RESEARCH

Open Access

Random non-autonomous second order linear differential equations: mean square analytic solutions and their statistical properties

J. Calatayud[1], J.-C. Cortés[1]* , M. Jornet[1] and L. Villafuerte[2]

*Correspondence:
jccortes@imm.upv.es
[1] Instituto Universitario de
Matemática Multidisciplinar,
Universitat Politècnica de València,
Valencia, Spain
Full list of author information is
available at the end of the article

Abstract

In this paper we study random non-autonomous second order linear differential equations by taking advantage of the powerful theory of random difference equations. The coefficients are assumed to be stochastic processes, and the initial conditions are random variables both defined in a common underlying complete probability space. Under appropriate assumptions established on the data stochastic processes and on the random initial conditions, and using key results on difference equations, we prove the existence of an analytic stochastic process solution in the random mean square sense. Truncating the random series that defines the solution process, we are able to approximate the main statistical properties of the solution, such as the expectation and the variance. We also obtain error *a priori* bounds to construct reliable approximations of both statistical moments. We include a set of numerical examples to illustrate the main theoretical results established throughout the paper. We finish with an example where our findings are combined with Monte Carlo simulations to model uncertainty using real data.

MSC: 39A50; 37H10; 34F05; 60H10; 60H35; 93E03; 65C05

Keywords: Random second order linear difference and differential equation; Analytic second order stochastic process; $L^p(\Omega)$ random calculus; Uncertainty quantification; Mathematical modelling and stochastic computation

1 Introduction

In this paper, we take advantage of the powerful theory of random difference equations to conduct a full probabilistic study to the random second order linear differential equation

$$\begin{cases} \ddot{X}(t) + A(t)\dot{X}(t) + B(t)X(t) = 0, & t \in \mathbb{R}, \\ X(t_0) = Y_0, \\ \dot{X}(t_0) = Y_1. \end{cases} \quad (1)$$

The data coefficients $A(t)$ and $B(t)$ are stochastic processes and the initial conditions Y_0 and Y_1 are random variables on an underlying complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$. In the triple of the probability space, Ω is the sample space, which consists of outcomes that will

be generically denoted by ω; $\mathcal{F}$ is a σ-algebra of events; and $\mathbb{P}$ is a probability measure. Naturally, the solution of (1), $X(t)$, is a stochastic process as well. Actually, we should write $A(t,\omega)$, $B(t,\omega)$, $Y_0(\omega)$ and $Y_1(\omega)$, but to simplify the notation, and in concordance with the literature, we will hide the generic outcome ω and just write $A(t)$, $B(t)$, Y_0 and Y_1 instead.

The aim of this paper is, in the first step, to specify the meaning of random differential equation (1) via the $L^p(\Omega)$ random calculus or, more concretely, using the so-called mean square calculus that corresponds to $p = 2$; secondly, to find a proper stochastic process solution to (1); and thirdly, to compute its main statistical information (expectation and variance) under mild conditions.

Particular cases of the random initial value problem (1) have been studied in previous contributions using $L^p(\Omega)$ random calculus. For instance, important deterministic models appearing in the area of mathematical physics like Airy, Hermite, Legendre, Laguerre and Bessel differential equations have been randomized and rigorously studied in [3, 4, 7, 8, 10], respectively. In these contributions, approximate solution stochastic processes together with their main statistical moments (mean and variance) are constructed by taking advantage of the random mean square calculus. Since in the case of Hermite, Legendre and Laguerre deterministic differential equations it is well known that they admit polynomial solutions, in those contributions the concept of random polynomial solution is introduced in the stochastic framework as well. In [15], the authors propose a homotopy technique to solve some particular random differential equations belonging to the class given in (1). A very important case of problem (1) is when its coefficients are random variables rather than stochastic processes, i.e., $A(t) = A$ and $B(t) = B$, corresponding to the autonomous case. In [6] the authors construct approximations of the first and second probability density function of the solution stochastic process using a complementary approach to mean square calculus. In [13, 14, 18–20, 33] one addresses significant advances for other random differential equations, dealing with the computation of the probability density function of the corresponding solution. Additional studies dealing with random differential equations via random mean square calculus include [12, 21–24, 26–28, 31, 36], for instance.

The structure of this paper is described as follows. In Sect. 2 we revise the notation and the theory of $L^p(\Omega)$ calculus necessary to understand the paper. In Sect. 3 we solve the random initial value problem (1) in a suitable way, and we describe the manner of approximating the main statistical information of the solution process (mean and variance). In Sect. 4 we compare our findings with the existing literature, and we also perform some numerical examples including an illustrative application in a modeling setting using real data. Finally, in Sect. 5 conclusions are drawn.

2 Preliminaries

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete probability space. In this paper we work with random variables $X : \Omega \to \mathbb{R}$ that belong to the so-called Lebesgue spaces $L^p(\Omega)$. Recall that we say that $X \in L^p(\Omega)$, $1 \le p < \infty$, if the norm

$$\|X\|_{L^p(\Omega)} := \left(\int_\Omega |X|^p \, d\mathbb{P} \right)^{\frac{1}{p}}$$

is finite. We say that $X \in L^\infty(\Omega)$ if the norm

$$\|X\|_{L^\infty(\Omega)} := \inf\big\{\sup\big\{\big|X(\omega)\big| : \omega \in \Omega \backslash N\big\} : \mathbb{P}(N) = 0\big\}$$

is finite (this norm is usually termed *essential supremum*). These spaces are Banach, and the particular case of $L^2(\Omega)$ is a Hilbert space. Sometimes, when we refer to convergence in $L^2(\Omega)$, we will say that the convergence holds in the mean square (m.s.) sense.

In general, the statistical information of the random variable X is given by a set of operators that give information about the distribution of X. In this paper we will deal with the expectation, $\mathbb{E}[X] = \int_\Omega X \, d\mathbb{P}$, and with the variance, $\mathbb{V}[X] = \mathbb{E}[(X - \mathbb{E}[X])^2]$. Mean square convergence is important because it preserves the limit of the expectation and variance: if $\{X_n\}_{n=1}^\infty$ is a sequence of random variables that converges in $L^2(\Omega)$ to X (that is, m.s. convergent), then

$$\lim_{n\to\infty} \mathbb{E}[X_n] = \mathbb{E}[X], \qquad \lim_{n\to\infty} \mathbb{V}[X_n] = \mathbb{V}[X] \tag{2}$$

(see Theorem 4.3.1 in [37]).

A very important inequality concerning the norm of a product of random variables is Hölder's inequality: for any two random variables X and Y, we have $\|XY\|_{L^r(\Omega)} \leq \|X\|_{L^p(\Omega)}\|Y\|_{L^q(\Omega)}$, where $1 \leq r, p, q \leq \infty$ and $1/r = 1/p + 1/q$. When $r = 1$ and $p = q = 2$, the inequality is known as Cauchy–Schwarz inequality. Another stochastic result that will be used throughout this paper is Jensen's inequality: if f is a convex function on $\mathbb{R}$ and assuming that the expectations $\mathbb{E}[X]$ and $\mathbb{E}[f(X)]$ both exist and are finite, then $f(\mathbb{E}[X]) \leq \mathbb{E}[f(X)]$. In particular, if $f(z) = |z|$ one gets $|\mathbb{E}[X]| \leq \mathbb{E}[|X|]$.

In this paper we will also deal with stochastic processes $X = \{X(t, \omega) : t \in I, \omega \in \Omega\}$, where $I \subseteq \mathbb{R}$. To simplify notation, we will just write X or $X(t)$ and we will make the dependence on $\omega \in \Omega$ implicit. Fixed $\omega \in \Omega$, the stochastic process $X(t)$ can be seen as a real mapping from $I \subseteq \mathbb{R}$ to $\mathbb{R}$, so any concept of real calculus, such as continuity, differentiability, etc., can be defined for the stochastic process.

However, sometimes it is more suitable to work with $L^p(\Omega)$ random calculus. In the case of $p = 2$, it is termed mean square calculus. To read a full exposition on this topic, see [37, Ch. 4], [29, Ch. XI], or [30, Ch. 5]. In [38] the authors combine $L^2(\Omega)$ and $L^4(\Omega)$, corresponding to mean square and mean fourth random calculus, to solve random differential equations.

We say that the stochastic process X is in $L^p(\Omega)$ if the random variable $X(t)$ belongs to $L^p(\Omega)$ for all $t \in I$. For such processes, we say that X is differentiable in the $L^p(\Omega)$ sense at $t_0 \in I$ if there exists a random variable $\dot{X}(t_0)$ in $L^p(\Omega)$ such that

$$\lim_{h\to 0}\left\|\frac{X(t_0 + h) - X(t_0)}{h} - \dot{X}(t_0)\right\|_{L^p(\Omega)} = 0.$$

The random variable $\dot{X}(t_0)$ is called the $L^p(\Omega)$ derivative of $X(t)$ at t_0. We say that the stochastic process X is differentiable on I in the $L^p(\Omega)$ sense if it is differentiable at every $t_0 \in I$.

The stochastic process X is analytic at t_0 if $X(t) = \sum_{n=0}^\infty X_n(t - t_0)^n$ for every t in a neighbourhood of t_0, where $X_0, X_1, \ldots$ are random variables and the sum is in the topology of $L^p(\Omega)$.

Thus, when we deal with the random differential equation (1), we will understand the derivatives in an $L^p(\Omega)$ random sense. More concretely, as we will see, the correct setting will be $L^2(\Omega)$ and the use of the mean square calculus.

An important difference with respect to the deterministic scenario when solving a random differential equation is the computation of the main statistical functions associated to the solution stochastic process, such as the mean and the variance functions.

3 Results

Our main goal is to find the solution stochastic process to the random initial value problem (1). We will assume that the data stochastic process $A(t)$ and $B(t)$ are analytic at t_0, in the following sense:

$$A(t) = \sum_{n=0}^{\infty} A_n(t - t_0)^n, \qquad B(t) = \sum_{n=0}^{\infty} B_n(t - t_0)^n,$$

for $t \in (t_0 - r, t_0 + r)$, being $r > 0$ fixed, and the sum is understood in the $\mathrm{L}^2(\Omega)$ setting. We search for an analytic solution process $X(t)$ of the form

$$X(t) = \sum_{n=0}^{\infty} X_n(t - t_0)^n,$$

for $t \in (t_0 - r, t_0 + r)$, where the sum is in $\mathrm{L}^2(\Omega)$. This stochastic process will be a solution to the random problem (1) in the sense of $\mathrm{L}^2(\Omega)$ (so, in particular twice differentiable in the mean square sense).

3.1 Auxiliary results concerning random power series

We need some auxiliary results to deal with random power series in the $\mathrm{L}^2(\Omega)$ setting. First of all, we need a result to differentiate a power series in the $\mathrm{L}^p(\Omega)$ sense (in this paper we will just use the cases $p = 1$ and $p = 2$, but we do the proof for a general $p \geq 1$ just for the sake of completeness). The particular case $p = 2$ is a consequence of Theorem 3.1 in [11].

Theorem 3.1 (Differentiation of a random power series in the $\mathrm{L}^p(\Omega)$ sense) *Let $A(t) = \sum_{n=0}^{\infty} A_n(t - t_0)^n$ be a random power series in the $\mathrm{L}^p(\Omega)$ setting $(p \geq 1)$ for $t \in (t_0 - r, t_0 + r)$, $r > 0$. Then the random power series $\sum_{n=1}^{\infty} nA_n(t - t_0)^{n-1}$ exists in $\mathrm{L}^p(\Omega)$ for $t \in (t_0 - r, t_0 + r)$; moreover, the $\mathrm{L}^p(\Omega)$ derivative of $A(t)$ is equal to it*:

$$\dot{A}(t) = \sum_{n=1}^{\infty} nA_n(t - t_0)^{n-1}$$

for all $t \in (t_0 - r, t_0 + r)$.

Proof Let us see first that the random power series $\sum_{n=1}^{\infty} nA_n(t - t_0)^{n-1}$ exists in $\mathrm{L}^p(\Omega)$ for $t \in (t_0 - r, t_0 + r)$. Given $0 < \rho < r$, we prove that

$$\sum_{n=1}^{\infty} n\|A_n\|_{\mathrm{L}^p(\Omega)}\rho^{n-1} < \infty. \tag{3}$$

Fix $s : 0 < \rho < s < r$. Since the sum $\sum_{n=0}^{\infty} A_n s^n$ exists in $\mathrm{L}^p(\Omega)$, $\lim_{n \to \infty} \|A_n\|_{\mathrm{L}^p(\Omega)} s^n = 0$, so there exists $K > 0$ such that $\|A_n\|_{\mathrm{L}^p(\Omega)} s^n \leq K$ for every $n \geq 0$. Then

$$n\|A_n\|_{\mathrm{L}^p(\Omega)}\rho^{n-1} \leq nK\frac{1}{s}\left(\frac{\rho}{s}\right)^{n-1}.$$

As $\sum_{n=1}^{\infty} n(\rho/s)^{n-1} < \infty$, by the comparison test for series, we conclude that (3) holds.

Let us see now that the $L^p(\Omega)$ derivative of $A(t)$ is equal to $\sum_{n=1}^{\infty} nA_n(t - t_0)^{n-1}$. By the definition of $L^p(\Omega)$ derivative, we have to check that

$$\lim_{h \to 0} \left\| \sum_{n=0}^{\infty} A_n \frac{(t + h - t_0)^n - (t - t_0)^n}{h} - \sum_{n=1}^{\infty} nA_n(t - t_0)^{n-1} \right\|_{L^p(\Omega)} = 0.$$

Fix s and h such that $0 < \rho < s < r$ and $0 < |h| < s - \rho$. Fix $t \in (t_0 - \rho, t_0 + \rho)$. By the triangular inequality,

$$\left\| \sum_{n=0}^{\infty} A_n \frac{(t + h - t_0)^n - (t - t_0)^n}{h} - \sum_{n=1}^{\infty} nA_n(t - t_0)^{n-1} \right\|_{L^p(\Omega)}$$
$$\leq \sum_{n=1}^{\infty} \|A_n\|_{L^p(\Omega)} \left| \frac{(t + h - t_0)^n - (t - t_0)^n}{h} - n(t - t_0)^{n-1} \right|. \tag{4}$$

We know that

$$\lim_{h \to 0} \left| \frac{(t + h - t_0)^n - (t - t_0)^n}{h} - n(t - t_0)^{n-1} \right| = 0 \tag{5}$$

(by the definition of the pointwise derivative of $(t - t_0)^n$). On the other hand, using the identity $a^n - b^n = (a - b)(\sum_{m=0}^{n-1} a^{n-1-m} b^m)$, we perform the following estimates:

$$\left| \frac{(t + h - t_0)^n - (t - t_0)^n}{h} \right| = \left| \frac{(t + h - t_0)^n - (t - t_0)^n}{(t + h - t_0) - (t - t_0)} \right|$$
$$= \left| \sum_{m=0}^{n-1} (t + h - t_0)^{n-1-m}(t - t_0)^m \right|$$
$$\leq \sum_{m=0}^{n-1} |t + h - t_0|^{n-1-m} |t - t_0|^m$$
$$\leq \sum_{m=0}^{n-1} \left(|t - t_0| + |h| \right)^{n-1-m} |t - t_0|^m$$
$$\leq n \left(|t - t_0| + |h| \right)^{n-1} \leq ns^{n-1},$$

where we have used that $|t - t_0| \leq |t - t_0| + |h|$, $|t - t_0| < \rho$ and $|h| < s - \rho$. Then

$$\|A_n\|_{L^p(\Omega)} \left| \frac{(t - t_0 + h)^n - (t - t_0)^n}{h} - n(t - t_0)^{n-1} \right|$$
$$\leq \|A_n\|_{L^p(\Omega)} \left(\left| \frac{(t - t_0 + h)^n - (t - t_0)^n}{h} \right| + n|t - t_0|^{n-1} \right)$$
$$\leq \|A_n\|_{L^p(\Omega)} \left(ns^{n-1} + n\rho^{n-1} \right) \leq \|A_n\|_{L^p(\Omega)} \left(ns^{n-1} + ns^{n-1} \right) = 2\|A_n\|_{L^p(\Omega)} ns^{n-1}, \tag{6}$$

being $\sum_{n=1}^{\infty} \|A_n\|_{L^p(\Omega)} ns^{n-1} < \infty$. By the dominated convergence theorem for series, both (5) and (6) permit us to conclude that (4) tends to 0 as $h \to 0$, as wanted. $\qquad \square$

As it shall be apparent later, we need a theorem to multiply random power series. In the deterministic setting, the so-called Merten's theorem allows multiplying two power series. The deterministic version of Merten's theorem is proved in Theorem 8.46 of [1]. We adapt the proof in [1] to a stochastic setting in Theorem 3.2. Notice that, from Theorem 3.2, when we multiply two random power series, we lose Lebesgue spaces of convergence. As we will see in the proof, this fact will be a consequence of the Cauchy–Schwarz inequality (for two real numbers u and v, we always have $|uv| = |u||v|$; however, for random variables U and V, we do not have $\|UV\|_{L^2(\Omega)} = \|U\|_{L^2(\Omega)}\|V\|_{L^2(\Omega)}$ in general, but $\|UV\|_{L^1(\Omega)} \leq \|U\|_{L^2(\Omega)}\|V\|_{L^2(\Omega)}$).

Theorem 3.2 (Merten's theorem for random series in the mean square sense) *Let $U = \sum_{n=0}^{\infty} U_n$ and $V = \sum_{n=0}^{\infty} V_n$ be two random series that converge in $L^2(\Omega)$. Suppose that one of the series converges absolutely, say $\sum_{n=0}^{\infty} \|V_n\|_{L^2(\Omega)} < \infty$. Then*

$$\left(\sum_{n=0}^{\infty} U_n\right)\left(\sum_{n=0}^{\infty} V_n\right) = \sum_{n=0}^{\infty} W_n,$$

where

$$W_n = \sum_{m=0}^{n} U_{n-m} V_m$$

and $\sum_{n=0}^{\infty} W_n$ is understood in $L^1(\Omega)$. The series $\sum_{n=0}^{\infty} W_n$ is known as the Cauchy product of the series $\sum_{n=0}^{\infty} U_n$ and $\sum_{n=0}^{\infty} V_n$.

Proof Let us write the Nth partial sum of $\sum_{n=0}^{\infty} W_n$ in an appropriate way:

$$\sum_{n=0}^{N} W_n = \sum_{n=0}^{N}\sum_{m=0}^{n} U_{n-m} V_m = \sum_{m=0}^{N} V_m \sum_{n=m}^{N} U_{n-m} = \sum_{m=0}^{N} V_m \sum_{n=0}^{N-m} U_n$$

$$= \sum_{m=0}^{N} V_m \left(U - \sum_{n=N-m+1}^{\infty} U_n\right) = U\sum_{m=0}^{N} V_m - \sum_{m=0}^{N} V_m \sum_{n=N-m+1}^{\infty} U_n.$$

The first addend, $U\sum_{m=0}^{N} V_m$, tends to UV in $L^1(\Omega)$ as $N \to \infty$. In fact, observe that because of the Cauchy–Schwarz inequality, one gets

$$\left\|U\sum_{m=0}^{N} V_m - UV\right\|_{L^1(\Omega)} = \left\|U\left(\sum_{m=0}^{N} V_m - V\right)\right\|_{L^1(\Omega)}$$

$$\leq \|U\|_{L^2(\Omega)}\left\|\sum_{m=0}^{N} V_m - V\right\|_{L^2(\Omega)} \xrightarrow[N\to\infty]{} 0,$$

where we have used that $\|U\|_{L^2(\Omega)} < \infty$ (since $U \in L^2(\Omega)$ because it is the m.s. limit of the series $\sum_{n=0}^{\infty} U_n$) and that $\|\sum_{m=0}^{N} V_m - V\|_{L^2(\Omega)} \xrightarrow[N\to\infty]{0}$ (since by hypothesis $\sum_{n=0}^{\infty} V_n$ converges to V in $L^2(\Omega)$).

Thus, it only remains to prove that the second addend, $\sum_{m=0}^{N} V_m \sum_{n=N-m+1}^{\infty} U_n$, goes to 0 in $L^1(\Omega)$ as $N \to \infty$.

Since $\lim_{N \to \infty} \sum_{n=N}^{\infty} U_n = 0$ in $L^2(\Omega)$, there exists $L > 0$ such that

$$\left\| \sum_{n=N}^{\infty} U_n \right\|_{L^2(\Omega)} \leq L, \quad \forall N \in \mathbb{N}.$$

Let $K = \sum_{n=0}^{\infty} \| V_n \|_{L^2(\Omega)}$. Fix $\epsilon > 0$. We can take N_ϵ such that, for all $N \geq N_\epsilon$,

$$\left\| \sum_{n=N}^{\infty} U_n \right\|_{L^2(\Omega)} < \frac{\epsilon}{2K}, \qquad \sum_{n=N+1}^{\infty} \| V_n \|_{L^2(\Omega)} < \frac{\epsilon}{2L}.$$

Then, for $N \geq 2N_\epsilon$, by the triangular inequality and the Cauchy–Schwarz inequality,

$$
\begin{aligned}
\left\| \sum_{m=0}^{N} V_m \sum_{n=N-m+1}^{\infty} U_n \right\|_{L^1(\Omega)} &\leq \sum_{m=0}^{N_\epsilon} \| V_m \|_{L^2(\Omega)} \left\| \sum_{n=N-m+1}^{\infty} U_n \right\|_{L^2(\Omega)} \\
&\quad + \sum_{m=N_\epsilon+1}^{N} \| V_m \|_{L^2(\Omega)} \left\| \sum_{n=N-m+1}^{\infty} U_n \right\|_{L^2(\Omega)} \\
&\leq \frac{\epsilon}{2K} \sum_{m=0}^{N_\epsilon} \| V_m \|_{L^2(\Omega)} + L \sum_{m=N_\epsilon+1}^{N} \| V_m \|_{L^2(\Omega)} \\
&\leq \frac{\epsilon}{2K} \sum_{m=0}^{\infty} \| V_m \|_{L^2(\Omega)} + L \sum_{m=N_\epsilon+1}^{\infty} \| V_m \|_{L^2(\Omega)} \\
&\leq \frac{\epsilon}{2K} K + L \frac{\epsilon}{2L} = \epsilon.
\end{aligned}
$$

This shows that $\sum_{m=0}^{N} V_m \sum_{n=N-m+1}^{\infty} U_n$ tends to 0 in $L^1(\Omega)$ as $N \to \infty$. □

3.2 Main result: constructing the solution stochastic process of the random non-autonomous second order linear differential equation

We present the main theorem of this paper. After stating and proving it, a deeper analysis of the hypotheses and consequences of the theorem will be performed.

Theorem 3.3 *Let $A(t) = \sum_{n=0}^{\infty} A_n (t - t_0)^n$ and $B(t) = \sum_{n=0}^{\infty} B_n (t - t_0)^n$ be two random series in the $L^2(\Omega)$ setting, for $t \in (t_0 - r, t_0 + r)$, being $r > 0$ finite and fixed. Assume that the initial conditions Y_0 and Y_1 belong to $L^2(\Omega)$. Suppose that there is a constant $C_r > 0$, maybe dependent on r, such that $\| A_n \|_{L^\infty(\Omega)} \leq C_r / r^n$ and $\| B_n \|_{L^\infty(\Omega)} \leq C_r / r^n$, $n \geq 0$. Then the stochastic process $X(t) = \sum_{n=0}^{\infty} X_n (t - t_0)^n$, $t \in (t_0 - r, t_0 + r)$, where*

$$X_0 = Y_0, \qquad X_1 = Y_1, \tag{7}$$

$$X_{n+2} = \frac{-1}{(n+2)(n+1)} \sum_{m=0}^{n} \big[(m+1) A_{n-m} X_{m+1} + B_{n-m} X_m \big], \quad n \geq 0, \tag{8}$$

is the unique analytic solution to the random initial value problem (1) in the mean square sense.

Proof Suppose that $X(t) = \sum_{n=0}^{\infty} X_n(t - t_0)^n$ is a solution to (1) in the $L^2(\Omega)$ sense for $t \in (t_0 - r, t_0 + r)$, $r > 0$. By Theorem 3.1 with $p = 2$, the mean square derivatives of $X(t)$ are given by

$$\dot{X}(t) = \sum_{n=1}^{\infty} n X_n(t - t_0)^{n-1} = \sum_{n=0}^{\infty}(n + 1)X_{n+1}(t - t_0)^n,$$

$$\ddot{X}(t) = \sum_{n=2}^{\infty} n(n - 1) X_n(t - t_0)^{n-2} = \sum_{n=0}^{\infty}(n + 2)(n + 1)X_{n+2}(t - t_0)^n.$$

By the random Merten's Theorem 3.2,

$$A(t)\dot{X}(t) = \sum_{n=0}^{\infty}\left(\sum_{m=0}^{n} A_{n-m}(m + 1)X_{m+1}\right)(t - t_0)^n,$$

$$B(t)X(t) = \sum_{n=0}^{\infty}\left(\sum_{m=0}^{n} B_{n-m}X_m\right)(t - t_0)^n,$$

where these two random series converge in $L^1(\Omega)$. From $\ddot{X}(t) + A(t)\dot{X}(t) + B(t)X(t) = 0$,

$$\sum_{n=0}^{\infty}\left[(n + 2)(n + 1)X_{n+2} + \sum_{m=0}^{n}\left(A_{n-m}(m + 1)X_{m+1} + B_{n-m}X_m\right)\right](t - t_0)^n = 0, \qquad (9)$$

for $t \in (t_0 - r, t_0 + r)$, where the random series is again in the $L^1(\Omega)$ sense. By Theorem 3.1 with $p = 1$, differentiating (9) over and over again in the $L^1(\Omega)$ sense and evaluating at $t = t_0$ yield

$$(n + 2)(n + 1)X_{n+2} + \sum_{m=0}^{n}\left(A_{n-m}(m + 1)X_{m+1} + B_{n-m}X_m\right) = 0.$$

Isolating X_{n+2} we obtain the recursive expression (8):

$$X_{n+2} = \frac{-1}{(n + 2)(n + 1)}\sum_{m=0}^{n}\left[(m + 1)A_{n-m}X_{m+1} + B_{n-m}X_m\right].$$

The initial conditions of the random initial value problem (1) give (7), and so $X(t)$ is uniquely determined with probability 1.

It remains to check that the random series $\sum_{n=0}^{\infty} X_n(t - t_0)^n$ is convergent in $L^2(\Omega)$. For that purpose, we will make use of the $L^{\infty}(\Omega)$ bounds for A_n and B_n quoted in the hypotheses.

From the hypothesis $Y_0, Y_1 \in L^2(\Omega)$ and by induction on n in expression (8), we obtain that $X_n \in L^2(\Omega)$ for all $n \geq 0$. By the triangular inequality, the hypotheses and Hölder's inequality with $r = 2$, $p = \infty$ and $q = 2$ (see the notation used in Sect. 2),

$$\|X_{n+2}\|_{L^2(\Omega)} \leq \frac{1}{(n + 2)(n + 1)}\sum_{m=0}^{n}\left[(m + 1)\|A_{n-m}X_{m+1}\|_{L^2(\Omega)} + \|B_{n-m}X_m\|_{L^2(\Omega)}\right]$$

$$\leq \frac{1}{(n + 2)(n + 1)}\frac{C_r}{r^n}\sum_{m=0}^{n} r^m\left((m + 1)\|X_{m+1}\|_{L^2(\Omega)} + \|X_m\|_{L^2(\Omega)}\right). \qquad (10)$$

Define $H_0 := \|Y_0\|_{L^2(\Omega)}$, $H_1 := \|Y_1\|_{L^2(\Omega)}$ and

$$H_{n+2} := \frac{1}{(n+2)(n+1)} \frac{C_r}{r^n} \sum_{m=0}^{n} r^m \big((m+1)H_{m+1} + H_m\big). \tag{11}$$

By induction on n it is trivially seen that $\|X_n\|_{L^2(\Omega)} \leq H_n$ for $n \geq 0$. Thus, given $0 < \rho < r$, it is enough to see that $\sum_{n=0}^{\infty} H_n \rho^n < \infty$. For that purpose, we rewrite (11) so that H_{n+2} is expressed as a function of H_{n+1} and H_n (second order recurrence equation):

$$H_{n+2} = \frac{1}{(n+2)(n+1)} \frac{C_r}{r^n} \left(\sum_{m=0}^{n-1} r^m \big((m+1)H_{m+1} + H_m\big) + r^n \big((n+1)H_{n+1} + H_n\big) \right)$$

$$= \frac{1}{(n+2)(n+1)} \frac{C_r}{r^n} \frac{(n+1)n}{C_r} r^{n-1} \underbrace{\left(\frac{1}{(n+1)n} \frac{C_r}{r^{n-1}} \sum_{m=0}^{n-1} r^m \big((m+1)H_{m+1} + H_m\big) \right)}_{=H_{n+1}}$$

$$+ \frac{C_r}{n+2} H_{n+1} + \frac{C_r}{(n+2)(n+1)} H_n$$

$$= \left(\frac{n}{(n+2)r} + \frac{C_r}{n+2} \right) H_{n+1} + \frac{C_r}{(n+2)(n+1)} H_n. \tag{12}$$

Fix $s : 0 < \rho < s < r$. We have

$$H_{n+2} s^{n+2} = \left(\frac{ns}{(n+2)r} + \frac{C_r s}{n+2} \right) H_{n+1} s^{n+1} + \frac{C_r s^2}{(n+2)(n+1)} H_n s^n.$$

Let $M_n = \max_{0 \leq m \leq n} H_m s^m$. We have

$$H_{n+2} s^{n+2} \leq \left(\frac{ns}{(n+2)r} + \frac{C_r s}{n+2} + \frac{C_r s^2}{(n+2)(n+1)} \right) M_{n+1}. \tag{13}$$

Since

$$\lim_{n \to \infty} \frac{ns}{(n+2)r} + \frac{C_r s}{n+2} + \frac{C_r s^2}{(n+2)(n+1)} = \frac{s}{r} < 1,$$

it holds $M_{n+2} = M_{n+1}$ for all large n, and call the common value M. Hence, $H_n s^n \leq M$ for all large n, therefore $H_n \rho^n \leq M(\rho/s)^n$. Since $\sum_{n=0}^{\infty} (\rho/s)^n < \infty$, by comparison the series $\sum_{n=0}^{\infty} H_n \rho^n$ converges, and we are done. □

3.3 Comments on the hypotheses of the theorem

The hypotheses concerning the $L^\infty(\Omega)$ growth of the coefficients A_n and B_n, $n \geq 0$, may seem quite restrictive. However, these hypotheses have been necessary to relate the $L^2(\Omega)$ norm of the coefficients $X_0, X_1, X_2, \ldots$ in (10), then define the random variables $H_0, H_1, H_2, \ldots$ and finally bound $\|X_n\|_{L^2(\Omega)} \leq H_n$ by induction on $n \geq 0$. Without the hypotheses $\|A_n\|_{L^\infty(\Omega)} \leq C_r/r^n$ and $\|B_n\|_{L^\infty(\Omega)} \leq C_r/r^n$ of Theorem 3.3, this would not have been possible.

Moreover, these $L^\infty(\Omega)$ hypotheses are equivalent to a growth condition on the moments of the random variables $A_0, A_1, \ldots$ and $B_0, B_1, \ldots$. The key fact is that, for a given

random variable Z, we have that $\mathbb{E}[|Z|^n] \leq HR^n$ for certain $H > 0$ and $R > 0$ if and only if $\|Z\|_{L^\infty(\Omega)} \leq R$.

This key fact is a direct consequence of the following result: if Z is a random variable, then $\lim_{n\to\infty} \|Z\|_{L^n(\Omega)} = \|Z\|_{L^\infty(\Omega)}$. For the sake of completeness, we show the proof. If $\|Z\|_{L^\infty(\Omega)} < \infty$, define $\Omega_\delta = \{|Z| \geq \|Z\|_{L^\infty(\Omega)} - \delta\}$ for $\delta > 0$ small. By the definition of $L^\infty(\Omega)$, $\mathbb{P}(\Omega_\delta) > 0$. Now,

$$\|Z\|_{L^n(\Omega)} = \left(\int_\Omega |Z|^n \, d\mathbb{P}\right)^{\frac{1}{n}} \geq \left(\int_{\Omega_\delta} |Z|^n \, d\mathbb{P}\right)^{\frac{1}{n}} \geq \left(\|Z\|_{L^\infty(\Omega)} - \delta\right)\mathbb{P}(\Omega_\delta)^{\frac{1}{n}}.$$

Since $\mathbb{P}(\Omega_\delta) > 0$, we obtain

$$\liminf_{n\to\infty} \|Z\|_{L^n(\Omega)} \geq \|Z\|_{L^\infty(\Omega)} - \delta.$$

As $\delta > 0$ is arbitrary, $\liminf_{n\to\infty} \|Z\|_{L^n(\Omega)} \geq \|Z\|_{L^\infty(\Omega)}$. To prove the reverse inequality, write

$$\|Z\|_{L^n(\Omega)} = \left(\int_\Omega |Z|^n \, d\mathbb{P}\right)^{\frac{1}{n}} = \left(\int_\Omega |Z|^{n-1}|Z| \, d\mathbb{P}\right)^{\frac{1}{n}} \leq \|Z\|_{L^\infty(\Omega)}^{\frac{n-1}{n}} \|Z\|_{L^1(\Omega)}^{\frac{1}{n}}.$$

Since $\|Z\|_{L^1(\Omega)} \leq \|Z\|_{L^\infty(\Omega)} < \infty$, it holds $\lim_{n\to\infty} \|Z\|_{L^1(\Omega)}^{\frac{1}{n}} = 1$ (the trivial case $Z \equiv 0$ is obvious, we can assume $Z \not\equiv 0$). Then

$$\limsup_{n\to\infty} \|Z\|_{L^n(\Omega)} \leq \|Z\|_{L^\infty(\Omega)}.$$

This shows the result when $\|Z\|_{L^\infty(\Omega)} < \infty$. If $\|Z\|_{L^\infty(\Omega)} = \infty$, then one has to define $\Omega_L = \{|Z| \geq L\}$ for $L > 0$ large. Proceeding similarly and by the arbitrariness of L, one arrives at $\liminf_{n\to\infty} \|Z\|_{L^n(\Omega)} = \infty = \|Z\|_{L^\infty(\Omega)}$.

Growth hypotheses of the form $\mathbb{E}[|Z|^n] \leq HR^n$, for certain $H > 0$ and $R > 0$, are common in the literature to find stochastic analytic solutions to particular cases of (1). See, for example, Airy's random differential equation in [7] and Hermite's random differential equation in [3]. We have proved in this subsection that controlling the growth of the moments is equivalent to controlling the $L^\infty(\Omega)$ norm. Hence, Theorem 3.3 will allow us to generalize the results obtained in previous articles, for instance [3, 7]. See Example 4.1 and Example 4.2 for the generalization.

3.4 Relationship between the random pointwise and classical differential equations approaches

The hypotheses of Theorem 3.3, besides providing a stochastic solution to our problem (1), also give a pointwise classical solution to (1) under the additional assumption $Y_0, Y_1 \in L^\infty(\Omega)$. Indeed, by hypothesis we have $r^n \|A_n\|_{L^\infty(\Omega)} \leq C_r$ and $r^n \|B_n\|_{L^\infty(\Omega)} \leq C_r$. For fixed $\delta > 0$ small enough, one gets

$$\left(\frac{r}{1+\delta}\right)^n \|A_n\|_{L^\infty(\Omega)} \leq \frac{C_r}{(1+\delta)^n}, \qquad \left(\frac{r}{1+\delta}\right)^n \|B_n\|_{L^\infty(\Omega)} \leq \frac{C_r}{(1+\delta)^n}.$$

From these two inequalities,

$$\sum_{n=0}^{\infty} \|A_n\|_{L^{\infty}(\Omega)} (t - t_0)^n < \infty, \qquad \sum_{n=0}^{\infty} \|B_n\|_{L^{\infty}(\Omega)} (t - t_0)^n < \infty$$

for $t \in (t_0 - r/(1 + \delta), t_0 + r/(1 + \delta))$. Letting δ approach 0, we arrive at the uniform convergence of the series for $t \in (t_0 - r, t_0 + r)$. The solution X also satisfies $\sum_{n=0}^{\infty} \|X_n\|_{L^{\infty}(\Omega)} (t - t_0)^n < \infty$ for $t \in (t_0 - r, t_0 + r)$, because we can apply the triangular inequality in (10) with the $L^{\infty}(\Omega)$ norm instead of the $L^2(\Omega)$ norm. Therefore, $X(t)$ is a pointwise classical solution for each $\omega \in \Omega$ fixed (it is a solution from a deterministic point of view). This manner of studying random differential equations is referred to as the sample approach [37, Appendix I], [9].

3.5 Statistical information of the solution stochastic process: mean and variance

The expectation and variance of the stochastic process $X(t) = \sum_{n=0}^{\infty} X_n(t - t_0)^n$ given by (7)–(8) can be approximated. Indeed, first, one has to obtain X_n as a function of Y_0, Y_1, $A_0, \ldots, A_{n-1}$ and $B_0, \ldots, B_{n-1}$ by recursion via (8) for $n = 0, 1, \ldots, N$. After this, we construct a truncation

$$X_N(t) = \sum_{n=0}^{N} X_n(t - t_0)^n \tag{14}$$

of the solution stochastic process $X(t)$. Since $X_N(t) \to X(t)$ in $L^2(\Omega)$ as $N \to \infty$, according to the key property (2), we have

$$\lim_{N \to \infty} \mathbb{E}[X_N(t)] = \mathbb{E}[X(t)], \qquad \lim_{N \to \infty} \mathbb{V}[X_N(t)] = \mathbb{V}[X(t)].$$

As an example of the manner one can proceed, we show by hand some random coefficients X_n:

$$X_2 = \frac{-1}{2}(A_0 Y_1 + B_0 Y_0),$$

$$X_3 = \frac{-1}{6}(A_1 Y_1 + B_1 Y_0 + 2A_0 X_2 + B_0 Y_1) = \frac{-1}{6}\left(A_1 Y_1 + B_1 Y_0 - A_0^2 Y_1 - A_0 B_0 Y_0 + B_0 Y_1\right),$$

$$X_4 = \frac{-1}{12}(A_2 Y_1 + B_2 Y_0 + 2A_1 X_2 + B_1 X_1 + 3A_0 X_3 + B_0 X_2)$$

$$= \frac{-1}{12}\Big(A_2 Y_1 + B_2 Y_0 - A_1 A_0 Y_1 - A_1 B_0 Y_0 + B_1 Y_1 - \frac{1}{2}A_0 A_1 Y_1 - \frac{1}{2}A_0 B_1 Y_0$$

$$+ \frac{1}{2}A_0^3 Y_1 + \frac{1}{2}A_0^2 B_0 Y_0 - \frac{1}{2}A_0 B_0 Y_1 - \frac{1}{2}B_0 A_0 Y_1 - \frac{1}{2}B_0^2 Y_0\Big).$$

From these computations, we have truncation (14) for $N = 4$, $X_4(t) = Y_0 + Y_1 t + X_2 t^2 + X_3 t^3 + X_4 t^4$. We need to compute $\mathbb{E}[X_4(t)]$ to approximate $\mathbb{E}[X(t)]$. By linearity of the expectation, $\mathbb{E}[X_4(t)] = \mathbb{E}[Y_0] + \mathbb{E}[Y_1]t + \mathbb{E}[X_2]t^2 + \mathbb{E}[X_3]t^3 + \mathbb{E}[X_4]t^4$. Assuming independence of $Y_0, Y_1, A_0, A_1, \ldots, B_0, B_1, \ldots$ and applying a property from [16, p. 93], we are able to compute the expectation of the addends by hand:

$$\mathbb{E}[X_2] = \frac{-1}{2}\big(\mathbb{E}[A_0]\mathbb{E}[Y_1] + \mathbb{E}[B_0]\mathbb{E}[Y_0]\big),$$

$$\mathbb{E}[X_3] = \frac{-1}{6}\big(\mathbb{E}[A_1]\mathbb{E}[Y_1] + \mathbb{E}[B_1]\mathbb{E}[Y_0] - \mathbb{E}\big[A_0^2\big]\mathbb{E}[Y_1]$$
$$- \mathbb{E}[A_0]\mathbb{E}[B_0]\mathbb{E}[Y_0] + \mathbb{E}[B_0]\mathbb{E}[Y_1]\big),$$

$$\mathbb{E}[X_4] = \frac{-1}{12}\bigg(\mathbb{E}[A_2]\mathbb{E}[Y_1] + \mathbb{E}[B_2]\mathbb{E}[Y_0] - \mathbb{E}[A_1]\mathbb{E}[A_0]\mathbb{E}[Y_1] - \mathbb{E}[A_1]\mathbb{E}[B_0]\mathbb{E}[Y_0]$$
$$+ \mathbb{E}[B_1]\mathbb{E}[Y_1] - \frac{1}{2}\mathbb{E}[A_0]\mathbb{E}[A_1]\mathbb{E}[Y_1] - \frac{1}{2}\mathbb{E}[A_0]\mathbb{E}[B_1]\mathbb{E}[Y_0]$$
$$+ \frac{1}{2}\mathbb{E}\big[A_0^3\big]\mathbb{E}[Y_1] + \frac{1}{2}\mathbb{E}\big[A_0^2\big]\mathbb{E}[B_0]\mathbb{E}[Y_0] - \frac{1}{2}\mathbb{E}[A_0]\mathbb{E}[B_0]\mathbb{E}[Y_1]$$
$$- \frac{1}{2}\mathbb{E}[B_0]\mathbb{E}[A_0]\mathbb{E}[Y_1] - \frac{1}{2}\mathbb{E}\big[B_0^2\big]\mathbb{E}[Y_0]\bigg).$$

For large values of n, we need a computer to manage the big expressions for X_n, and as a consequence $\mathbb{E}[X_N(t)]$ and $\mathbb{V}[X_N(t)]$ for large values of N. We show how to implement the necessary formulas to compute the expectation and variance of the truncated series in the software Mathematica®. The recurrence relation (7)–(8) is defined as follows:

```
X[n_?NonPositive] := Y0;
X[1] = Y1;
X[n_] := -1/(n*(n - 1))*Sum[(m + 1)*A[n - 2 - m]*X[m + 1]
    + B[n - 2 - m]*X[m], {m, 0, n - 2}].
```

Truncation (14) is implemented by writing

```
seriesX[t_, t0_, N_] := X[0] + Sum[X[n]*(t - t0)^n, {n, 1, N}].
```

Using the `Expectation` function, in which one can set the distributions of `A[n]`, `B[n]`, `Y0` and `Y1`, both the expectation and variance of (14) can be calculated by the computer.

There are other approaches to approximating the expectation of the solution to the random initial value problem (1). One of these approaches is the so-called dishonest method [2], [17, p. 149], which assumes that $A(t)$ and $\dot{X}(t)$ are independent and that $B(t)$ and $X(t)$ are independent. Denoting $\mu_X(t) = \mathbb{E}[X(t)]$, the idea is that, since $\mathbb{E}[\ddot{X}(t)] = \frac{d^2}{dt^2}(\mu_X(t))$ and $\mathbb{E}[\dot{X}(t)] = \frac{d}{dt}(\mu_X(t))$, because of the commutation between the mean square limit and the expectation operator (see [37, Ch. 4]), by the assumed independence we arrive at a deterministic initial value problem to compute $\mu_X(t)$:

$$\begin{cases} \frac{d^2}{dt^2}(\mu_X(t)) + \mathbb{E}[A(t)]\frac{d}{dt}(\mu_X(t)) + \mathbb{E}[B(t)]\mu_X(t) = 0, \quad t \in \mathbb{R}, \\ \mu_X(t_0) = \mathbb{E}[Y_0], \\ \frac{d}{dt}(\mu_X(t_0)) = \mathbb{E}[Y_1]. \end{cases} \tag{15}$$

In [2] and [17, p. 149] this method is used to handle the problem of computing the expectation of the solution stochastic process of certain random differential equations. In [3, 7] approximations of the expectation of the corresponding solution stochastic process obtained via specific methods have been compared with the ones calculated by the dishonest approach. In our context, the dishonest method will work on cases where $\mathbb{C}\text{ov}[A(t), \dot{X}(t)]$ and $\mathbb{C}\text{ov}[B(t), X(t)]$ are small, but in general, there is no certainty that this may hold. In Example 4.1 and Example 4.2, we approximate $\mathbb{C}\text{ov}[A(t), \dot{X}(t)]$ and $\mathbb{C}\text{ov}[B(t), X(t)]$ in order to understand better the accuracy of the dishonest method. Nevertheless, the approximations via the truncation method previously described allow us to obtain reliable approximations for the expectation, and also for the variance, of the solution process.

Another popular approach consists in using Monte Carlo simulations. Sample from the distributions of $A(t)$, $B(t)$, Y_0 and Y_1 to obtain, say M realizations, for M large. That is, we have $A(t, \omega_1), \ldots, A(t, \omega_M)$, $B(t, \omega_1), \ldots, B(t, \omega_M)$, $Y_0(\omega_1), \ldots, Y_0(\omega_M)$ and $Y_1(\omega_1), \ldots, Y_1(\omega_M)$ for M outcomes $\omega_1, \ldots, \omega_M \in \Omega$. Then we solve the M deterministic initial value problems

$$\begin{cases} \ddot{X}(t, \omega_i) + A(t, \omega_i)\dot{X}(t, \omega_i) + B(t, \omega_i)X(t, \omega_i) = 0, & t \in \mathbb{R}, \\ X(t_0, \omega_i) = Y_0(\omega_i), \\ \dot{X}(t_0, \omega_i) = Y_1(\omega_i), \end{cases} \tag{16}$$

so that we obtain M realizations of $X(t)$: $X(t, \omega_1), \ldots, X(t, \omega_M)$. The law of large numbers permits approximating $\mathbb{E}[X(t)]$ and $\mathbb{V}[X(t)]$ by computing the sample mean and sample variance of $X(t, \omega_1), \ldots, X(t, \omega_M)$.

Monte Carlo simulations, in contrast to the dishonest method, give always correct approximations, and as M grows, these approximations are more accurate, although Monte Carlo method possesses a slow convergence rate, namely, $\mathcal{O}(1/\sqrt{M})$ [39, p. 53]. Thereby, the statistical information computed by means of Monte Carlo simulations must approximate our truncation method.

3.6 Obtaining error estimates for the approximation of the solution stochastic process, its mean and its variance

Given an error ϵ, we want to obtain N_ϵ so that $\|X_N(t) - X(t)\|_{L^2(\Omega)} < \epsilon$ for all $N \geq N_\epsilon$. Notice that, in such a case, by Jensen's and the Cauchy–Schwarz inequalities, we would have

$$\left| \mathbb{E}[X_N(t)] - \mathbb{E}[X(t)] \right| = \left| \mathbb{E}[X_N(t) - X(t)] \right| \leq \mathbb{E}[\left| X_N(t) - X(t) \right|] \leq \left\| X_N(t) - X(t) \right\|_{L^2(\Omega)} < \epsilon,$$

therefore we will be able to estimate the error when approximating the mean $\mathbb{E}[X(t)]$ via $\mathbb{E}[X_N(t)]$.

The method to estimate errors for $\|X_N(t) - X(t)\|_{L^2(\Omega)}$ is as follows. We use the notation from the proof of Theorem 3.3. If we denote $\rho = |t - t_0|$ and take $\rho < s < r$, we have

$$\left\| X_N(t) - X(t) \right\|_{L^2(\Omega)} = \left\| \sum_{n=N+1}^{\infty} X_n(t - t_0)^n \right\|_{L^2(\Omega)} \leq \sum_{n=N+1}^{\infty} \|X_n\|_{L^2(\Omega)} \rho^n \leq \sum_{n=N+1}^{\infty} H_n \rho^n. \tag{17}$$

To bound $H_n s^n$, we base our reasoning on (13). Given $M_n = \max_{0 \leq m \leq n} H_m s^m$, we saw in the proof of Theorem 3.3 that $M_n = M$ for sufficiently large n. In fact, for n satisfying

$$\frac{ns}{(n+2)r} + \frac{C_r s}{n+2} + \frac{C_r s^2}{(n+2)(n+1)} < 1, \tag{18}$$

it holds $M_n = M$, so M can be computed just by knowing r, C_r, s, $\|Y_0\|_{L^2(\Omega)}$ and $\|Y_1\|_{L^2(\Omega)}$, because from these values we can see when (18) holds and compute H_n via the recursion (11) or (12), and thereby M. Notice that, if a lot of the random variables A_n and B_n are 0, then H_n will not be a tight bound for $\|X_n\|_{L^2(\Omega)}$, in general.

Once we know M, recall from the end of the proof of Theorem 3.3 that $H_n \rho^n \leq M(\rho/s)^n$. We bound from (17):

$$\left\| X_N(t) - X(t) \right\|_{L^2(\Omega)} \leq M \sum_{n=N+1}^{\infty} \left(\frac{\rho}{s} \right)^n = M \frac{(\frac{\rho}{s})^{N+1}}{1 - \frac{\rho}{s}}.$$

If we want $\|X_N(t) - X(t)\|_{L^2(\Omega)}$ to be smaller than a prefixed error $\epsilon > 0$, we impose

$$M\frac{(\frac{\rho}{s})^{N+1}}{1 - \frac{\rho}{s}} < \epsilon.$$

This yields

$$N_\epsilon = \left\lceil \frac{\log(\frac{(1-\frac{\rho}{s})\epsilon}{M})}{\log(\frac{\rho}{s})} - 1 \right\rceil \tag{19}$$

(here, $\lceil x \rceil$ denotes the least integer that is greater than or equal to x, commonly known as the ceiling of x).

In Example 4.1 and Example 4.4, we will apply these computations to find an N_ϵ for which the approximation of $\mathbb{E}[X(t)]$ via $\mathbb{E}[X_N(t)]$ at the points $t = 1$ and $t = 0.25$, respectively, gives an error smaller than ϵ.

We develop a similar method to estimate the errors in the approximations for the variance. That is to say, given an error $\epsilon > 0$, we want to find N_ϵ such that $|\mathbb{V}[X_N(t)] - \mathbb{V}[X(t)]| < \epsilon$ for all $N \geq N_\epsilon$. We start bounding the difference $|\mathbb{V}[X_N(t)] - \mathbb{V}[X(t)]|$ using triangular, Jensen's and the Cauchy–Schwarz inequalities:

$$
\begin{aligned}
&\left| \mathbb{V}[X_N(t)] - \mathbb{V}[X(t)] \right| \\
&\quad = \left| \mathbb{E}[(X_N(t))^2] - (\mathbb{E}[X_N(t)])^2 - \mathbb{E}[(X(t))^2] + (\mathbb{E}[X(t)])^2 \right| \\
&\quad \leq \mathbb{E}\left[\left| (X_N(t))^2 - (X(t))^2 \right| \right] + \left| (\mathbb{E}[X_N(t)])^2 - (\mathbb{E}[X(t)])^2 \right| \\
&\quad = \mathbb{E}\left[|X_N(t) - X(t)| |X_N(t) + X(t)| \right] + \left| \mathbb{E}[X_N(t)] - \mathbb{E}[X(t)] \right| \left| \mathbb{E}[X_N(t)] + \mathbb{E}[X(t)] \right| \\
&\quad \leq \left\| X_N(t) - X(t) \right\|_{L^2(\Omega)} \left\| X_N(t) + X(t) \right\|_{L^2(\Omega)} \\
&\qquad + \left| \mathbb{E}[X_N(t)] - \mathbb{E}[X(t)] \right| \left(\left| \mathbb{E}[X_N(t)] \right| + \left| \mathbb{E}[X(t)] \right| \right) \\
&\quad \leq \left\| X_N(t) - X(t) \right\|_{L^2(\Omega)} \left(\left\| X_N(t) \right\|_{L^2(\Omega)} + \left\| X(t) \right\|_{L^2(\Omega)} \right) \\
&\qquad + \left| \mathbb{E}[X_N(t)] - \mathbb{E}[X(t)] \right| \left(\left| \mathbb{E}[X_N(t)] \right| + \left| \mathbb{E}[X(t)] \right| \right).
\end{aligned}
$$

Let $\delta > 0$ that will be determined later on to make the error smaller than ϵ. By the results previously obtained in this subsection, we can choose N_δ such that $\|X_N(t) - X(t)\|_{L^2(\Omega)} < \delta$ for all $N \geq N_\delta$ (just applying (19) with $\epsilon = \delta > 0$). Moreover, $|\mathbb{E}[X_N(t)] - \mathbb{E}[X(t)]| \leq \|X_N(t) - X(t)\|_{L^2(\Omega)} < \delta$, by Jensen's and the Cauchy–Schwarz inequalities. Then

$$
\begin{aligned}
&\left| \mathbb{V}[X_N(t)] - \mathbb{V}[X(t)] \right| \\
&\quad \leq \delta \left(\left\| X_N(t) \right\|_{L^2(\Omega)} + \left\| X(t) \right\|_{L^2(\Omega)} \right) + \delta \left(\left| \mathbb{E}[X_N(t)] \right| + \left| \mathbb{E}[X(t)] \right| \right) \\
&\quad \leq \delta \left(\left\| X_N(t) - X(t) + X(t) \right\|_{L^2(\Omega)} + \left\| X(t) \right\|_{L^2(\Omega)} \right) \\
&\qquad + \delta \left(\left| \mathbb{E}[X_N(t)] - \mathbb{E}[X_N(t)] + \mathbb{E}[X_N(t)] \right| + \left| \mathbb{E}[X(t)] \right| \right) \\
&\quad \leq \delta \left(\left\| X_N(t) - X(t) \right\|_{L^2(\Omega)} + \left\| X(t) \right\|_{L^2(\Omega)} + \left\| X(t) \right\|_{L^2(\Omega)} \right) \\
&\qquad + \delta \left(\left| \mathbb{E}[X_N(t)] - \mathbb{E}[X_N(t)] \right| + \left| \mathbb{E}[X(t)] \right| + \left| \mathbb{E}[X(t)] \right| \right)
\end{aligned}
$$

$$\leq \delta\big(\delta + 2\|X(t)\|_{L^2(\Omega)}\big) + \delta\big(\delta + 2\big|\mathbb{E}\big[X(t)\big]\big|\big).$$

To bound $\|X(t)\|_{L^2(\Omega)}$, write

$$\big\|X(t)\big\|_{L^2(\Omega)} = \left\|\sum_{n=0}^{\infty} X_n(t - t_0)^n\right\|_{L^2(\Omega)} \leq \sum_{n=0}^{\infty} \|X_n\|_{L^2(\Omega)}\rho^n$$

$$\leq \sum_{n=0}^{\infty} H_n\rho^n \leq M \sum_{n=0}^{\infty}\left(\frac{\rho}{s}\right)^n = M\frac{1}{1 - \frac{\rho}{s}} =: \gamma,$$

where $\rho = |t - t_0|$ and s is any number satisfying $\rho < s < r$. Before, we saw how to compute M, so we have obtained a computable bound for $\|X(t)\|_{L^2(\Omega)}$. On the other hand, to bound $|\mathbb{E}[X(t)]|$ we have two options. One option consists in using Jensen's and the Cauchy–Schwarz inequalities to derive $|\mathbb{E}[X(t)]| \leq \|X(t)\|_{L^2(\Omega)} \leq \gamma$, and we are done. The second option, which provides a tighter bound for $|\mathbb{E}[X(t)]|$, consists in using the approximations performed for $\mathbb{E}[X(t)]$ via $\mathbb{E}[X_N(t)]$, and from them deducing an upper bound for $\mathbb{E}[X(t)]$. For any of these two options, we denote the upper bound obtained for $|\mathbb{E}[X(t)]|$ by $\beta > 0$. Thus, for $N \geq N_\delta$,

$$\big|\mathbb{V}\big[X_N(t)\big] - \mathbb{V}\big[X(t)\big]\big| \leq \delta(2\gamma + \delta) + \delta(2\beta + \delta).$$

Now choose δ so that

$$\delta(2\gamma + \delta) + \delta(2\beta + \delta) \leq \epsilon.$$

From here, we take

$$\delta = \frac{-(\gamma + \beta) + \sqrt{(\gamma + \beta)^2 + 2\epsilon}}{2} > 0. \tag{20}$$

To sum up, given a prefixed error $\epsilon > 0$, the steps to be done in order to guarantee the approximations $\mathbb{V}[X_N(t)]$ of the exact variance $\mathbb{V}[X(t)]$ to satisfy $|\mathbb{V}[X_N(t)] - \mathbb{V}[X(t)]| \leq \epsilon$ are the following ones:

1. Compute $\gamma = M/(1 - \rho/s)$;
2. Compute $\beta > 0$, upper bound of $|\mathbb{E}[X(t)]|$;
3. Obtain $\delta > 0$ from (20);
4. Take N_δ as in the approximation of the mean (expression (19), but with δ instead of ϵ).

To put forward these ideas, in Example 4.1 and Example 4.4 we will apply these computations to find an N_ϵ for which, given *a priori* error $\epsilon > 0$, the approximation of $\mathbb{V}[X(t)]$ by means of $\mathbb{V}[X_N(t)]$ at the points $t = 1$ and $t = 0.25$, respectively, gives an error smaller than ϵ.

4 Examples

The main goal of this section is to approximate the expectation and variance of the solution process $X(t)$ to particular random initial value problems (1). Our tools will be the ones described in Sect. 3.5. That is, computing $\mathbb{E}[X_N(t)]$ and $\mathbb{V}[X_N(t)]$ for truncation (14), the

dishonest method and Monte Carlo simulations. We will compare the three approaches in order to realize the potentiality of using the truncation method. Our truncation method and Monte Carlo simulation must give similar approximations of the statistical moments (and equal and exact results in the limit).

In addition, we will take some particular problems (1) that have been already studied in the literature such as Airy's and Hermite's random differential equations [3, 7]. As we will see, our findings will generalize the results obtained in those papers (recall Sect. 3.3).

Example 4.1 (Airy's random differential equation) Airy's random differential equation is the following:

$$
\begin{cases}
\ddot{X}(t) + At X(t) = 0, & t \in \mathbb{R}, \\
X(0) = Y_0, \\
\dot{X}(0) = Y_1,
\end{cases}
\tag{21}
$$

where A, Y_0 and Y_1 are random variables.

In [7], the hypothesis used in order to obtain a mean square analytic solution $X(t)$ is $\mathbb{E}[|A|^n] \le H R^n$, $n \ge n_0$. Notice that this hypothesis is equivalent to $\|A\|_{L^\infty(\Omega)} \le R$ by Sect. 3.3.

In the case $\mathbb{E}[|A|^n] \le H R^n$, $n \ge n_0$ (that is, $\|A\|_{L^\infty(\Omega)} \le R$), we are under the hypotheses of Theorem 3.3. Indeed, in the notation of Theorem 3.3, $A_n = 0$ for all $n \ge 0$, $B_1 = A$ and $B_n = 0$ for every $n \ne 1$. For a fixed and finite $r > 0$ and $t_0 = 0$, we have $\|B_1\|_{L^\infty(\Omega)} \le C_r/r^1$, being $C_r = r\|A\|_{L^\infty(\Omega)}$, for instance. Then the stochastic process $X(t) = \sum_{n=0}^\infty X_n t^n$, defined as in Theorem 3.3, is a mean square analytic solution to (21) in $(-r, r)$. As $r > 0$ is arbitrary, in fact $X(t) = \sum_{n=0}^\infty X_n t^n$ is a mean square analytic solution to (21) in $\mathbb{R}$.

Let us carry out a practical case. As in [7], consider $A \sim \text{Beta}(2, 3)$ and Y_0, Y_1 independent random variables such that $Y_0 \sim \text{Normal}(1, 1)$ and $Y_1 \sim \text{Normal}(2, 1)$. Tables 1 and 2 collect the simulations obtained in [7]. In Table 1 we show, for distinct values of t, $\mathbb{E}[X_N(t)]$ for $N = 15$ and $N = 16$, the expectation of the solution stochastic process obtained via the dishonest method and also using Monte Carlo simulations with samples of size 50,000 and 100,000. In Table 2 we present, for distinct values of t, $\mathbb{V}[X_N(t)]$ for $N = 15$ and $N = 16$ and the corresponding approximations computed via Monte Carlo sampling with 50,000 and 100,000 simulations.

We observe that convergence has been achieved for small N. Compare with Monte Carlo simulation, in which a lot of realizations are required in order to obtain good approximations. Nevertheless, it must be remarked that small N is needed for the truncation order, because Airy's random differential equation is not specially complex. For more complex data processes, as in Example 4.3, Example 4.4 and Example 4.5, a larger order of truncation N may be needed. This may imply a computational expense greater than or similar to Monte Carlo simulation.

On the other hand, it is remarkable how well the dishonest method approximates the correct expectation, although the required independence between $A(t), \dot{X}(t)$ and $B(t), X(t)$ does not hold. The key point is that $\mathbb{C}\text{ov}[A(t), \dot{X}(t)]$ and $\mathbb{C}\text{ov}[B(t), X(t)]$ are small, as Table 3 shows, which justifies the accuracy of the dishonest method, especially for small t. Notice that, in this example, $\mathbb{C}\text{ov}[A(t), \dot{X}(t)] = 0$, because $A(t) \equiv 0$ is deterministic. The value of $\mathbb{C}\text{ov}[B(t), X(t)]$ is calculated by considering approximations $X_N(t)$ of $X(t)$ with $N = 16$,

Table 1 Approximation of the expectation of the solution stochastic process. Example 4.1, assuming independent random data

t	$\mathbb{E}[X_{15}(t)]$	$\mathbb{E}[X_{16}(t)]$	Dishonest	MC 50,000	MC 100,000
0.00	1	1	1	0.99701	1.00138
0.25	1.49870	1.49870	1.49870	1.49519	1.49976
0.50	1.98752	1.98752	1.98752	1.98353	1.98829
0.75	2.45108	2.45108	2.45102	2.44667	2.45160
1.00	2.86856	2.86856	2.86818	2.86383	2.86893
1.25	3.21494	3.21494	3.21339	3.21008	3.21534
1.50	3.46310	3.46310	3.45812	3.45831	3.46376
1.75	3.58660	3.58660	3.57340	3.58215	3.58784
2.00	3.56336	3.56335	3.53286	3.55948	3.56552

Table 2 Approximation of the variance of the solution stochastic process. Example 4.1, assuming independent random data

t	$\mathbb{V}[X_{15}(t)]$	$\mathbb{V}[X_{16}(t)]$	MC 50,000	MC 100,000
0.00	1	1	0.99610	0.99530
0.25	1.06035	1.06035	1.05902	1.05642
0.50	1.23142	1.23142	1.23408	1.22793
0.75	1.49261	1.49261	1.50041	1.48944
1.00	1.81392	1.81392	1.82744	1.81127
1.25	2.15870	2.15870	2.17768	2.15721
1.50	2.49379	2.49379	2.51690	2.49462
1.75	2.80560	2.80560	2.83029	2.81030
2.00	3.11530	3.11530	3.13783	3.12559

Table 3 Approximation of $\mathbb{C}\mathrm{ov}[A(t), \dot{X}(t)]$ and $\mathbb{C}\mathrm{ov}[B(t), X(t)]$ via accurate truncations $\dot{X}_{16}(t)$ and $X_{16}(t)$, respectively. Example 4.1, assuming independent random data

t	$\mathbb{C}\mathrm{ov}[A(t), \dot{X}_{16}(t)]$	$\mathbb{C}\mathrm{ov}[B(t), X_{16}(t)]$
0.00	0	0
0.25	0	−0.0000325384
0.50	0	−0.000622983
0.75	0	−0.00365252
1.00	0	−0.0130099
1.25	0	−0.0349332
1.50	0	−0.0778424
1.75	0	−0.151444
2.00	0	−0.264968

since for this order of truncation one gets good approximations of $X(t)$; in other words, approximations can be consider as fairly exact.

We perform another example of Airy's random differential equation (21), but this time the input random variables A, Y_0 and Y_1 will not be independent. Indeed, take a random vector (A, Y_0, Y_1) that follows a multivariate Gaussian distribution, with mean vector and covariance matrix

$$\mu = \begin{pmatrix} 0.4 \\ 1 \\ 2 \end{pmatrix}, \qquad \Sigma = \begin{pmatrix} 0.04 & 0.0001 & -0.05 \\ 0.0001 & 1 & 0.5 \\ -0.005 & 0.5 & 1 \end{pmatrix},$$

respectively. In order for the hypotheses of Theorem 3.3 to be satisfied, we need to truncate A (because the normal distribution is unbounded). Since A follows a normal distribution with mean $\mu_A = 0.4$ and variance $\sigma_A^2 = 0.04$, the interval $[\mu_A - 3\sigma_A, \mu_A + 3\sigma_A] = [-0.2, 1]$

Table 4 Approximation of the expectation of the solution stochastic process. Example 4.1, assuming dependent random data

t	$\mathbb{E}[X_{15}(t)]$	$\mathbb{E}[X_{16}(t)]$	Dishonest	MC 50,000	MC 100,000
0.00	1	1	1	1.00287	1.00114
0.25	1.50619	1.50619	1.49870	1.50202	1.50134
0.50	1.98755	1.98755	1.98750	1.99129	1.99168
0.75	2.45120	2.45120	2.45102	2.45529	2.45678
1.00	2.86893	2.86893	2.86818	2.87321	2.87583
1.25	3.23538	3.23538	3.21339	3.22008	3.22386
1.50	3.46485	3.46485	3.45812	3.46878	3.47374
1.75	3.58966	3.58966	3.57340	3.59290	3.59902
2.00	3.56817	3.56817	3.53286	3.57032	3.57755

Table 5 Approximation of the variance of the solution stochastic process. Example 4.1, assuming dependent random data

t	$\mathbb{V}[X_{15}(t)]$	$\mathbb{V}[X_{16}(t)]$	MC 50,000	MC 100,000
0.00	1	1	1.01144	1.00603
0.25	1.31128	1.31128	1.32213	1.31611
0.50	1.75161	1.75161	1.73855	1.73153
0.75	2.22370	2.22370	2.22680	2.21842
1.00	2.72094	2.72094	2.73668	2.72659
1.25	3.06515	3.06515	3.20829	3.19619
1.50	3.57031	3.57031	3.58843	3.57424
1.75	3.83591	3.83590	3.85507	3.83916
2.00	4.02095	4.02100	4.04117	4.02478

Table 6 Approximation of $\mathbb{C}\mathrm{ov}[A(t),\dot{X}(t)]$ and $\mathbb{C}\mathrm{ov}[B(t),X(t)]$ via accurate truncations $\dot{X}_{16}(t)$ and $X_{16}(t)$, respectively. Example 4.1, assuming dependent random data

t	$\mathbb{C}\mathrm{ov}[A(t),\dot{X}_{16}(t)]$	$\mathbb{C}\mathrm{ov}[B(t),X_{16}(t)]$
0.00	0	0
0.25	0	−0.00106042
0.50	0	−0.00177052
0.75	0	−0.00618508
1.00	0	−0.0173
1.25	0	−0.0509053
1.50	0	−0.0858552
1.75	0	−0.161068
2.00	0	−0.276206

contains 99.7% of the observations of A. Thus, the multivariate Gaussian distribution will be truncated to $[-0.2, 1] \times \mathbb{R} \times \mathbb{R}$.

In Tables 4 and 5, we present the numerical experiments. We use truncation (14) with $N = 15$ and $N = 16$, the dishonest method and Monte Carlo simulation.

Once again, convergence has been achieved quite quickly compared with Monte Carlo simulation. The results obtained are more accurate than via the dishonest method and Monte Carlo simulation. Nonetheless, the accuracy of the dishonest method is remarkable again, particularly in the time interval $t \in [0, 1]$, although not as good as in the previous case (see Table 1). In Table 6, we show approximations of the covariances $\mathbb{C}\mathrm{ov}[A(t), \dot{X}(t)]$ and $\mathbb{C}\mathrm{ov}[B(t), X(t)]$. These covariances are small, especially for small t, which explains the good approximation of the expectation via the dishonest method.

As an application of the error estimates studied in Sect. 3.6, we estimate for which index N_ϵ the error obtained in the approximation of $\mathbb{E}[X(t)]$ via $\mathbb{E}[X_N(t)]$ is smaller than

$\epsilon = 0.00001$. Take $t = 1$ and $r = 2$. Let, for instance, $s = 1.5$. In both cases in this example (assuming independent and dependent random data), we have $C_r = r\|A\|_{L^\infty(\Omega)}$. Since $\|A\|_{L^\infty(\Omega)} = 1$ (recall that in the two cases considered throughout this example, the realizations $A(\omega)$, $\omega \in \Omega$, of the random variable A lie either in $[0,1]$ or in $[-0.2,1]$, thus being less than 1), we take $C_r = r = 2$. From these values, the least n_0 such that (18) holds for all $n \geq n_0$ is $n_0 = 7$ (this value is obtained by plotting the left-hand side of (18) and looking at the point n_0 from which the graph is less than 1). Then, from (13), $M = M_8 = \max_{0 \leq m \leq 8} H_m s^m = 2024.49$. Finally, using (19), one gets $N_\epsilon = 10$. For $N \geq N_\epsilon = 10$, it holds $|\mathbb{E}[X_N(t)] - \mathbb{E}[X(t)]| < 0.00001$.

Now, given $\epsilon = 0.00001$, we obtain an N_ϵ such that $|\mathbb{V}[X_N(t)] - \mathbb{V}[X(t)]| < \epsilon$ at $t = 1$ for every $N \geq N_\epsilon$. We use the ideas and notation from Sect. 3.6. We have $t = 1$, $\rho = 1$, $r = 2$ and $s = 1.5$. We saw that $M = 2024.49$. Then $\gamma = M/(1 - \rho/s) = 6073.47$. Recall that we could choose β equal to γ or, for a tighter bound, use Tables 1 and 4. We see that $|\mathbb{E}[X(t)]| \leq 2.869 =: \beta$. From these values, we obtain $\delta = 8.22638 \cdot 10^{-10}$. Finally, choose N_δ so that $\|X_N(1) - X(1)\|_{L^2(\Omega)} < \delta$. Use formula (19) (with δ instead of ϵ) to get $N_\delta = 73$. Thus, for $N \geq 73$, the inequality $|\mathbb{V}[X_N(t)] - \mathbb{V}[X(t)]| < 0.00001$ holds for sure.

Example 4.2 (Hermite's random differential equation) Hermite's random differential equation is defined as follows:

$$\begin{cases} \ddot{X}(t) - 2t\dot{X}(t) + AX(t) = 0, & t \in \mathbb{R}, \\ X(0) = Y_0, \\ \dot{X}(0) = Y_1, \end{cases} \tag{22}$$

where A, Y_0 and Y_1 are random variables.

In [3], the moments of A are controlled as $\mathbb{E}[|A|^n] \leq HR^n$, $n \geq n_0$, to prove the existence of a mean square analytic solution to random initial value problem (22). As we saw in Sect. 3.3, this hypothesis reduces to $\|A\|_{L^\infty(\Omega)} \leq R$.

If $\mathbb{E}[|A|^n] \leq HR^n$, $n \geq n_0$ (that is, $\|A\|_{L^\infty(\Omega)} \leq R$), we are under the assumptions of Theorem 3.3. Indeed, in the notation of Theorem 3.3, $A_1 = -2$, $A_n = 0$ for all $n \neq 1$, $B_0 = A$ and $B_n = 0$ for every $n \neq 0$. For a fixed and finite $r > 0$ and $t_0 = 0$, we have $\|A_1\|_{L^\infty(\Omega)} \leq C_r/r^1$ and $\|B_0\|_{L^\infty(\Omega)} \leq C_r/r^0 = C_r$, being $C_r = \max\{2r, \|A\|_{L^\infty(\Omega)}\}$ for example. Then $X(t) = \sum_{n=0}^\infty X_n t^n$ defined as in Theorem 3.3 is an analytic solution to (22) in $(-r, r)$. Again, as $r > 0$ is arbitrary, $X(t) = \sum_{n=0}^\infty X_n t^n$ is a mean square analytic solution stochastic process to random initial value problem (22) in $\mathbb{R}$.

As in [3], let $A \sim \text{Normal}(\mu = 5, \sigma^2 = 1)$ and Y_0, Y_1 independent random variables such that $Y_0 \sim \text{Normal}(1, 1)$ and $Y_1 \sim \text{Normal}(2, 1)$. In this case, since the normal distribution is unbounded, it does not fulfill the hypotheses, therefore we need to truncate it: in [3] it has been truncated to the interval $[\mu - 3\sigma, \mu + 3\sigma] = [2, 8]$, which contains approximately 99.7% of the observations of a Gaussian random variable. Tables 7 and 8 simulate the results obtained in [3]. In Table 7 we show, for distinct values of t, $\mathbb{E}[X_N(t)]$ for $N = 15$ and $N = 16$, the corresponding approximation obtained via the dishonest method and Monte Carlo simulations with samples of size 50,000 and 100,000. In Table 8 we present, for distinct values of t, $\mathbb{V}[X_N(t)]$ for $N = 15$ and $N = 16$ and Monte Carlo simulations with samples of size 50,000 and 100,000. As it occurred in the previous example, convergence has been achieved for small N. In Table 9, we show approximations of the covariances

Table 7 Approximation of the expectation of the solution stochastic process. Example 4.2

t	$\mathbb{E}[X_{15}(t)]$	$\mathbb{E}[X_{16}(t)]$	Dishonest	MC 50,000	MC 100,000
0.00	1	1	1	0.99750	0.99919
0.25	1.32907	1.32907	1.32889	1.32619	1.32703
0.50	1.26473	1.26473	1.26175	1.26351	1.26219
0.75	0.74510	0.74510	0.72906	0.74737	0.74316
1.00	−0.27157	−0.27157	−0.32467	−0.26484	−0.27144
1.25	−1.80636	−1.80635	−1.93991	−1.79597	−1.80237
1.50	−3.85882	−3.85868	−4.13681	−3.84872	−3.84906
1.75	−6.40911	−6.40754	−6.90081	−6.40873	−6.39186
2.00	−9.43553	−9.42222	−10.1448	−9.46498	−9.41066

Table 8 Approximation of the variance of the solution stochastic process. Example 4.2

t	$\mathbb{V}[X_{15}(t)]$	$\mathbb{V}[X_{16}(t)]$	MC 50,000	MC 100,000
0.00	1	1	0.98671	1.00327
0.25	0.77433	0.77433	0.76716	0.77821
0.50	0.37752	0.37752	0.37822	0.37992
0.75	0.54181	0.54181	0.53554	0.54357
1.00	2.10396	2.10396	2.06444	2.10993
1.25	5.48674	5.48670	5.40047	5.50378
1.50	10.4476	10.4467	10.3456	10.4828
1.75	18.0186	18.0108	17.9539	18.0963
2.00	43.5731	43.6462	43.5773	44.1340

Table 9 Approximation of $\mathbb{C}\text{ov}[A(t), \dot{X}(t)]$ and $\mathbb{C}\text{ov}[B(t), X(t)]$ via accurate truncations $\dot{X}_{16}(t)$ and $X_{16}(t)$, respectively. Example 4.2

t	$\mathbb{C}\text{ov}[A(t), \dot{X}_{16}(t)]$	$\mathbb{C}\text{ov}[B(t), X_{16}(t)]$
0.00	0	0
0.25	0	−0.0345004
0.50	0	−0.145742
0.75	0	−0.322733
1.00	0	−0.515142
1.25	0	−0.612961
1.50	0	−0.403919
1.75	0	0.522644
2.00	0	3.0078

$\mathbb{C}\text{ov}[A(t), \dot{X}(t)]$ and $\mathbb{C}\text{ov}[B(t), X(t)]$ to understand better the accuracy of the dishonest method.

Example 4.3 (Random linear differential equation with polynomial data processes) Let us consider more complex data processes in our random differential equation (1). The data stochastic processes will be random polynomials. For example,

$$\begin{cases} \ddot{X}(t) + (A_0 + A_1 t)\dot{X}(t) + (B_0 + B_1 t)X(t) = 0, & t \in \mathbb{R}, \\ X(0) = Y_0, \\ \dot{X}(0) = Y_1, \end{cases} \tag{23}$$

where $A_0 = 4$, $A_1 \sim \text{Uniform}(0, 1)$, $B_0 \sim \text{Gamma}(2, 2)$, $B_1 \sim \text{Bernoulli}(0.35)$, $Y_0 = -1$ and $Y_1 \sim \text{Binomial}(2, 0.29)$ are assumed to be independent. In order for the hypotheses of Theorem 3.3 to be satisfied, the gamma distribution will be truncated. For the gamma distri-

Table 10 Approximation of the expectation of the solution stochastic process. Example 4.3

t	$\mathbb{E}[X_{19}(t)]$	$\mathbb{E}[X_{20}(t)]$	Dishonest	MC 50,000	MC 100,000
0.00	−1	−1	−1	−1	−1
0.25	−0.886467	−0.886467	−0.886418	−0.886789	−0.886432
0.50	−0.809269	−0.809269	−0.808743	−0.809370	−0.809219
0.75	−0.747589	−0.747589	−0.745742	−0.747321	−0.747526
1.00	−0.693453	−0.693453	−0.689284	−0.692816	−0.693375
1.25	−0.643943	−0.643944	−0.636462	−0.642985	−0.643845
1.50	−0.598053	−0.598082	−0.586360	−0.596854	−0.597952

Table 11 Approximation of the variance of the solution stochastic process. Example 4.3

t	$\mathbb{V}[X_{15}(t)]$	$\mathbb{V}[X_{16}(t)]$	MC 50,000	MC 100,000
0.00	0	0	0	0
0.25	0.0102077	0.0102074	0.0101172	0.0102664
0.50	0.0190996	0.0190999	0.0189214	0.0192053
0.75	0.0237400	0.0237403	0.0235191	0.0238499
1.00	0.0268721	0.0268711	0.0266311	0.0269620
1.25	0.0297852	0.0297465	0.0295049	0.0298201
1.50	0.0333309	0.0325867	0.0325021	0.0328009

bution with shape and rate 2, it can straightforwardly be checked that the interval $[0, 4]$ contains approximately 99.7% of the observations.

By Theorem 3.3, the mean square solution of (23) can be written as a random power series $X(t) = \sum_{n=0}^{\infty} X_n t^n$ that is mean square convergent for all $t \in \mathbb{R}$.

In Tables 10 and 11, the numerical experiments for the expectation and variance are presented.

Example 4.4 (Random linear differential equation with infinite series data processes) In this example, the data stochastic processes in the random differential equation (1) are non-polynomial analytic stochastic process:

$$\begin{cases} \ddot{X}(t) + A(t)\dot{X}(t) + B(t)X(t) = 0, & t \in \mathbb{R}, \\ X(0) = Y_0, \\ \dot{X}(0) = Y_1, \end{cases} \tag{24}$$

where $A_n \sim \text{Beta}(11, 15)$ for $n \geq 0$, $B_n = 1/n^2$, for $n \geq 1$, and $Y_0 \sim \text{Poisson}(2)$ and $Y_1 \sim \text{Uniform}(0, 1)$ are assumed to be independent. We have $\mathbb{E}[A_n] = 11/26$ and $\mathbb{V}[A_n] = 55/6084$, therefore $\|A_n\|_{\text{L}^2(\Omega)} = \sqrt{\mathbb{V}[A_n] + \mathbb{E}[A_n]^2} = 0.12908$, $n \geq 0$. Then

$$\sum_{n=0}^{\infty} \|A_n\|_{\text{L}^2(\Omega)} t^n = 0.12908 \sum_{n=0}^{\infty} t^n,$$

which is convergent for $t \in (-1, 1)$. On the other hand,

$$\sum_{n=0}^{\infty} \|B_n\|_{\text{L}^2(\Omega)} t^n = \sum_{n=1}^{\infty} \frac{1}{n^2} t^n,$$

which is convergent for $t \in (-1, 1)$ as well. Therefore, the maximum r, that we can take so that the random differential equation (24) makes sense, is $r = 1$.

Table 12 Approximation of the expectation of the solution stochastic process. Example 4.4, assuming independent initial conditions

t	$\mathbb{E}[X_{16}(t)]$	$\mathbb{E}[X_{17}(t)]$	Dishonest	MC 50,000	MC 100,000
0.00	2	2	2	2.01406	1.99858
0.25	2.11267	2.11267	2.11266	2.12676	2.11130
0.50	2.17662	2.17662	2.17654	2.19050	2.17531
0.75	2.15693	2.15693	2.15675	2.17021	2.15573

Table 13 Approximation of the variance of the solution stochastic process. Example 4.4, assuming independent initial conditions

t	$\mathbb{V}[X_{17}(t)]$	$\mathbb{V}[X_{18}(t)]$	MC 50,000	MC 100,000
0.00	2	2	2.00274	2.00822
0.25	1.99421	1.99421	1.99725	2.00356
0.50	1.93408	1.93408	1.93725	1.94407
0.75	1.76917	1.76919	1.77222	1.77899

Since $|A_n(\omega)| \leq 1$ for all $\omega \in \Omega$, $|B_n| \leq 1$ and $r = 1$, we can take $C_r = 1$ in Theorem 3.3 and the hypotheses hold. By Theorem 3.3, the mean square solution of (24), $X(t)$, is defined and is mean square analytic on $(-1, 1)$.

In Tables 12 and 13 we present the numerical experiments. To apply the dishonest method, we need the following two computations:

$$\mathbb{E}[A(t)] = \frac{11}{26} \sum_{n=0}^{\infty} t^n = \frac{11}{26(1-t)}, \quad t \in (-1, 1)$$

and

$$\mathbb{E}[B(t)] = \sum_{n=1}^{\infty} \frac{t^n}{n^2}, \quad t \in (-1, 1).$$

To apply Monte Carlo simulation, we need realizations of the stochastic process $A(t)$, that is, realizations of the random variables $A_0, A_1, \ldots$. As we cannot obtain infinite realizations of a beta distribution in the computer, we will approximate $A(t, \omega) \approx \sum_{n=0}^{100} A_n(t, \omega) t^n$. So, from realizations of $A_0, \ldots, A_{100}$, we will obtain an approximation of a realization of $A(t)$.

By contrast, our approximations using truncation $X_N(t)$, $t \in (-1, 1)$, do not require realizations of the infinite data stochastic process $A(t)$.

As it is observed in Tables 12 and 13, the convergence has been practically achieved for $N = 17$.

As an application of the error estimates analyzed in Sect. 3.6, we estimate for which index N_ϵ the error obtained in the approximation of $\mathbb{E}[X(t)]$ via $\mathbb{E}[X_N(t)]$ is smaller than $\epsilon = 0.00001$. Take $t = 0.25$ and, for instance, $s = 0.5$. We have $r = 1$ and $C_r = 1$, and from these values we obtain the least n_0 such that (18) holds for all $n \geq n_0$ by trial and error. We obtain $n_0 = 0$. From (13), $M = M_1 = \max\{H_0, H_1 s\} = \max\{\|Y_0\|_{L^2(\Omega)}, \|Y_1\|_{L^2(\Omega)} \cdot 0.5\} = \max\{\sqrt{6}, 1/(2\sqrt{3})\} = \sqrt{6}$, whence $N_\epsilon = 18$ by using (19).

Now we bound the error made when approximating the variance. Given an error $\epsilon = 0.00001$, we obtain an N_ϵ such that $|\mathbb{V}[X_N(t)] - \mathbb{V}[X(t)]| < \epsilon$ at the point $t = 0.25$ for all $N \geq N_\epsilon$. We use the ideas and notation from Sect. 3.6. We have $t = 0.25$, $\rho = 0.25$, $r = 1$

Table 14 Approximation of the expectation of the solution stochastic process. Example 4.4, assuming dependent initial conditions

t	$\mathbb{E}[X_{15}(t)]$	$\mathbb{E}[X_{16}(t)]$	Dishonest	MC 50,000	MC 100,000
0.00	0.87	0.87	0.87	0.866780	0.869210
0.25	0.973828	0.973828	0.973819	0.970205	0.973499
0.50	1.04817	1.04817	1.0481	1.04427	1.04823
0.75	1.07358	1.07358	1.0734	1.06961	1.07394

Table 15 Approximation of the variance of the solution stochastic process. Example 4.4, assuming dependent initial conditions

t	$\mathbb{V}[X_{14}(t)]$	$\mathbb{V}[X_{15}(t)]$	MC 50,000	MC 100,000
0.00	0.617700	0.617700	0.616645	0.621350
0.25	0.574373	0.574373	0.574355	0.577586
0.50	0.553754	0.553754	0.554522	0.556844
0.75	0.525683	0.525704	0.526991	0.528727

and $s = 0.5$. We computed $M = \sqrt{6}$, whence $\gamma = M/(1 - \rho/s) = 4.89898$. Recall that we could choose β equal to γ or, for a tighter bound, use Table 12. In Table 12, we see that $|\mathbb{E}[X(0.25)]| \leq 2.113 =: \beta$. From these values, $\delta = 7.13065 \cdot 10^{-7}$. To end up, pick N_δ such that $\|X_N(0.25) - X(0.25)\|_{L^2(\Omega)} < \delta$. Using expression (19) (with δ instead of ϵ), we get $N_\delta = 22$. Thereby, $|\mathbb{V}[X_N(0.25)] - \mathbb{V}[X(0.25)]| < 0.00001$ for $N \geq 22$.

We perform another example for the random initial value problem (24), again with $A_n \sim \text{Beta}(11, 15)$ for $n \geq 0$, $B_n = 1/n^2$ for $n \geq 1$, but now the random vector (Y_0, Y_1) follows a multinomial distribution with three repetitions and probabilities 0.29 and 0.15. The random variables/vectors $A_0, A_1, \ldots$ and (Y_0, Y_1) are independent, but, obviously, Y_0 and Y_1 are not independent. Again, the solution stochastic process $X(t)$ is defined on $(-1, 1)$.

In Tables 14 and 15, we show the numerical experiments. Convergence has been practically achieved.
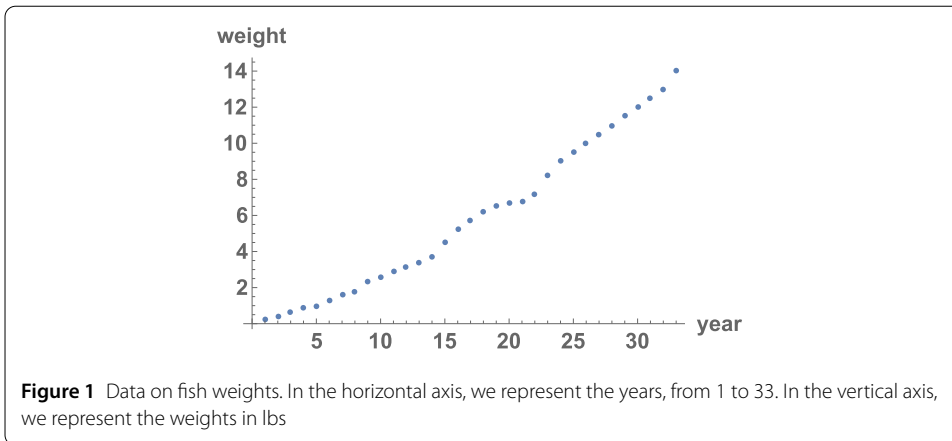
Example 4.5 (An application of the truncation method and Monte Carlo simulation to modeling) In order to see a real application of our theoretical development, let us fit data that describe the fish weight growth over the time via a random second order linear differential equation. In Fig. 1, we show the fish weight in lbs (vertical axis) per year (horizontal axis). The fish weight datum at the ith year will be denoted by w_i for $1 \leq i \leq 33$.

These data were previously used in [5], where a randomized Bernoulli differential equation was used, taking as reference the Bertalanffy model [34, p. 331].

Let W be the stochastic process that models the fish weight. The random variable $W(t)$ models the fish weight at year t, $1 \leq t \leq 33$. Since $W(t)$ is a positive random variable, we will work with $X(t) = \log(W(t))$ instead. The observed data become $\log(w_1), \ldots, \log(w_{33})$. We use the random initial value problem

$$\begin{cases} \ddot{X}(t) + A_0 \dot{X}(t) + (B_0 + B_1 t)X(t) = 0, & t \in \mathbb{R}, \\ X(0) = Y_0, \\ \dot{X}(0) = Y_1 \end{cases} \tag{25}$$

to model the logarithm of the fish weight growth. The stochastic processes $A(t) = A_0$ and $B(t) = B_0 + B_1 t$ have been chosen by numerical fit trials and computational viability. Notice that (25) is a generalization of Airy's random differential equation from Example 4.1.

Figure 1 Data on fish weights. In the horizontal axis, we represent the years, from 1 to 33. In the vertical axis, we represent the weights in lbs

Using the data drawn in Fig. 1, we would like to find the best random variables A_0, B_0, B_1, Y_0 and Y_1 so that $W(t)$ fits appropriately the uncertainty associated to the fish weight growth. Since we do not have an explicit solution process X, we will use truncation (14), $X_N(t)$, to approximate it in the $L^2(\Omega)$ sense. Using the truncation with a high N, we will be able to give a suitable distribution for $(A_0, B_0, B_1, Y_0, Y_1)$.

There are two statistical approaches to dealing with this problem: the frequentist and the Bayesian techniques. Reference [25] provides an introduction to Bayesian statistics. We do not carry out a Bayesian approach, because $X_N(t)$ has a very large expression, which makes the use of Bayesian estimation impracticable in the computer. Thereby, we use the ideas of the so-called inverse frequentist technique for parameter estimation exhibited in [5] and [35, Chap. 7]. In order not to depart from our reasoning, we will explain our concrete frequentist approach in Remark 4.6 at the end of this example.

Without entering into the theoretical details that will be explained in Remark 4.6, we specify the steps to solve our modeling problem. Computational viability makes us choose $N = 24$ as the order of truncation. We give the random vector $(A_0, B_0, B_1, Y_0, Y_1)$ a sixth dimensional multinormal random distribution (in the end, it will be truncated so that the hypotheses of Theorem 3.3 are fulfilled). The mean vector μ is determined as the solution of the deterministic minimization problem

$$\min_{a_0, b_0, b_1, y_0, y_1 \in \mathbb{R}} \sum_{i=1}^{33} \big(\log(w_i) - X_{24}(t_i | a_0, b_0, b_1, y_0, y_1)\big)^2,$$

where $X_{24}(t | a_0, b_0, b_1, y_0, y_1)$ corresponds to the value of $X_{24}(t)$ substituting A_0, B_0, B_1, Y_0 and Y_1 by the real numbers a_0, b_0, b_1, y_0 and y_1. This minimization problem can be solved with the built-in function `FindFit` with the option `Method -> NMinimize` in Mathematica®. We obtain

$$\mu = \begin{pmatrix} 0.169695 \\ -0.0123653 \\ 0.000347771 \\ -2.09309 \\ 0.672599 \end{pmatrix}.$$

Figure 2 Fit of the fish weight data. The blue points represent the real weights, the red points represent the estimated weights (the mean) and the green lines cover a 95% confidence interval constructed with the Gaussian rule [mean $\pm\, 2 \cdot$ standard deviation]
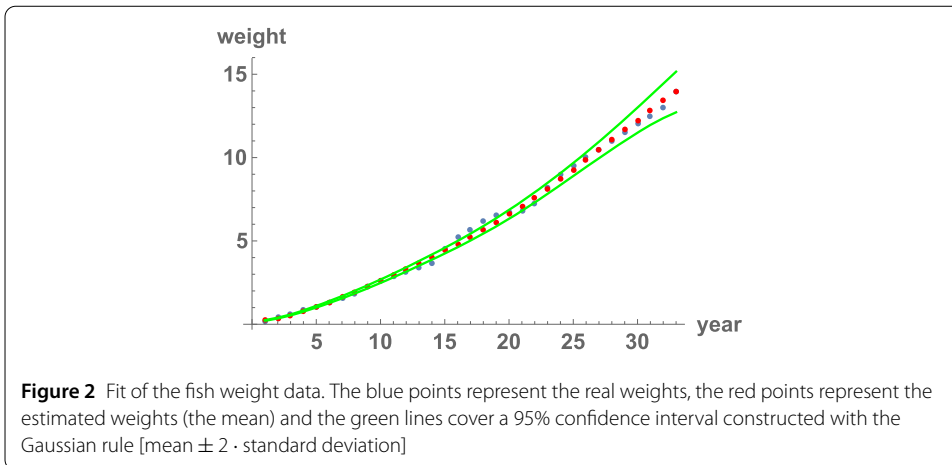
The covariance matrix Σ is estimated with $\sigma^2(J^T J)^{-1}$, where

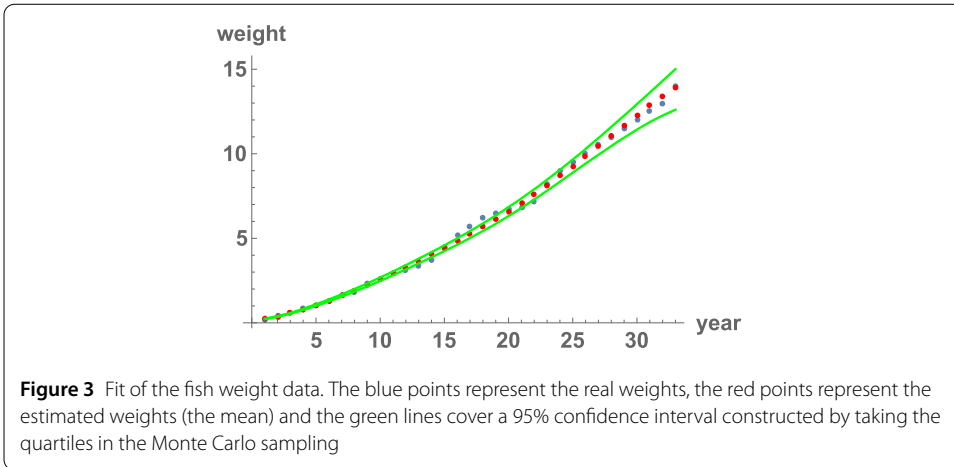$$\sigma^2 = \frac{\sum_{i=1}^{33}(\log(w_i) - X_{24}(t_i|\mu))^2}{33 - 5} = 0.00369358,$$

and J is the Jacobian matrix of $X_{24}(t)$ with respect to $(A_0, B_0, B_1, Y_0, Y_1)$ evaluated at μ. We obtain

$$\Sigma = \begin{pmatrix} 0.000109461 & -0.000010458 & 1.44986 \cdot 10^{-7} & -0.000645456 & 0.000435313 \\ -0.000010458 & 2.47732 \cdot 10^{-6} & -7.94312 \cdot 10^{-8} & 0.0000461867 & -0.0000398354 \\ 1.44986 \cdot 10^{-7} & -7.94312 \cdot 10^{-8} & 3.23082 \cdot 10^{-9} & -3.16049 \cdot 10^{-7} & 5.53045 \cdot 10^{-7} \\ -0.000645456 & 0.0000461867 & -3.16049 \cdot 10^{-7} & 0.00624016 & -0.00312264 \\ 0.000435313 & -0.0000398354 & 5.53045 \cdot 10^{-7} & -0.00312264 & 0.00186506 \end{pmatrix}.$$

Once we know the estimated distribution of the random vector $(A_0, B_0, B_1, Y_0, Y_1)$, we know the distribution of $X_{24}(t)$, at least theoretically. The computational complexity of $X_{24}(t)$ is very big, so computing its exact expectation or even good approximations of it is nearly impossible. Due to the complexity of both the truncation expression and the distribution of $(A_0, B_0, B_1, Y_0, Y_1)$, it is better to perform Monte Carlo simulation directly on (25) via simulations of $(A_0, B_0, B_1, Y_0, Y_1)$, which follows a (truncated) multivariate Gaussian distribution (μ, Σ).

By means of Monte Carlo simulation with 100,000 iterates, we obtain samples of $X(i)$, $i = 1, \ldots, 33$. Applying exponential, we have samples of $W(i)$, $i = 1, \ldots, 33$. Hence, approximations of both $\mathbb{E}[W(i)]$ and $\mathbb{V}[W(i)]$ can be calculated. A confidence interval can be computed in two ways: either considering $[\mathbb{E}[W(t)] \pm 2\sqrt{\mathbb{V}[W(t)]}]$ (this is based on how confidence intervals are constructed in a Gaussian setting) or obtaining an accurate approximation using the quartiles of the sample produced by Monte Carlo. In Figs. 2 and 3, the results are shown. As it is observed in both plots, the mean approximates well the real data. However, the confidence interval grows as we move away from 0. Intuitively, this may hold because a truncated random power series centered at t_0 works better near t_0. This phenomenon may be resolved by making the order of truncation N larger and larger (if the computer permits it).

Remark 4.6 (Frequentist technique used in Example 4.5) Let X be a random vector of size n to be modeled (in our case, $(\log(W(1)), \ldots, \log(W(33)))$). We set a model of the form

Figure 3 Fit of the fish weight data. The blue points represent the real weights, the red points represent the estimated weights (the mean) and the green lines cover a 95% confidence interval constructed by taking the quartiles in the Monte Carlo sampling

$X = f(V)$, where V is a random vector with p components (in our case $(A_0, B_0, B_1, Y_0, Y_1)$) that follows a multinormal distribution with parameters (μ, Σ), and $f : \mathbb{R}^p \to \mathbb{R}^n$ is a function, maybe non-linear (in our case, $n = 33$ and $f_i(a_0, b_0, b_1, y_0, y_1) = X_{24}(i|a_0, b_0, b_1, y_0, y_1)$, $i = 1, \dots, 33$, where $X_{24}(t|a_0, b_0, b_1, y_0, y_1)$ is the truncation that approximates the solution of the random differential equation). Let $x = X(\omega)$ be a vector realization of X (in our example, the real data $x = (\log(w_1), \dots, \log(w_{33}))$). From x, we want to estimate the best μ and Σ so that the model $X = f(V)$ can be considered correct. Let $\hat{v} \in \mathbb{R}^p$ be the minimizer of

$$\min_{v \in \mathbb{R}^p} \sum_{i=1}^{n} (x_i - f_i(v))^2.$$

Using Taylor's expansion,

$$X \approx f(\hat{v}) + Jf(\hat{v})(V - \hat{v}),$$

where J stands for the Jacobian. Then

$$Z := X - f(\hat{v}) + Jf(\hat{v})\hat{v} \approx \underbrace{Jf(\hat{v})}_{J} V.$$

We derive that $Z \approx J\mu + E$, where E follows a multivariate normal distribution with parameters $(0, J\Sigma J^T)$ (here T stands for the transpose matrix operator), i.e., $E \sim \mathrm{MN}(0, J\Sigma J^T)$. Write $J\Sigma J^T = P^T DP$, where P and D are an orthogonal and a diagonal matrix, respectively. Multiplying by P, we have $\bar{Z} \approx \bar{J}\mu + \bar{E}$, where $\bar{Z} = PZ$, $\bar{J} = PJ$ and $\bar{E} = PE \sim \mathrm{MN}(0, D)$. Therefore, $\bar{Z} \approx \bar{J}\mu + \bar{E}$ is a classical linear model (see [32, Chap. 7], [35, Chap. 7]) with normal and independent errors. In a linear model, one should assume homoscedasticity so that the estimation of the parameters makes sense. Thus, we impose $D = \sigma^2 I_n$, where σ^2 is the variance of the errors in the linear model. As a consequence, $\bar{Z} \approx \bar{J}\mu + \bar{E}$ is a classical linear model with homoscedasticity, and the estimations follow from general theory: $\hat{\mu}$ is the minimizer of

$$\min_{\mu \in \mathbb{R}^p} \|\bar{z} - \bar{J}\mu\|_2^2 = \min_{\mu \in \mathbb{R}^p} \|Pz - PJ\mu\|_2^2 = \min_{\mu \in \mathbb{R}^p} \|z - J\mu\|_2^2,$$

where $z = x - f(\hat{v}) + J\hat{v}$ is the vector realization of Z, $\bar{z} = Pz$ is the vector realization of $\bar{Z}$ and $\|\cdot\|_2$ is the Euclidean norm. Now,

$$\|z - J\mu\|_2^2 = \left\|x - \left(f(\hat{v}) + J(\mu - \hat{v})\right)\right\|_2^2 \approx \left\|x - f(\mu)\right\|_2^2,$$

so we can take $\hat{\mu} = \hat{v}$, which justifies our choice for the mean in the example. On the other hand, by the general theory of linear models,

$$\hat{\sigma}^2 = \frac{\|\bar{z} - \bar{J}\hat{\mu}\|_2^2}{n - p} = \frac{\|Pz - PJ\hat{\mu}\|_2^2}{n - p} = \frac{\|z - J\hat{\mu}\|_2^2}{n - p} \approx \frac{\|x - f(\hat{\mu})\|_2^2}{n - p}.$$

Finally, from $J\Sigma J^T = P^T DP = \sigma^2 P^T P = \sigma^2 I_n$, we derive $\Sigma = \sigma^2 (J^T J)^{-1}$, by multiplying by $(J^T J)^{-1} J^T$ to the left and by $J(J^T J)^{-1}$ to the right at both sides of the equality (we assume rank$(J) = p$ so that $(J^T J)^{-1}$ exists). Thus, we choose the estimator $\hat{\Sigma} = \hat{\sigma}^2 (J^T J)^{-1}$.

5 Conclusions

In this paper we have determined analytic stochastic processes that are solutions to the random non-autonomous second order linear differential equation in the mean square sense taking advantage of the powerful theory of random difference equations. After reviewing the $L^p(\Omega)$ random calculus and results concerning random power series (differentiation of a random power series in the $L^p(\Omega)$ sense and Merten's theorem for random series in the mean square sense), we stated the main theorem of the paper, Theorem 3.3. This theorem gives assumptions on the coefficient stochastic processes and on the random initial conditions of a random non-autonomous second order linear differential equation, so that there exists an analytic stochastic process that is a solution in the mean square sense. This mean square approach permitted approximating the main statistical information of the solution stochastic process, expectation and variance. These approximations for the expectation and variance were compared with other methods previously used in the literature: the dishonest method and Monte Carlo simulation.

The numerical examples presented illustrate the potentiality of our results. The examples show that our findings allow for much more complex random non-autonomous second order linear differential equations than those from the existing literature. The ideas of this paper permit dealing with any random non-autonomous second order linear differential equation in a general form. The statistical information of the stochastic process solution can be computed up to any degree of accuracy. These achievements have been reached thanks to the powerful theory of difference equations.

Moreover, our truncation method provides a methodology to estimate the parameters of the multivariate normally distributed explanatory random vector in the modeling of real data via random non-autonomous second order linear differential equations. This procedure together with Monte Carlo simulations gives fitting approximations of the real data.

We mention that a future line of research could consist in extending our ideas to a random non-autonomous linear differential equation of order higher than two using appropriate results belonging to the theory of random difference equations. The extension would require a more complex solution process than that given in Theorem 3.3, which, in turn, would increase the computational expense when approximating accurately its main statistical features.

Availability of data and materials

No applicable.

Competing interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Authors' contributions

The authors read and approved the final manuscript.

Author details

[1]Instituto Universitario de Matemática Multidisciplinar, Universitat Politècnica de València, Valencia, Spain. [2]Department of Mathematics, University of Texas at Austin, Austin, USA.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

1. Apostol, T.M.: Mathematical Analysis, 2nd edn. Pearson, New York (1976)
2. Boyce, W.E.: Probabilistic Methods in Applied Mathematics I. Academic Press, New York (1968)
3. Calbo, G., Cortés, J.C., Jódar, L.: Random Hermite differential equations: mean square power series solutions and statistical properties. Appl. Math. Comput. **218**(7), 3654–3666 (2011)
4. Calbo, G., Cortés, J.C., Jódar, L., Villafuerte, L.: Solving the random Legendre differential equation: mean square power series solution and its statistical functions. Comput. Math. Appl. **61**(9), 2782–2792 (2011)
5. Casabán, M.C., Cortés, J.C., Navarro-Quiles, A., Romero, J.V., Roselló, M.D., Villanueva, R.J.: Computing probabilistic solutions of the Bernoulli random differential equation. J. Comput. Appl. Math. **309**, 396–407 (2017)
6. Casabán, M.C., Cortés, J.C., Romero, J.V., Roselló, M.D.: Solving random homogeneous linear second-order differential equations: a full probabilistic description. Mediterr. J. Math. **13**(6), 3817–3836 (2016)
7. Cortés, J.C., Jódar, L., Camacho, J., Villafuerte, L.: Random Airy type differential equations: mean square exact and numerical solutions. Comput. Math. Appl. **60**(5), 1237–1244 (2010)
8. Cortés, J.C., Jódar, L., Company, R., Villafuerte, L.: Laguerre random polynomials: definition, differential and statistical properties. Util. Math. **98**, 283–295 (2015)
9. Cortés, J.C., Jódar, L., Villafuerte, L.: Random linear-quadratic mathematical models: computing explicit solutions and applications. Math. Comput. Simul. **79**(7), 2076–2090 (2009)
10. Cortés, J.C., Jódar, L., Villafuerte, L.: Mean square solution of Bessel differential equation with uncertainties. J. Comput. Appl. Math. **309**(1), 383–395 (2017)
11. Cortés, J.C., Sevilla-Peris, P., Jódar, L.: Analytic-numerical approximating processes of diffusion equation with data uncertainty. Comput. Math. Appl. **49**(7–8), 1255–1266 (2005)
12. Díaz-Infante, S., Jerez, S.: Convergence and asymptotic stability of the explicit Steklov method for stochastic differential equations. J. Comput. Appl. Math. **291**(1), 36–47 (2016)
13. Dorini, F., Cunha, M.: Statistical moments of the random linear transport equation. J. Comput. Phys. **227**(19), 8541–8550 (2008)
14. Dorini, F.A., Cecconello, M.S., Dorini, M.B.: On the logistic equation subject to uncertainties in the environmental carrying capacity and initial population density. Commun. Nonlinear Sci. Numer. Simul. **33**, 160–173 (2016)
15. Golmankhaneh, A.K., Porghoveh, N.A., Baleanu, D.: Mean square solutions of second-order random differential equations by using homotopy analysis method. Rom. Rep. Phys. **65**(2), 350–362 (2013)
16. Grimmett, G.R., Stirzaker, D.R.: Probability and Random Processes. Clarendon Press, Oxford (2000)
17. Henderson, D., Plaschko, P.: Stochastic Differential Equations in Science and Engineering. Cambridge Texts in Applied Mathematics. World Scientific, Singapore (2006)
18. Hussein, A., Selim, M.M.: A developed solution of the stochastic Milne problem using probabilistic transformations. Appl. Math. Comput. **216**(10), 2910–2919 (2009)
19. Hussein, A., Selim, M.M.: Solution of the stochastic transport equation of neutral particles with anisotropic scattering using RVT technique. Appl. Math. Comput. **213**(1), 250–261 (2009)
20. Hussein, A., Selim, M.M.: Solution of the stochastic radiative transfer equation with Rayleigh scattering using RVT technique. Appl. Math. Comput. **218**(13), 7193–7203 (2012)
21. Khodabin, M., Maleknejad, K., Rostami, M., Nouri, M.: Numerical solution of stochastic differential equations by second order Runge–Kutta methods. Math. Comput. Model. **53**(9–10), 1910–1920 (2011)
22. Khodabin, M., Rostami, M.: Mean square numerical solution of stochastic differential equations by fourth order Runge–Kutta method and its application in the electric circuits with noise. Adv. Differ. Equ. **2015**, 62 (2015)
23. Khudair, A.K., Ameen, A.A., Khalaf, S.L.: Mean square solutions of second-order random differential equations by using Adomian decomposition method. Appl. Math. Sci. **51**(5), 2521–2535 (2011)
24. Khudair, A.K., Haddad, S.A.M., Khalaf, S.L.: Mean square solutions of second-order random differential equations by using the differential transformation method. Open J. Appl. Sci. **6**, 287–297 (2016)
25. Lesaffre, E., Lawson, A.B.: Bayesian Biostatistics. Statistics in Practice. Wiley, New York (2012)

26. Li, X., Fu, X.: Stability analysis of stochastic functional differential equations with infinite delay and its application to recurrent neural networks. J. Comput. Appl. Math. **234**(2), 407–417 (2010)
27. Licea, J.A., Villafuerte, L., Chen-Charpentier, B.M.: Analytic and numerical solutions of a Riccati differential equation with random coefficients. J. Comput. Appl. Math. **309**(1), 208–219 (2013)
28. Liu, S., Debbouche, A., Wang, J.: On the iterative learning control for stochastic impulsive differential equations with randomly varying trial lengths. J. Comput. Appl. Math. **312**, 47–57 (2017)
29. Loève, M.: Probability Theory. Vol. I. Springer, Mineola (1977)
30. Lord, G.J., Powell, C.E., Shardlow, T.: An Introduction to Computational Stochastic PDEs. Cambridge Texts in Applied Mathematics. Dover, New York (2014)
31. Nouri, K., Ranjbar, H.: Mean square convergence of the numerical solution of random differential equations. Mediterr. J. Math. **12**(3), 1123–1140 (2015)
32. Rencher, A.C., Schaalje, G.B.: Linear Models in Statistics, 2nd edn. Wiley, New York (2008)
33. Santos, L.T., Dorini, F.A., Cunha, M.C.C.: The probability density function to the random linear transport equation. Appl. Math. Comput. **216**(5), 1524–1530 (2010)
34. Seber, G.A.F., Wild, C.J.: Nonlinear Regression. Cambridge Texts in Applied Mathematics. Wiley, New York (2003)
35. Smith, R.C.: Uncertainty Quantification: Theory, Implementation, and Applications. SIAM, Philadelphia (2014)
36. Soheili, A.R., Toutounian, F., Soleymani, F.: A fast convergent numerical method for matrix sign function with application in SDEs (Stochastic Differential Equations). J. Comput. Appl. Math. **282**, 167–178 (2015)
37. Soong, T.T.: Random Differential Equations in Science and Engineering. Academic Press, New York (1973)
38. Villafuerte, L., Braumann, C.A., Cortés, J.C., Jódar, L.: Random differential operational calculus: theory and applications. Comput. Math. Appl. **59**(1), 115–125 (2010)
39. Xiu, D.: Numerical Methods for Stochastic Computations. A Spectral Method Approach. Cambridge Texts in Applied Mathematics. Princeton University Press, New York (2010)