k, t_{k+1})}(u) \right|^p du.
\end{aligned} \tag{19}$$

Let $j = \max\{i \in \{0, 1, 2, \ldots, n-1\},\ t_i \leqslant t \leqslant T\}$. Denote that

$$J_{t_k,t}(u) = \left[f(x_u, u) - F\left(x_u^n, u; x_{t_k}^n\right) \right] I_{[t_k,t)}(u),$$
$$\tilde{J}_{t_k,t}(u) = \left[g(x_u, u) - G\left(x_u^n, u; x_{t_k}^n\right) \right] I_{[t_k,t)}(u).$$

Then, (19) can be written as

$$E \sup_{s\in[t_0-\tau,t]} \left|x(s) - x^n(s)\right|^p \leqslant 2^{p-1} (t - t_0)^{p-1} \int_{t_0}^{t} E \left| \sum_{k=0}^{j-1} J_{t_k,t_{k+1}}(u) + J_{t_j,t}(u) \right|^p du + 2^{p-1} c_p (t - t_0)^{\frac{p}{2}-1} \int_{t_0}^{t} E \left| \sum_{k=0}^{j-1} \tilde{J}_{t_k,t_{k+1}}(u) + \tilde{J}_{t_j,t}(u) \right|^p du. \tag{20}$$

Since

$$\sum_{k=0}^{j-1} J_{t_k,t_{k+1}}(u) + J_{t_j,t}(u) = \begin{cases} f(x_u, u) - F(x_u^n, u; x_{t_k}^n), & u \in [t_k, t_{k+1}), \\ f(x_u, u) - F(x_u^n, u; x_{t_j}^n), & u \in [t_j, t), \end{cases}$$

$$\sum_{k=0}^{j-1} \tilde{J}_{t_k,t_{k+1}}(u) + \tilde{J}_{t_j,t}(u) = \begin{cases} g(x_u, u) - G(x_u^n, u, x_{t_k}^n), & u \in [t_k, t_{k+1}), \\ g(x_u, u) - G(x_u^n, u; x_{t_j}^n), & u \in [t_j, t), \end{cases}$$

whenever $k = 0, 1, \ldots, j-1$, the relation (20) becomes

$$E \sup_{s\in[t_0-\tau,t]} \left|x(s) - x^n(s)\right|^p \leqslant 2^{p-1} (T - t_0)^{p-1} \sum_{k=0}^{j-1} \int_{t_k}^{t_{k+1}} E\left| f(x_u, u) - F\left(x_u^n, u; x_{t_k}^n\right) \right|^p du + 2^{p-1} (T - t_0)^{p-1} \int_{t_j}^{t} E\left| f(x_u, u) - F\left(x_u^n, u; x_{t_j}^n\right) \right|^p du$$

$$+ 2^{p-1} c_p (T - t_0)^{\frac{p}{2}-1} \sum_{k=0}^{j-1} \int_{t_k}^{t_{k+1}} E\big|g(x_u, u) - G\big(x_u^n, u; x_{t_k}^n\big)\big|^p \, du$$
$$+ 2^{p-1} c_p (T - t_0)^{\frac{p}{2}-1} \int_{t_j}^{t} E\big|g(x_u, u) - G\big(x_u^n, u; x_{t_j}^n\big)\big|^p \, du. \tag{21}$$

The application of the Lipschitz condition (3), the assumption $\mathcal{A}_2$ and Proposition 2 yield

$$\int_{t_k}^{t} E\big|f(x_u, u) - F\big(x_u^n, u; x_{t_k}^n\big)\big|^p \, du$$
$$\leqslant 2^{p-1}\left[\int_{t_k}^{t} E\big|f(x_u, u) - f\big(x_u^n, u\big)\big|^p \, du + \int_{t_k}^{t} E\big|f\big(x_u^n, u\big) - F\big(x_u^n, u; x_{t_k}^n\big)\big|^p \, du\right]$$
$$\leqslant 2^{p-1} K^p \int_{t_k}^{t} E\big\|x_u - x_u^n\big\|^p \, du + 2^{p-1} \int_{t_k}^{t} E\left\| \frac{f^{(m_1+1)}_{(x_{t_k}^n + \bar{\theta}(x_u^n - x_{t_k}^n), u)}(x_u^n - x_{t_k}^n, \ldots, x_u^n - x_{t_k}^n)}{(m_1+1)!} \right\|^p du$$
$$\leqslant 2^{p-1}\left[K^p \int_{t_k}^{t} E\big\|x_u - x_u^n\big\|^p \, du + \frac{L_1^p}{[(m_1+1)!]^p} \int_{t_k}^{t} E\big\|x_u^n - x_{t_k}^n\big\|^{(m_1+1)p} \, du\right]$$
$$\leqslant 2^{p-1}\left[K^p \int_{t_k}^{t} E\big\|x_u - x_u^n\big\|^p \, du + \frac{L_1^p B}{[(m_1+1)!]^p} n^{-(m_1+1)p/2} (t - t_k)\right], \tag{22}$$

whenever $k = 0, 1, \ldots, j$ and $t \in [t_k, t_{k+1}]$. Analogously,

$$\int_{t_k}^{t} E\big|g(x_u, u) - G\big(x_u^n, u; x_{t_k}^n\big)\big|^p \, du \leqslant 2^{p-1}\left[K^p \int_{t_k}^{t} E\big\|x_u - x_u^n\big\|^p \, du + \frac{L_2^p B}{[(m_2+1)!]^p} n^{-(m_2+1)p/2} (t - t_k)\right]. \tag{23}$$

Now, the estimates (22) and (23) together with (21) yield

$$E \sup_{s \in [t_0 - \tau\, t]} \big|x(s) - x^n(s)\big|^p \leqslant \alpha_1 \int_{t_0}^{t} E\big\|x_u - x_u^n\big\|^p \, du + \alpha_2 n^{-(m+1)p/2} (t - t_0), \tag{24}$$

where $m = \min\{m_1, m_2\}$ and α_1, α_2 are generic constants independent of n.

In order to estimate the $E\|x_u - x_u^n\|^p$, we distinguish two cases. First, let $u - \tau < t_0$. Then,

$$E\big\|x_u - x_u^n\big\|^p \leqslant E \sup_{\theta \in [-\tau, 0]} \big|x(u+\theta) - x^n(u+\theta)\big|^p = E \sup_{r \in [u-\tau, u]} \big|x(r) - x^n(r)\big|^p$$
$$\leqslant E \sup_{r \in [u-\tau, t_0]} \big|x(r) - x^n(r)\big|^p + E \sup_{r \in [t_0, u]} \big|x(r) - x^n(r)\big|^p = E \sup_{r \in [t_0, u]} \big|x(r) - x^n(r)\big|^p$$
$$\leqslant E \sup_{r \in [t_0 - \tau, u]} \big|x(r) - x^n(r)\big|^p.$$

Let $u - \tau \geqslant t_0$. Then,

$$E\big\|x_u - x_u^n\big\|^p \leqslant E \sup_{r \in [u-\tau, u]} \big|x(r) - x^n(r)\big|^p \leqslant E \sup_{r \in [t_0 - \tau, u]} \big|x(r) - x^n(r)\big|^p.$$

The last estimates and (24) imply

$$E \sup_{s \in [t_0 - \tau, t]} \big|x(s) - x^n(s)\big|^p \leqslant \alpha_1 \int_{t_0}^{t} E \sup_{r \in [t_0 - \tau, u]} \big|x(r) - x^n(r)\big|^p \, du + \alpha_2\, n^{-(m+1)p/2} (t - t_0).$$

The application of the Gronwall–Bellman lemma gives

$$E \sup_{s\in[t_0-\tau,t]} \left|x(s)-x^n(s)\right|^p \leqslant \alpha_2 n^{-(m+1)p/2}(T-t_0)e^{\alpha_1(T-t_0)} \equiv Hn^{-(m+1)p/2},$$

where H is a constant. Since the last inequality holds for all $t \in [t_0, T]$, it follows that

$$E \sup_{s\in[t_0-\tau,T]} \left|x(s)-x^n(s)\right|^p \leqslant Hn^{-(m+1)p/2},$$

which completes the proof. □

Therefore, $E\sup_{t\in[t_0-\tau,T]} |x(t)-x^n(t)|^p \to 0$ as $n\to\infty$, that is, $x^n \xrightarrow{L^p} x$ as $n\to\infty$. Let us note that the rate of the convergence decreases if the degrees of Taylor expansions of the functionals f and g increase, similarly to Taylor approximation in real analysis. Moreover, the following assertion is a direct consequence of Theorem 1, that the approximate solutions x^n pathwise converge to the solution x of Eq. (8).

Theorem 2. *Let the conditions of Theorem* 1 *be satisfied. Then, the sequence* $\{x^n,\ n\in N\}$ *of approximate solutions determined with Eq.* (10)*, pathwise converges to the solution x of Eq.* (8)*.*

Proof. By applying Chebyshev inequality and Theorem 1, we find for an arbitrary $\eta > 0$ that

$$\sum_{n=1}^{\infty} P\Big(\sup_{t\in[t_0-\tau,T]} \left|x(t)-x^n(t)\right|^{\frac{p}{2}} \geqslant n^{-\eta}\Big) \leqslant \sum_{n=1}^{\infty} E \sup_{t\in[t_0-\tau,T]} \left|x(t)-x^n(t)\right|^p \cdot n^{2\eta} \leqslant H\sum_{n=1}^{\infty} n^{-[(m+1)p-4\eta]/2}.$$

The series of the right-hand side converges if we choose, for example, $\eta < 1/2$ for $p=2$, and $\eta < (p/2-1)/2$ for $p>2$. Then, $x^n \xrightarrow{\text{a.s.}} x$ as $n\to\infty$, in view of the Borel–Cantelli lemma. □

3. Conclusions and remarks

• Any other form of the residuum in Taylor expansion for the functionals f and g could be used instead of the Lagrange form (6). The residuum of the Cauchy form $W(x,h)=\frac{1}{n!}T^{(n+1)}_{(x+th)}((1-t)h,\dots,(1-t)h,h)$ for some $t\in(0,1)$, for example, can be estimated with $\|W(x,h)\| \leqslant \frac{1}{n!}\|T^{(n+1)}_{(x+th)}\|\cdot\|h\|^{n+1}$, and the residuum in integral form, $W(x,h)=\frac{1}{n!}\int_0^1 T^{(n+1)}_{(a+th)}((1-t)h,\dots,(1-t)h,h)\,dt$ for some $t\in(0,1)$, with $\|W(x,h)\| \leqslant \frac{1}{n!}\sup_{t\in[0,1]}\|T^{(n+1)}_{(x+th)}\|\cdot\|h\|^{n+1}$. Then, Propositions 1 and 2 and Theorem 1 hold under the same assumptions, while the appropriate estimates differ only in constants.

• The rate of the closeness between the solution to Eq. (1) and the approximate solution x^n could be improved by using a sequence $\{q_n,\ n\in N\}$, $q_n>n$, and a partition with equidistant discretization points $t_k=t_0+k\delta_n$, $k=0,1,\dots,q_n$, with the time step $\delta_n=(T-t_0)/q_n$.

• If Eq. (1) is autonomous, the presented approximate method is reduced for $m=0$ to the well-known Euler–Maruyama method (see [12, for example]).

• In papers [6,7] the proofs of the assertions are based on the Ito formula and on some difference inequalities. Recall that similar procedures do not yield desired results in this paper. For this reason, although analogous problems are studied here, the proofs of the assertions are completely different with respect to the ones in the previously cited papers.

• The method presented in the paper could be appropriately extended to stochastic functional differential equations including martingales and martingale measures instead of the Brownian motion process.

• The fact that the rate of convergence between the solution x to Eq. (1) and the approximate solution x^n decreases when the degrees of Taylor expansions for f and g increase, indicates that it would be convenient to combine the presented analytic method with numerical approximations based on Ito–Taylor expansions of higher degrees, described, above all, by Kloeden and Platen [8,9]. Moreover, in order to derive numerical schemes of higher order, it seems to be reasonable to replace the solutions in approximate equations, that is, in Taylor polynomials, by lower order approximations, analogously to the recent papers [10,11] by Kloeden and Jentzen treating random ordinary differential equations.

References

[1] M.A. Atalla, Finite-difference approximations for stochastic differential equations, in: Probabilistic Methods for the Investigation of Systems with an Infinite Number of Degrees of Freedom, Inst. of Math. Acad. of Science USSR, Kiev, 1986, pp. 11–16 (in Russian).
[2] M.A. Atalla, On one approximating method for stochastic differential equations, in: Asymptotic Methods for the Theory of Stochastic Processes, Inst. of Math. Acad. of Science USSR, Kiev, 1987, pp. 15–22 (in Russian).
[3] E. Buckwar, One-step approximation for stochastic functional differential equations, Appl. Numer. Math. 56 (2006) 667–681.
[4] D.H. Griffel, Applied Functional Analysis, Dover Publications, 2002.
[5] V. Huston, J.S. Pym, M.J. Cloud, Applications of Functional Analysis and Operator Theory, Elsevier, 2005.
[6] S. Janković, D. Ilić, An analytic approximation of solutions of stochastic differential equations, Comput. Math. Appl. 47 (2004) 903–912.
[7] S. Janković, D. Ilić, An analytic approximate method for solving stochastic integrodifferential equations, J. Math. Anal. Appl. 320 (2006) 230–245.
[8] P.E. Kloeden, E. Platen, Numerical Solution of Stochastic Differential Equations, third ed., Springer-Verlag, Berlin, 1999.
[9] P.E. Kloeden, E. Platen, H. Schurz, Numerical Solution of SDE through Computer Experiments, third ed., Springer-Verlag, Berlin, 1999.

[10] P.E. Kloeden, A. Jentzen, Pathwise convergent higher order numerical schemes for random ordinary differential equations, Proc. R. Soc. Lond. Ser. A 463 (2007) 2929–2944.
[11] A. Jentzen, P.E. Kloeden, Pathwise Taylor schemes for random ordinary differential equations, BIT 49 (1) (2009) 113–140.
[12] X. Mao, Stochastic Differential Equations and Applications, second ed., Horvood, Chichester, UK, 2007.
[13] X. Mao, Exponential Stability for Stochastic Differential Equations, Marcel Dekker, 1994.
[14] X. Mao, Numerical solutions of stochastic functional differential equations, London Math. Soc. 6 (2003) 14–161.
[15] S.E.A. Mohammed, Stochastic Functional Differential Equations, Longman Scientific and Technical, Harlow, UK, 1986.