Rate of Convergence for Wong–Zakai-Type Approximations of Itô Stochastic Differential Equations

Published Version:
Rate of Convergence for Wong–Zakai-Type Approximations of Itô Stochastic Differential Equations / Ben Ammou B.K.; Lanconelli A.. - In: JOURNAL OF THEORETICAL PROBABILITY. - ISSN 0894-9840. - STAMPA. - 32:4(2019), pp. 1780-1803. [10.1007/s10959-018-0837-x]

Availability:
This version is available at: https://hdl.handle.net/11585/710977 since: 2019-12-30

Published:
DOI: http://doi.org/10.1007/s10959-018-0837-x

Rate of convergence for Wong-Zakai-type approximations of Itô stochastic differential equations

Bilel Kacem Ben Ammou* Alberto Lanconelli†

Abstract

We consider a class of stochastic differential equations driven by a one dimensional Brownian motion and we investigate the rate of convergence for Wong-Zakai-type approximated solutions. We first consider the Stratonovich case, obtained through the point-wise multiplication between the diffusion coefficient and a smoothed version of the noise; then, we consider Itô equations where the diffusion coefficient is Wick-multiplied by the regularized noise. We discover that in both cases the speed of convergence to the exact solution coincides with the speed of convergence of the smoothed noise towards the original Brownian motion. We also prove, in analogy with a well known property for exact solutions, that the solutions of approximated Itô equations solve approximated Stratonovich equations with a certain correction term in the drift.

Key words and phrases: stochastic differential equations, Wong-Zakai theorem, Wick product

AMS 2000 classification: 60H10; 60H30; 60H05

1 Introduction and statement of the main results

From a modeling point of view, the celebrated Wong-Zakai theorem [24],[25] provides a crucial insight in the theory of stochastic differential equations. It asserts that the solution $\{X_t^{(n)}\}_{t\in[0,T]}$ of the random ordinary differential equation

$$\frac{dX_t^{(n)}}{dt} = b(t, X_t^{(n)}) + \sigma(t, X_t^{(n)}) \cdot \frac{dB_t^{(n)}}{dt}, \tag{1.1}$$

where $\{B_t^{(n)}\}_{t\in[0,T]}$ is a suitable smooth approximation of the Brownian motion $\{B_t\}_{t\in[0,T]}$, converges in the mean, as n goes to infinity, to the solution of the Stratonovich stochastic differential equation (SDE, for short)

$$dX_t = b(t, X_t)dt + \sigma(t, X_t) \circ dB_t. \tag{1.2}$$

*Department of Mathematics, Universtiy of Tunis - El Manar, Street Mohamed Alaya Kacem, Nabeul - Tunisia. E-mail: *bilelbenammou@gmail.com*

†Dipartimento di Matematica, Universitá degli Studi di Bari Aldo Moro, Via E. Orabona 4, 70125 Bari - Italia. E-mail: *alberto.lanconelli@uniba.it*

At a first sight, it may look a bit surprising the fact that the sequence $\{X_t^{(n)}\}_{t\in[0,T]}$ does not converge to the Itô's interpretation of the corresponding stochastic equation, i.e.

$$dX_t = b(t, X_t)dt + \sigma(t, X_t)dB_t. \tag{1.3}$$

What makes the sequence $\{X_t^{(n)}\}_{t\in[0,T]}$ prefer to converge to the Stratonovich SDE (1.2) instead of the Itô SDE (1.3) is the presence of the point-wise product $\cdot$ appearing in (1.1) between the diffusion coefficient σ and the smoothed noise. In fact, Hu and Øksendal [14] proved, when the diffusion coefficient is linear, that the solution of

$$\frac{dY_t^{(n)}}{dt} = b(t, Y_t^{(n)}) + \sigma(t)Y_t^{(n)} \diamond \frac{dB_t^{(n)}}{dt}, \tag{1.4}$$

where $\diamond$ stands for the Wick product, converges as n goes to infinity to the solution of the Itô SDE

$$dY_t = b(t, Y_t)dt + \sigma(t)Y_t dB_t. \tag{1.5}$$

Along this direction, Da Pelo et al. [5] introduced a family of products interpolating between the point-wise and Wick products and proved convergence for Wong-Zakai-type approximations toward stochastic differential equations where the stochastic integrals are defined via suitable evaluation points in the Riemann sums.
Approximation procedures based on Wong-Zakai-type theorems have attracted the attention of several authors. First of all, Stroock and Varadhan [22] proved the multidimensional version of the Wong-Zakai theorem. Then, generalizations to SDEs driven by different type of noises and to stochastic partial differential equations have been the most investigated directions. For instance, Konecny [17] proved a Wong-Zakai-type theorem for one-dimensional SDEs driven by a semimartingale, Gyoñgy and G. Michaletzky [6] considered δ-martingales while Naganuma [20] examined the case of Gaussian rough paths. In the theory of stochastic partial differential equations, Hairer and Pardoux [9] proved a version of the Wong-Zakai theorem for one-dimensional parabolic nonlinear stochastic PDEs driven by space-time white noise utilizing the recent theory of regularity structures; Brezniak and Flandoli [2] proved almost sure convergence to the solution to a Stratonovich stochastic partial differential equation; Tessitore and Zabczyk [23] obtained results on the weak convergence of the laws of the Wong-Zakai approximations for stochastic evolution equations. We also mention that Londono and Villegas [19] proposed to use a Wong-Zakai type approximation method for the numerical evaluation of the solutions of SDEs.
The aim of the present paper is to compare the rate of convergence for approximations of Stratonovich and Itô quasi-linear SDEs and to investigate whether the connection between exact solutions of the two different interpretations can be restored for the corresponding approximating sequences (see the discussion after Corollary 1.6 below).

We remark that the rate of convergence for Wong-Zakai approximations, in the Stratonovich case, has been already investigated by other authors. We recall Hu and Nualart [13] dealing with almost sure convergence in Hölder norms; Hu, Kallianpur and Xiong [12] studying approximations for the Zakai equation and Gyongy and Shmatkov [7] and Gyongy and Stinga [8] treating general linear stochastic partial differential equations. We also refer the reader to the book by Hu [11] where Wong-Zakai approximations are considered in the framework of Euler-Maruyama discretization schemes.We will discuss in Remark 1.4 below the details of the comparison between our convergence rate for Stratonovich equations and the one in [11].

While Wong-Zakai-type theorems for Stratonovich SDEs have been largely investigated, approximations for Itô SDEs are very rare in the literature. In fact, as the paper by Hu and Øksendal shows, to recover the Itô interpretation of the SDE one has to deal with the Wick product and in most cases this multiplication is not easy to handle. This is the reason why we focus on equations with linear diffusion coefficient (it is in fact not known whether the fully non linear version of (1.4) admits a solution [14]). Nevertheless, to find the speed of convergence of the approximation to the solution of the Itô equation, we had to utilize some tools from the Malliavin calculus (see Lemma 3.5 below). The present paper can be considered as a continuation of the work presented in Da Pelo et al. [5], where the issue of the rate of convergence has not been studied.
To state our main results we briefly describe our framework. Let $(W, \mathcal{A}, \mu)$ be the classical Wiener space over the time interval $[0, T]$, where T is an arbitrary positive constant, and denote by $\{B_t\}_{t\in[0,T]}$ the coordinate process, i.e.

$$\begin{aligned} B_t : W &\to \mathbb{R} \\ \omega &\mapsto B_t(\omega) = \omega(t). \end{aligned}$$

By construction, the process $\{B_t\}_{t\in[0,T]}$ is, under the measure μ, a one dimensional Brownian motion. We now introduce a smooth (continuously differentiable) approximation of $\{B_t\}_{t\in[0,T]}$ by means of a kernel satisfying certain technical assumptions. In the sequel, the symbol $|f|$ will denote the norm of $f \in L^2([0, T])$ while $\|X\|_p$ will denote the norm of $X \in \mathcal{L}^p(W, \mu)$ for any $p \geq 1$.

Assumption 1.1 *For any $\varepsilon > 0$ let $K_\varepsilon : [0, T]^2 \to \mathbb{R}$ be such that*

- *the function $t \mapsto K_\varepsilon(t, s)$ belongs to $C^1([0, T])$ for almost all $s \in [0, T]$;*
- *the functions $s \mapsto K_\varepsilon(t, s)$ and $s \mapsto \partial_t K_\varepsilon(t, s)$ belong to $L^2([0, T])$ for all $t \in [0, T]$.*

Moreover, we assume that

$$\lim_{\varepsilon\to 0^+} \sup_{t\in[0,T]} |K_\varepsilon(t, \cdot) - 1_{[0,t]}(\cdot)| = 0 \tag{1.6}$$

and

$$M := \sup_{\varepsilon>0} \sup_{t\in[0,T]} |K_\varepsilon(t, \cdot)| < +\infty.$$

Now, if we let

$$B_t^\varepsilon := \int_0^T K_\varepsilon(t, s) dB_s, \quad t \in [0, T],$$

and recall that $B_t = \int_0^T 1_{[0,t]}(s) dB_s$, then Assumption 1.1 implies that $\{B_t^\varepsilon\}_{t\in[0,T]}$ is a continuosly differentiable Gaussian process and that B_t^ε converges to B_t in $\mathcal{L}^2(W, \mu)$ uniformly with respect to $t \in [0, T]$. In fact, condition (1.6) is equivalent to

$$\lim_{\varepsilon\to 0^+} \sup_{t\in[0,T]} \|B_t^\varepsilon - B_t\|_2 = 0.$$

Therefore, we deal with a quite general class of smooth approximations of the Brownian motion $\{B_t\}_{t\geq 0}$. In the sequel we will be studying SDEs of the type (1.5) both in the Stratonovich and Itô senses. We now state the assumptions on the coefficients b and σ which are supposed to be valid for the rest of the present paper.

Assumption 1.2 *There exist two positive constants C_1 and C_2 such that for all $t \in [0,T]$ and $x, y \in \mathbb{R}$ one has*

$$|b(t,x) - b(t,y)| \leq C_1|x-y| \quad \textit{and} \quad |b(t,x)| \leq C_2(1+|x|). \tag{1.7}$$

Moreover, the function σ belongs to $\mathcal{L}^\infty([0,T])$.

For $f \in L^2([0,T])$ we denote

$$\mathcal{E}(f) := \exp\left\{\int_0^T f(s)dB_s - \frac{1}{2}\int_0^T f^2(s)ds\right\}$$

and we call it *stochastic exponential*. The set $\{\mathcal{E}(f), f \in L^2([0,T])\}$ turns out to be total in $\mathcal{L}^p(W,\mu)$ for any $p \geq 1$. Given $f, g \in L^2([0,T])$, the *Wick product* of $\mathcal{E}(f)$ and $\mathcal{E}(g)$ is defined to be

$$\mathcal{E}(f) \diamond \mathcal{E}(g) := \mathcal{E}(f+g).$$

This multiplication can be extended by linearity and density to an unbounded bilinear form on a proper subset of $\mathcal{L}^p(W,\mu) \times \mathcal{L}^p(W,\mu)$ (see Holden et al. [10] and Janson [15] for its connection to Itô-Skorohod integration theory). For $g \in L^2([0,T])$ we also define the *translation operator* T_g as the operator that shifts the Brownian path by the function $\int_0^{\cdot} g(s)ds$; more precisely, the action of T_g on stochastic exponentials is given by

$$T_g\mathcal{E}(f) := \mathcal{E}(f) \cdot \exp\{\langle f, g\rangle\}.$$

where $\langle \cdot, \cdot \rangle$ denotes the inner product in $L^2([0,T])$ (see Holden et al. [10] and Janson [15] for details).
We are now ready to state the first two main theorems of the present paper. The proofs are postponed to Section 2 and Section 3, respectively.

Theorem 1.3 *Let $\{X_t\}_{t\in[0,T]}$ be the unique solution of the Stratonovich SDE*

$$dX_t = b(t,X_t)dt + \sigma(t)X_t \circ dB_t, \quad t \in]0,T] \qquad X_0 = x \tag{1.8}$$

and for any $\varepsilon > 0$ let $\{X_t^\varepsilon\}_{t\in[0,T]}$ be the unique solution of

$$\frac{dX_t^\varepsilon}{dt} = b(t,X_t^\varepsilon) + \sigma(t)X_t^\varepsilon \cdot \frac{dB_t^\varepsilon}{dt}, \quad X_0^\varepsilon = x. \tag{1.9}$$

Then, for any $p \geq 1$ there exists a positive constant C (depending on p, $|x|$, T, C_1, C_2 and M) such that for any q greater than p

$$\sup_{t\in[0,T]} \|X_t^\varepsilon - X_t\|_p \leq C \cdot \mathcal{S}_q\left(\sup_{t\in[0,T]} |K_\varepsilon(t,\cdot) - 1_{[0,t]}(\cdot)|\right) \tag{1.10}$$

where

$$\mathcal{S}_q(\lambda) := \lambda \exp\left\{q\lambda^2\right\} + \exp\{\lambda^2/2\} - 1, \quad \lambda \in \mathbb{R} \tag{1.11}$$

Remark 1.4 *In Theorem 11.6 of [11] it is proved that*

$$\left\| \sup_{t\in[0,T]} |X_t^\pi - X_t| \right\|_p \leq C_{p,T}(\log|\pi|)^2|\pi|^{\frac{1}{2}+\frac{1}{\log|\pi|}} \tag{1.12}$$

where π is a partition of the interval $[0,T]$, $|\pi|$ denotes the mesh of the partition π and $\{X_t^\pi\}_{t\in[0,T]}$ stands for the solution of

$$\frac{dX_t^\pi}{dt} = b(t, X_t^\pi) + \sigma(t, X_t^\pi) \cdot \frac{dB_t^\pi}{dt},$$

with $\{B_t^\pi\}_{t\in[0,T]}$ being the polygonal approximation of $\{B_t\}_{t\in[0,T]}$ associated to the partition π. The above result is stated and proved for general nonlinear systems of SDEs driven by a multidimensional Brownian motion. Moreover, the topology utilized in (1.12) is stronger than the one in (1.10) (where the supremum is outside of the $\mathcal{L}^p(W,\mu)$-norm). It is not difficult to see that the polygonal approximation $\{B_t^\pi\}_{t\in[0,T]}$ is included in the family of approximations $\{B_t^\varepsilon\}_{t\in[0,T]}$ considered in the present paper (the parameter ε reduces to the mesh of the partition $|\pi|$) and in that case we get

$$\sup_{t\in[0,T]} \|B_t^\pi - B_t\|_2 = \sup_{t\in[0,T]} |K_{|\pi|}(t,\cdot) - 1_{[0,t]}(\cdot)| = C\sqrt{|\pi|}.$$

Substituting in (1.10) we obtain

$$\sup_{t\in[0,T]} \|X_t^\pi - X_t\|_p \leq C \cdot \mathcal{S}_q(C\sqrt{|\pi|})$$

which behaves like $\sqrt{|\pi|}$ for $|\pi|$ going to zero. A comparison with (1.12) shows that Theorem 1.3 provides a highest rate of convergence, at the price of a weaker topology and more restrictive conditions on the class of the SDEs considered.
The result and proof of Theorem 1.3 are however necessary for the comparison proposed in the present paper.

Theorem 1.5 *Let $\{Y_t\}_{t\in[0,T]}$ be the unique solution of the Itô SDE*

$$dY_t = b(t, Y_t)dt + \sigma(t)Y_t dB_t, \quad t \in]0,T] \qquad Y_0 = x \tag{1.13}$$

and for any $\varepsilon > 0$ let $\{Y_t^\varepsilon\}_{t\in[0,T]}$ be the unique solution of

$$\frac{dY_t^\varepsilon}{dt} = b(t, Y_t^\varepsilon) + \sigma(t)Y_t^\varepsilon \diamond \frac{dB_t^\varepsilon}{dt}, \quad Y_0^\varepsilon = x. \tag{1.14}$$

Then, for any $p \geq 1$ there exists a positive constant C (depending on p, $|x|$, T, C_1, C_2 and M) such that for any q greater than p

$$\sup_{t\in[0,T]} \|Y_t^\varepsilon - Y_t\|_p \leq C \cdot \mathcal{S}_q\left(\sqrt{2}\sup_{t\in[0,T]} |K_\varepsilon(t,\cdot) - 1_{[0,t]}(\cdot)|\right)$$

where $\mathcal{S}$ is the function defined in (1.11).

Corollary 1.6 *In the notation of Theorem 1.3 and Theorem 1.5, we have for any* $p \geq 1$ *that*

$$\lim_{\varepsilon \to 0^+} \sup_{t\in[0,T]} \|X_t^\varepsilon - X_t\|_p = \lim_{\varepsilon \to 0^+} \sup_{t\in[0,T]} \|Y_t^\varepsilon - Y_t\|_p = 0$$

where both limits have rate of convergence of order

$$\sup_{t\in[0,T]} |K_\varepsilon(t,\cdot) - 1_{[0,t]}(\cdot)| \quad \text{as } \varepsilon \text{ tends to zero.}$$

Proof. It follows from

$$\lim_{\lambda \to 0^+} \frac{\mathcal{S}_q(\lambda)}{\lambda} = 1. \tag{1.15}$$

for all $q \geq 1$. ∎

It is well known (see for instance Karatzas and Shreve [16]) that the Itô SDE (1.13) can be reformulated as the Stratonovich SDE

$$dY_t = \left(b(t, Y_t) - \frac{1}{2}\sigma(t) Y_t \right) dt + \sigma(t) Y_t \circ dB_t, \quad t \in]0,T] \qquad Y_0 = x. \tag{1.16}$$

The next theorem provides a similar representation for the approximated Itô equation (1.14) in terms of a suitable approximated Stratonovich equation. The proof can be found in Section 4.

Theorem 1.7 *For any* $\varepsilon > 0$ *let* $\{Y_t^\varepsilon\}_{t\in[0,T]}$ *be the unique solution of*

$$\frac{dY_t^\varepsilon}{dt} = b(t, Y_t^\varepsilon) + \sigma(t) Y_t^\varepsilon \diamond \frac{dB_t^\varepsilon}{dt}, \quad X_0^\varepsilon = x.$$

Then, for any $t \in [0,T]$ *we have*

$$Y_t^\varepsilon = T_{-K_\varepsilon(t,\cdot)} S_t^\varepsilon, \tag{1.17}$$

where $\{S_t^\varepsilon\}_{t\in[0,T]}$ *is the unique solution of*

$$\frac{dS_t^\varepsilon}{dt} = b(t, S_t^\varepsilon) + \frac{1}{2}\frac{d|K_\varepsilon(t,\cdot)|^2}{dt} \cdot S_t^\varepsilon + \sigma(t) S_t^\varepsilon \cdot \frac{dB_t^\varepsilon}{dt}, \quad S_0^\varepsilon = x.$$

The paper is organized as follows: Section 2 and Section 3 are devoted to the proofs of Theorem 1.3 and Theorem 1.5, respectively. Both sections also contain some preliminary results and estimates utilized in the proofs of the main results, which are divided in two major steps. Section 4 contains two different proofs of Theorem 1.7, the second one being a direct verification of the identity (1.17).

2 Proof of Theorem 1.3

2.1 Auxiliary results and remarks: Stratonovich case

The proof of Theorem 1.3 will be carried for the simplified equation where b does not depend on t and σ is identically equal to one. Straightforward modifications will lead to the general case.
The existence and uniqueness for the solutions of (1.9) and (1.8) follow, in view of Assumption 1.1 and Assumption 1.2, by standard results in the theory of stochastic and ordinary differential equations. We also refer the reader to Theorem 5.5 in [5] for a proof using the techniques adopted in this paper. To ease the notation we define

$$E_\varepsilon(t) := \exp\{\delta(-K_\varepsilon(t,\cdot))\} \quad \text{and} \quad E_0(t) := \exp\{\delta(-1_{[0,t]}(\cdot))\}.$$

Here and in the sequel the symbol $\delta(f)$ stands for $\int_0^T f(s)dB_s$. We begin by observing that (see the proof of Theorem 5.5 in [5]) the solution $\{X_t^\varepsilon\}_{t\in[0,T]}$ from Theorem 1.3 can be represented as

$$X_t^\varepsilon = Z_t^\varepsilon \cdot E_\varepsilon(t)^{-1}$$

where

$$\frac{dZ_t^\varepsilon}{dt} = b(Z_t^\varepsilon \cdot E_\varepsilon^{-1}(t)) \cdot E_\varepsilon(t), \quad Z_0^\varepsilon = x.$$

The same holds true for $\{X_t\}_{t\in[0,T]}$; more precisely,

$$X_t = Z_t \cdot E_0(t)^{-1}$$

where

$$\frac{dZ_t}{dt} = b(Z_t \cdot E_0^{-1}(t)) \cdot E_0(t), \quad Z_0 = x.$$

Moreover, we have the estimates

$$\begin{aligned}
|Z_t^\varepsilon| &\leq |x| + \int_0^t |b(Z_s^\varepsilon \cdot E_\varepsilon(s)^{-1}) \cdot E_\varepsilon(s)|ds \\
&\leq |x| + \int_0^t C_2\left(1 + |Z_s^\varepsilon \cdot E_\varepsilon(s)^{-1}|\right) \cdot E_\varepsilon(s)ds \\
&= |x| + \int_0^t C_2 E_\varepsilon(s)ds + \int_0^t C_2|Z_s^\varepsilon|ds \\
&\leq |x| + \int_0^T C_2 E_\varepsilon(s)ds + \int_0^t C_2|Z_s^\varepsilon|ds.
\end{aligned}$$

By the Gronwall inequality,

$$|Z_t^\varepsilon| \leq \left(|x| + \int_0^T C_2 E_\varepsilon(s)ds\right)e^{C_2 t}.$$

This shows that for any $q \geq 1$ we have the bound

$$\left\|\sup_{t\in[0,T]}|Z_t^\varepsilon|\right\|_q \leq \left(|x| + \int_0^T C_2\|E_\varepsilon(s)\|_q ds\right)e^{C_2 T}$$

$$
\begin{aligned}
&= \left(|x| + \int_0^T C_2 \exp\left\{\frac{q}{2}|K_\varepsilon(s,\cdot)|^2\right\} ds\right) e^{C_2 T} \\
&\leq \left(|x| + C_2 T \exp\left\{\frac{q}{2} \sup_{s\in[0,T]} |K_\varepsilon(s,\cdot)|^2\right\}\right) e^{C_2 T}. \qquad (2.1)
\end{aligned}
$$

To prove Theorem 1.3 we need the following estimate which is of independent interest.

Proposition 2.1 *Let $f, g \in L^2([0,T])$. Then, for any $p \geq 1$ we have*

$$
\| \exp\{\delta(f)\} - \exp\{\delta(g)\}\|_p \leq C\mathcal{S}_p(|f-g|)
$$

where

$$
\mathcal{S}_p(\lambda) := \lambda \exp\left\{p\lambda^2\right\} + \exp\{\lambda^2/2\} - 1, \quad \lambda \in \mathbb{R}
$$

and C is a constant depending on p and $|g|$.

Proof. The proof involves few notions of Malliavin calculus. We refer the reader to the books of Nualart [21] and Bogachev [1]. Let $f \in L^2([0,T])$ and $p \geq 1$; then, according to the Poincaré inequality (see Theorem 5.5.11 in Bogachev [1]), we can write

$$
\begin{aligned}
\| \exp\{\delta(f)\} - 1\|_p &\leq \| \exp\{\delta(f)\} - E[\exp\{\delta(f)\}]\|_p + |E[\exp\{\delta(f)\}] - 1| \\
&= \| \exp\{\delta(f)\} - \exp\{|f|^2/2\}\|_p + \exp\{|f|^2/2\} - 1 \\
&\leq \mathcal{C}(p) \left\| |D\exp\{\delta(f)\}|_{L^2([0,T])}\right\|_p + \exp\{|f|^2/2\} - 1 \\
&= \mathcal{C}(p) \left\| |\exp\{\delta(f)\} f|_{L^2([0,T])}\right\|_p + \exp\{|f|^2/2\} - 1 \\
&= \mathcal{C}(p)|f| \| \exp\{\delta(f)\}\|_p + \exp\{|f|^2/2\} - 1 \\
&= \mathcal{C}(p)|f| \exp\left\{\frac{p}{2}|f|^2\right\} + \exp\{|f|^2/2\} - 1
\end{aligned}
$$

where D denotes the Malliavin derivative and $\mathcal{C}(p)$ is a positive constant depending only on p. Therefore, for any $f, g \in L^2([0,T])$ and $p \geq 1$ we have

$$
\begin{aligned}
\| \exp\{\delta(f)\} - \exp\{\delta(g)\}\|_p &= \| \exp\{\delta(g)\} \left(\exp\{\delta(f-g)\} - 1\right) \|_p \\
&\leq \| \exp\{\delta(g)\}\|_{2p} \| \exp\{\delta(f-g)\} - 1\|_{2p} \\
&\leq e^{p|g|^2} \left(\mathcal{C}(2p)|f-g| \exp\left\{p|f-g|^2\right\} + \exp\{|f-g|^2/2\} - 1\right) \\
&\leq C\mathcal{S}_p(|f-g|)
\end{aligned}
$$

where we utilized the Hölder inequality. ∎

2.2 Proof of Theorem 1.3

The proof is divided in two steps.

Step one: We prove that for any $p \geq 1$ there exists a positive constant C (depending on p, $|x|$, T, C_1, C_2 and M) such that for any q greater than p

$$
\left\| \sup_{t\in[0,T]} |Z_t^\varepsilon - Z_t| \right\|_p \leq C \cdot \mathcal{S}_q\left(\sup_{s\in[0,T]} |K_\varepsilon(s,\cdot) - 1_{[0,s]}(\cdot)|\right) \qquad (2.2)
$$

We begin by using the equations solved by Z^ε_t and Z_t and the assumptions on b to get

$$\begin{aligned}
|Z^\varepsilon_t - Z_t| &= \Big|\int_0^t b(Z^\varepsilon_s E_\varepsilon(s)^{-1})E_\varepsilon(s)ds - \int_0^t b(Z_s E_0(s)^{-1})E_0(s)ds\Big| \\
&\le \Big|\int_0^t b(Z^\varepsilon_s E_\varepsilon(s)^{-1})E_\varepsilon(s) - b(Z_s E_0(s)^{-1})E_\varepsilon(s)ds\Big| \\
&\quad +\Big|\int_0^t b(Z_s E_0(s)^{-1})E_\varepsilon(s) - b(Z_s E_0(s)^{-1})E_0(s)ds\Big| \\
&\le \int_0^t |b(Z^\varepsilon_s E_\varepsilon(s)^{-1}) - b(Z_s E_0(s)^{-1})|E_\varepsilon(s)ds \\
&\quad + \int_0^t |b(Z_s E_0(s)^{-1})||E_\varepsilon(s) - E_0(s)|ds \\
&\le \int_0^t C_1|Z^\varepsilon_s E_\varepsilon(s)^{-1} - Z_s E_0(s)^{-1}|E_\varepsilon(s)ds \\
&\quad + \int_0^t C_2(1+|Z_s E_0(s)^{-1}|)|E_\varepsilon(s) - E_0(s)|ds \\
&\le C_1\int_0^t |Z^\varepsilon_s E_\varepsilon(s)^{-1} - Z_s E_\varepsilon(s)^{-1}|E_\varepsilon(s) + |Z_s E_\varepsilon(s)^{-1} - Z_s E_0(s)^{-1}|E_\varepsilon(s)ds \\
&\quad +C_2\int_0^t (1+|Z_s|E_0(s)^{-1})|E_\varepsilon(s) - E_0(s)|ds \\
&= C_1\int_0^t |Z^\varepsilon_s - Z_s|ds + C_1\int_0^t |Z_s||E_\varepsilon(s)^{-1} - E_0(s)^{-1}|E_\varepsilon(s)ds \\
&\quad +C_2\int_0^t (1+|Z_s|E_0(s)^{-1})|E_\varepsilon(s) - E_0(s)|ds \\
&\le C_1\int_0^t |Z^\varepsilon_s - Z_s|ds + C_1\int_0^T |Z_s||E_\varepsilon(s)^{-1} - E_0(s)^{-1}|E_\varepsilon(s)ds \\
&\quad +C_2\int_0^T (1+|Z_s|E_0(s)^{-1})|E_\varepsilon(s) - E_0(s)|ds \\
&= \Lambda_\varepsilon + C_1\int_0^t |Z^\varepsilon_s - Z_s|ds
\end{aligned}$$

where

$$\begin{aligned}
\Lambda_\varepsilon &:= C_1\int_0^T |Z_s||E_\varepsilon(s)^{-1} - E_0(s)^{-1}|E_\varepsilon(s)ds \\
&\quad +C_2\int_0^T (1+|Z_s|E_0(s)^{-1})|E_\varepsilon(s) - E_0(s)|ds.
\end{aligned}$$

By the Gronwall inequality we deduce that

$$|Z^\varepsilon_t - Z_t| \le \Lambda_\varepsilon e^{C_1 t}, \quad t\in[0,T]$$

and hence for $p\ge 1$ the inequality

$$\Big\|\sup_{t\in[0,T]}|Z^\varepsilon_t - Z_t|\Big\|_p \le e^{C_1 T}\|\Lambda_\varepsilon\|_p.$$

We now estimate $\|\Lambda_\varepsilon\|_p$ by writing $\Lambda_\varepsilon = \Lambda_\varepsilon^1 + \Lambda_\varepsilon^2$ where

$$\Lambda_\varepsilon^1 := C_1 \int_0^T |Z_s||E_\varepsilon(s)^{-1} - E_0(s)^{-1}|E_\varepsilon(s)ds$$

and

$$\Lambda_\varepsilon^2 := C_2 \int_0^T (1 + |Z_s|E_0(s)^{-1})|E_\varepsilon(s) - E_0(s)|ds.$$

Applying the triangle and Hölder inequalities we get

$$\begin{aligned}
\|\Lambda_\varepsilon^1\|_p &\leq C_1 \int_0^T \||Z_s||E_\varepsilon(s)^{-1} - E_0(s)^{-1}|E_\varepsilon(s)\|_p ds \\
&\leq C_1 \int_0^T \|Z_s\|_{p_1}\|E_\varepsilon(s)^{-1} - E_0(s)^{-1}\|_{p_2}\|E_\varepsilon(s)\|_{p_3} ds
\end{aligned}$$

where $p_1, p_2, p_3 \in [1, +\infty[$ satisfy $\frac{1}{p_1} + \frac{1}{p_2} + \frac{1}{p_3} = \frac{1}{p}$. From the estimate (2.1) and the identity

$$\|E_\varepsilon(s)\|_{p_3} = \exp\left\{\frac{p_3}{2}|K_\varepsilon(s,\cdot)|^2\right\}$$

we can write

$$\|\Lambda_\varepsilon^1\|_p \leq C \int_0^T \|E_\varepsilon(s)^{-1} - E_0(s)^{-1}\|_{p_2} ds$$

where C denotes a positive constant depending on C_1, C_2, $|x|$, T, p and M (in the sequel C will denote a generic constant, depending on the previously specified parameters, which may vary from one line to another). Moreover, employing Proposition 2.1 with $f(\cdot) = K_\varepsilon(s,\cdot)$ and $g(\cdot) = 1_{[0,s]}(\cdot)$ we conclude that

$$\begin{aligned}
\|\Lambda_\varepsilon^1\|_p &\leq C \int_0^T \mathcal{S}_{p_2}(|K_\varepsilon(s,\cdot) - 1_{[0,s]}(\cdot)|)ds \\
&\leq C \cdot \mathcal{S}_{p_2}\left(\sup_{s\in[0,T]} |K_\varepsilon(s,\cdot) - 1_{[0,s]}(\cdot)|\right). \qquad (2.3)
\end{aligned}$$

Note that for any $p \geq 1$ the function $\lambda \mapsto \mathcal{S}(\lambda)$ is increasing on $[0, +\infty]$. Let us now consider Λ_ε^2; if we apply one more time the triangle and Hölder inequalities, then we get

$$\begin{aligned}
\|\Lambda_\varepsilon^2\|_p &\leq C_2 \int_0^T \|(1 + |Z_s|E_0(s)^{-1})|E_\varepsilon(s) - E_0(s)|\|_p ds \\
&\leq C_2 \int_0^T \|1 + |Z_s|E_0(s)^{-1}\|_{q_1}\|E_\varepsilon(s) - E_0(s)\|_{q_2} ds \\
&\leq C_2 \int_0^T (1 + \||Z_s|E_0(s)^{-1}\|_{q_1})\|E_\varepsilon(s) - E_0(s)\|_{q_2} ds
\end{aligned}$$

where $q_1, q_2 \in [1, +\infty[$ satisfy $\frac{1}{q_1} + \frac{1}{q_2} = \frac{1}{p}$. We observe that

$$1 + \||Z_s|E_0(s)^{-1}\|_{q_1} \leq 1 + \|Z_s\|_{r_1} \cdot \|E_0(s)^{-1}\|_{r_2} \qquad (2.4)$$

where $\frac{1}{q_1} = \frac{1}{r_1} + \frac{1}{r_2}$ and that, according to estimate (2.1), the right hand side of (2.4) is bounded uniformly in $s \in [0, T]$ by a constant C depending on C_1, C_2, $|x|$, T, p and M. Therefore,

$$\|\Lambda_\varepsilon^2\|_p \leq C \cdot \mathcal{S}_{q_2} \left(\sup_{s \in [0,T]} |K_\varepsilon(s, \cdot) - 1_{[0,s]}(\cdot)| \right). \tag{2.5}$$

Here we utilized Proposition 2.1 with $f(\cdot) = -K_\varepsilon(s, \cdot)$ and $g(\cdot) = -1_{[0,s]}(\cdot)$. Finally, combining (2.3) with (2.5) we obtain

$$\left\| \sup_{t \in [0,T]} |Z_t^\varepsilon - Z_t| \right\|_p \leq C \cdot \mathcal{S}_p \left(\sup_{s \in [0,T]} |K_\varepsilon(s, \cdot) - 1_{[0,s]}(\cdot)|^2 \right).$$

Step two: We prove that for any $p \geq 1$ there exists a positive constant C (depending on p, $|x|$, T, C_1, C_2 and M) such that for any q greater than p

$$\sup_{t \in [0,T]} \|X_t^\varepsilon - X_t\|_p \leq C \cdot \mathcal{S}_q \left(\sup_{s \in [0,T]} |K_\varepsilon(s, \cdot) - 1_{[0,s]}(\cdot)| \right).$$

We first note that

$$\begin{aligned} X_t^\varepsilon - X_t &= Z_t^\varepsilon \cdot E_\varepsilon(t)^{-1} - Z_t \cdot E_0(t)^{-1} \\ &= Z_t^\varepsilon \cdot E_\varepsilon(t)^{-1} - Z_t^\varepsilon \cdot E_0(t)^{-1} + Z_t^\varepsilon \cdot E_0(t)^{-1} - Z_t \cdot E_0(t)^{-1} \\ &= Z_t^\varepsilon \cdot (E_\varepsilon(t)^{-1} - E_0(t)^{-1}) + (Z_t^\varepsilon - Z_t) \cdot E_0(t)^{-1}. \end{aligned}$$

Now we take $p \geq 1$ and apply the triangle and Hölder inequalities to get

$$\|X_t^\varepsilon - X_t\|_p \leq \|Z_t^\varepsilon\|_{p_1} \cdot \|E_\varepsilon(t)^{-1} - E_0(t)^{-1}\|_{p_2} + \|Z_t^\varepsilon - Z_t\|_{q_1} \cdot \|E_0(t)^{-1}\|_{q_2}$$

where $\frac{1}{p} = \frac{1}{p_1} + \frac{1}{p_2} = \frac{1}{q_1} + \frac{1}{q_2}$. From estimate (2.1) we know that $\|Z_t^\varepsilon\|_{p_1}$ is bounded uniformly in $t \in [0, T]$ for any $p_1 \geq 1$ while Proposition 2.1 ensures that

$$\|E_\varepsilon(t)^{-1} - E_0(t)^{-1}\|_{p_2} \leq C\mathcal{S}_{p_2} \left(|K_\varepsilon(t, \cdot) - 1_{[0,t]}(\cdot)| \right)$$

with a constant independent of $t \in [0, T]$. Moreover, inequality (2.2) from *Step one* gives for $r > q_1$

$$\|Z_t^\varepsilon - Z_t\|_{q_1} \leq C \cdot \mathcal{S}_r \left(\sup_{s \in [0,T]} |K_\varepsilon(s, \cdot) - 1_{[0,s]}(\cdot)| \right).$$

These last assertions imply

$$\sup_{t \in [0,T]} \|X_t^\varepsilon - X_t\|_p \leq C \cdot \mathcal{S}_q \left(\sup_{s \in [0,T]} |K_\varepsilon(s, \cdot) - 1_{[0,s]}(\cdot)| \right).$$

The proof is complete.

3 Proof of Theorem 1.5

3.1 Auxiliary results and remarks: Itô case

The proof of Theorem 1.5 will be carried for the simplified equation where b does not depend on t and σ is identically equal to one. Straightforward modifications will lead to the general case. To ease the notation, we denote for $t \in [0,T]$

$$\mathcal{E}_\varepsilon(t) := \mathcal{E}(-K_\varepsilon(t,\cdot)) = \exp\left\{\delta(-K_\varepsilon(t,\cdot)) - \frac{1}{2}|K_\varepsilon(t,\cdot)|^2\right\}$$

and

$$\mathcal{E}_0(t) := \mathcal{E}(-1_{[0,t]}) = \exp\left\{\delta(-1_{[0,t]}(\cdot)) - \frac{t}{2}\right\}.$$

The existence and uniqueness for the solutions of (1.14) and (1.13) can be found in Theorem 5.5 from [5]. There it was observed that the solution $\{Y_t^\varepsilon\}_{t\in[0,T]}$ of (1.14) can be represented as

$$Y_t^\varepsilon = V_t^\varepsilon \diamond \mathcal{E}_\varepsilon(t)^{\diamond -1}$$

where

$$\frac{dV_t^\varepsilon}{dt} = b\left(V_t^\varepsilon \diamond (\mathcal{E}_\varepsilon(t))^{\diamond -1}\right) \diamond \mathcal{E}_\varepsilon(t), \quad V_0^\varepsilon = x \tag{3.1}$$

while the solution $\{Y_t\}_{t\in[0,T]}$ of (1.13) can be represented as

$$Y_t = V_t \diamond \mathcal{E}_0(t)^{\diamond -1}$$

where

$$\frac{dV_t}{dt} = b\left(V_t \diamond (\mathcal{E}_0(t))^{\diamond -1}\right) \diamond \mathcal{E}_0(t), \quad V_0 = x. \tag{3.2}$$

Here, for $f \in L^2([0,T])$ the symbol $\mathcal{E}(f)^{\diamond -1}$ stands for the so called *Wick inverse* of $\mathcal{E}(f)$ which, in this particular case, coincides with $\mathcal{E}(-f)$. The next lemma will serve to write equations (3.1) and (3.2) in a Wick product-free form.

Lemma 3.1 *If $F \in \mathcal{L}^p(W,\mu)$ for some $p > 1$ and $\Psi : \mathbb{R} \to \mathbb{R}$ is measurable and with at most linear growth at infinity, then for all $h \in L^2([0,T])$ we have:*

$$\Psi\left(F \diamond \mathcal{E}(h)\right) \diamond \mathcal{E}(-h) = \Psi\left(F \cdot \mathcal{E}(-h)^{-1}\right) \cdot \mathcal{E}(-h).$$

Proof. We apply twice Gjessing's lemma (see Holden et al. [10]) to get:

$$\begin{aligned}
\Psi\left(F \diamond \mathcal{E}(h)\right) \diamond \mathcal{E}(-h) &= \Psi\left(T_{-h}F \cdot \mathcal{E}(h)\right) \diamond \mathcal{E}(-h) \\
&= T_h\left(\Psi(T_{-h}F \cdot \mathcal{E}(h))\right) \cdot \mathcal{E}(-h) \\
&= \Psi\left(F \cdot T_h\mathcal{E}(h)\right) \cdot \mathcal{E}(-h) \\
&= \Psi\left(F \cdot \mathcal{E}(-h)^{-1}\right) \cdot \mathcal{E}(-h).
\end{aligned}$$

Here we utilized the identities

$$\begin{aligned} T_h\mathcal{E}(h) &= \mathcal{E}(h)\exp\left\{\int_0^T h(s)^2 ds\right\} \\ &= \exp\left\{\int_0^T h(s)dB_s + \frac{1}{2}\int_0^T h(s)^2 ds\right\} \\ &= \mathcal{E}(-h)^{-1}. \end{aligned}$$

The proof is complete. ∎

Therefore, by Lemma 3.1 we can rewrite equation (3.1) as

$$\frac{dV_t^\varepsilon}{dt} = b\left(V_t^\varepsilon \cdot (\mathcal{E}_\varepsilon(t))^{-1}\right) \cdot \mathcal{E}_\varepsilon(t)$$

and equation (3.2) as

$$\frac{dV_t}{dt} = b\left(V_t \cdot (\mathcal{E}_0(t))^{-1}\right) \cdot \mathcal{E}_0(t)$$

since, as we mentioned before,

$$\mathcal{E}_\varepsilon(t)^{\diamond-1} = \mathcal{E}(K_\varepsilon(t,\cdot)) \quad \text{and} \quad \mathcal{E}_0(t)^{\diamond-1} = \mathcal{E}(1_{[0,t]}(\cdot)).$$

The following two propositions are the stochastic exponential's counterparts of Proposition 2.1.

Proposition 3.2 *Let $f, g \in L^2([0,T])$. Then, for any $p \geq 1$ we have*

$$\|\mathcal{E}(f) - \mathcal{E}(g)\|_p \leq C \cdot \mathcal{S}_p(|f-g|)$$

where, as before,

$$\mathcal{S}_p(\lambda) = \lambda \exp\left\{p\lambda^2\right\} + \exp\{\lambda^2/2\} - 1, \quad \lambda \in \mathbb{R}$$

and C is a constant depending on p and $|g|$.

Proof. Let $f \in L^2([0,T])$ and $p \geq 1$; then, according to the Poincaré inequality (see Theorem 5.5.11 in Bogachev [1]), we can write

$$\begin{aligned} \|\mathcal{E}(f) - 1\|_p &\leq \mathcal{C}(p)\left\||D\mathcal{E}(f)|_{L^2([0,T])}\right\|_p \\ &= \mathcal{C}(p)\left\||\mathcal{E}(f)f|_{L^2([0,T])}\right\|_p \\ &= \mathcal{C}(p)|f|\|\mathcal{E}(f)\|_p \\ &= \mathcal{C}(p)|f|\exp\left\{\frac{p-1}{2}|f|^2\right\} \end{aligned}$$

where D denotes the Malliavin derivative and $\mathcal{C}(p)$ is a positive constant depending only on p. Therefore, for any $f, g \in L^2([0,T])$ and $p \geq 1$ we have

$$\begin{aligned} \|\mathcal{E}(f) - \mathcal{E}(g)\|_p &= \|\mathcal{E}(g) \diamond (\mathcal{E}(f-g) - 1)\|_p \\ &\leq \|\mathcal{E}(\sqrt{2}g)\|_p \|\mathcal{E}(\sqrt{2}(f-g)) - 1\|_p \\ &\leq e^{(p-1)|g|^2}\mathcal{C}(p)|f-g|\exp\left\{(p-1)|f-g|^2\right\} \\ &\leq C\mathcal{S}_p(|f-g|) \end{aligned}$$

where we utilized an inequality for the Wick product from Da Pelo et al. [4]. ∎

Proposition 3.3 *Let $f, g \in L^2([0,T])$. Then, for any $p \geq 1$ we have*

$$\|\mathcal{E}(f)^{-1} - \mathcal{E}(g)^{-1}\|_p \leq C \cdot \mathcal{S}_p(\sqrt{2}|f-g|)$$

where C is a constant depending on p and $|g|$.

Proof. Denote by $\Gamma(1/\sqrt{2})$ the bounded linear operator acting on stochastic exponentials according to the prescription

$$\Gamma(1/\sqrt{2})\mathcal{E}(f) := \mathcal{E}(f/\sqrt{2}).$$

This operator coincides with the Ornstein-Uhlenbeck semigroup $\{P_t\}_{t\geq 0}$ for a proper choice of the parameter t (see Janson [15] for details) and therefore it is a contraction on any $\mathcal{L}^p(W,\mu)$ for $p \geq 1$. Moreover, by a direct verification one can see that

$$\mathcal{E}(f)^{-1} = \Gamma(1/\sqrt{2}) \exp\{-\delta(\sqrt{2}f)\}. \tag{3.3}$$

Hence, we can write

$$\begin{aligned} \|\mathcal{E}(f)^{-1} - \mathcal{E}(g)^{-1}\|_p &= \|\Gamma(1/\sqrt{2}) \exp\{-\delta(\sqrt{2}f)\} - \Gamma(1/\sqrt{2}) \exp\{-\delta(\sqrt{2}g)\}\|_p \\ &\leq \|\exp\{-\delta(\sqrt{2}f)\} - \exp\{-\delta(\sqrt{2}g)\}\|_p. \end{aligned}$$

Therefore, by means of Proposition 2.1 we can conclude that

$$\begin{aligned} \|\mathcal{E}(f)^{-1} - \mathcal{E}(g)^{-1}\|_p &\leq \|\exp\{-\delta(\sqrt{2}f)\} - \exp\{-\delta(\sqrt{2}g)\}\|_p \\ &\leq C \cdot \mathcal{S}_p(\sqrt{2}|f-g|). \end{aligned}$$

∎

Remark 3.4 *The idea of the proof of the previous proposition, and in particular identity (3.3), is inspired by the investigation carried in Da Pelo and Lanconelli [3], where a new probabilistic representation for the solution of the heat equation is derived in terms of the operator $\Gamma(1/\sqrt{2})$ and its inverse.*

3.2 Proof of Theorem 1.5

As before, we divide the proof in two steps.

Step one: We prove that for any $p \geq 1$ there exists a positive constant C (depending on p, $|x|$, T, C_1, C_2 and M) such that for any q greater than p

$$\left\| \sup_{t\in[0,T]} |V_t^{\varepsilon} - V_t| \right\|_p \leq C \cdot \mathcal{S}_q \left(\sqrt{2} \sup_{s\in[0,T]} |K_{\varepsilon}(s,\cdot) - 1_{[0,s]}(\cdot)| \right). \tag{3.4}$$

The proof can be carried following the same line of the proof of *Step one* of Theorem 1.3; we have simply to replace $\{Z_t\}_{t\in[0,T]}$ and $\{Z_t^{\varepsilon}\}_{t\in[0,T]}$ with $\{V_t\}_{t\in[0,T]}$ and $\{V_t^{\varepsilon}\}_{t\in[0,T]}$, respectively. Moreover, the exponentials $\{E_{\varepsilon}(t)\}_{t\in[0,T]}$ and $\{E_0(t)\}_{t\in[0,T]}$ have to be replaced by $\{\mathcal{E}_{\varepsilon}(t)\}_{t\in[0,T]}$ and $\{\mathcal{E}_0(t)\}_{t\in[0,T]}$, respectively. The estimate (2.1) changes to

$$\left\| \sup_{t\in[0,T]} |V_t^{\varepsilon}| \right\|_q \leq \left(|x| + C_2 T \exp\left\{ \frac{q-1}{2} \sup_{s\in[0,T]} |K_{\varepsilon}(s,\cdot)|^2 \right\} \right) e^{C_2 T}.$$

We remark that for all $r \geq 1$ we have

$$\|E_\varepsilon(t)\|_r = \exp\left\{\frac{r}{2}|K_\varepsilon(s,\cdot)|^2\right\}$$

while

$$\|\mathcal{E}_\varepsilon(t)\|_r = \exp\left\{\frac{r-1}{2}|K_\varepsilon(s,\cdot)|^2\right\} \quad \text{and} \quad \|\mathcal{E}_\varepsilon(t)^{-1}\|_r = \exp\left\{\frac{r+1}{2}|K_\varepsilon(s,\cdot)|^2\right\}.$$

Moreover, we utilize Proposition 3.2 and Proposition 3.3 with $f(\cdot) = K_\varepsilon(s,\cdot)$ and $g(\cdot) = 1_{[0,s]}(\cdot)$ instead of Proposition 2.1.

Step two: We prove that for any $p \geq 1$ there exists a positive constant C (depending on p, $|x|$, T, C_1, C_2 and M) such that for any q greater than p

$$\sup_{t\in[0,T]} \|Y_t^\varepsilon - Y_t\|_p \leq C \cdot \mathcal{S}_q\left(\sqrt{2}\sup_{s\in[0,T]}|K_\varepsilon(s,\cdot) - 1_{[0,s]}(\cdot)|\right).$$

We first note that

$$Y_t^\varepsilon - Y_t = V_t^\varepsilon \diamond \mathcal{E}_\varepsilon(t)^{\diamond-1} - V_t \diamond \mathcal{E}_0(t)^{\diamond-1}.$$

To ease the readability of the formulas we will adopt, only for this part of the proof, the notation

$$\tilde{\mathcal{E}}_\varepsilon(t) := \mathcal{E}_\varepsilon(t)^{\diamond-1} \quad \text{and} \quad \tilde{\mathcal{E}}_0(t) := \mathcal{E}_0(t)^{\diamond-1}.$$

Then, by mean of Gjessing's Lemma we have

$$\begin{aligned}
Y_t^\varepsilon - Y_t &= V_t^\varepsilon \diamond \tilde{\mathcal{E}}_\varepsilon(t) - V_t \diamond \tilde{\mathcal{E}}_0(t) \\
&= V_t^\varepsilon \diamond \tilde{\mathcal{E}}_\varepsilon(t) - V_t^\varepsilon \diamond \tilde{\mathcal{E}}_0(t) + V_t^\varepsilon \diamond \tilde{\mathcal{E}}_0(t) - V_t \diamond \tilde{\mathcal{E}}_0(t) \\
&= T_{-K_\varepsilon(t,\cdot)}V_t^\varepsilon \cdot \tilde{\mathcal{E}}_\varepsilon(t) - T_{-1_{[0,t]}(\cdot)}V_t^\varepsilon \cdot \tilde{\mathcal{E}}_0(t) + (V_t^\varepsilon - V_t) \diamond \tilde{\mathcal{E}}_0(t) \\
&= T_{-K_\varepsilon(t,\cdot)}V_t^\varepsilon \cdot \tilde{\mathcal{E}}_\varepsilon(t) - T_{-K_\varepsilon(t,\cdot)}V_t^\varepsilon \cdot \tilde{\mathcal{E}}_0(t) + T_{-K_\varepsilon(t,\cdot)}V_t^\varepsilon \cdot \tilde{\mathcal{E}}_0(t) \\
&\quad -T_{-1_{[0,t]}(\cdot)}V_t^\varepsilon \cdot \tilde{\mathcal{E}}_0(t) + (V_t^\varepsilon - V_t) \diamond \tilde{\mathcal{E}}_0(t) \\
&= T_{-K_\varepsilon(t,\cdot)}V_t^\varepsilon \cdot \left(\tilde{\mathcal{E}}_\varepsilon(t) - \tilde{\mathcal{E}}_0(t)\right) + \left(T_{-K_\varepsilon(t,\cdot)}V_t^\varepsilon - T_{-1_{[0,t]}(\cdot)}V_t^\varepsilon\right) \cdot \tilde{\mathcal{E}}_0(t) \\
&\quad +T_{-1_{[0,t]}(\cdot)}\left(V_t^\varepsilon - V_t\right) \cdot \tilde{\mathcal{E}}_0(t) \\
&= \mathcal{F}_1 + \mathcal{F}_2 + \mathcal{F}_3
\end{aligned}$$

where we set

$$\mathcal{F}_1 := T_{-K_\varepsilon(t,\cdot)}V_t^\varepsilon \cdot \left(\tilde{\mathcal{E}}_\varepsilon(t) - \tilde{\mathcal{E}}_0(t)\right) \qquad \mathcal{F}_2 := \left(T_{-K_\varepsilon(t,\cdot)}V_t^\varepsilon - T_{-1_{[0,t]}(\cdot)}V_t^\varepsilon\right) \cdot \tilde{\mathcal{E}}_0(t)$$

and

$$\mathcal{F}_3 := T_{-1_{[0,t]}(\cdot)}\left(V_t^\varepsilon - V_t\right) \cdot \tilde{\mathcal{E}}_0(t).$$

Hence, for any $p \geq 1$ we can write

$$\|Y_t^\varepsilon - Y_t\|_p \leq \|\mathcal{F}_1\|_p + \|\mathcal{F}_2\|_p + \|\mathcal{F}_3\|_p.$$

We recall (see Theorem 14.1 in Janson [15]) that for any $g \in L^2([0,T])$ the linear operator T_g is bounded from $\mathcal{L}^q(W,\mu)$ to $\mathcal{L}^p(W,\mu)$ for any $p < q$. Therefore, by the Hölder inequality and Proposition 3.2 we deduce

$$\begin{aligned}
\|\mathcal{F}_1\|_p &= \left\| T_{-K_\varepsilon(t,\cdot)} V_t^\varepsilon \cdot \left(\tilde{\mathcal{E}}_\varepsilon(t) - \tilde{\mathcal{E}}_0(t) \right) \right\|_p \\
&\leq \left\| T_{-K_\varepsilon(t,\cdot)} V_t^\varepsilon \right\|_{q_1} \cdot \left\| \tilde{\mathcal{E}}_\varepsilon(t) - \tilde{\mathcal{E}}_0(t) \right\|_{q_2} \\
&\leq C \left\| V_t^\varepsilon \right\|_r \cdot \left\| \tilde{\mathcal{E}}_\varepsilon(t) - \tilde{\mathcal{E}}_0(t) \right\|_{q_2} \\
&\leq C \cdot \mathcal{S}_{q_2} \left(\sqrt{2} \sup_{s \in [0,T]} |K_\varepsilon(s,\cdot) - 1_{[0,s]}(\cdot)| \right).
\end{aligned}$$

where $p < q_1 < r$, C is a constant depending on the parameters appearing in the statement of the theorem and $\frac{1}{p} = \frac{1}{q_1} + \frac{1}{q_2}$. The term $\|\mathcal{F}_3\|_p$ is treated similarly with the help of inequality (3.4). Let us now focus on $\|\mathcal{F}_2\|_p$. We first observe that

$$\begin{aligned}
\|\mathcal{F}_2\|_p &= \left\| \left(T_{-K_\varepsilon(t,\cdot)} V_t^\varepsilon - T_{-1_{[0,t]}(\cdot)} V_t^\varepsilon \right) \cdot \tilde{\mathcal{E}}_0(t) \right\|_p \\
&\leq \left\| T_{-K_\varepsilon(t,\cdot)} V_t^\varepsilon - T_{-1_{[0,t]}(\cdot)} V_t^\varepsilon \right\|_q \cdot \left\| \tilde{\mathcal{E}}_0(t) \right\|_r.
\end{aligned}$$

According to Theorem 14.1 in Janson [15] the map $T_g X$ is jointly continuous in the variables (g, X) from $L^2([0,T]) \times \mathcal{L}^q(W,\mu)$ to $\mathcal{L}^p(W,\mu)$ for $p < q$. Therefore, the first term in the last member of the previous inequality tends to zero as $\varepsilon \to 0^+$. However, we need to know the speed of such convergence. The following lemma will help us in this direction.

Lemma 3.5 *For any $X \in \mathbb{D}^{1,q}$ and $h \in L^2([0,T])$ with $|h| < \delta$ one has*

$$\|T_h X - X\|_p \leq C|h| \|X\|_{\mathbb{D}^{1,q}}$$

where $p < q$ and C depends on δ, p and q.

Proof. Since the linear span of the stochastic exponentials is dense in $\mathcal{L}^p(W,\mu)$ and in $\mathbb{D}^{1,q}$, we will prove the lemma for $X = \sum_{j=1}^n \alpha_j \mathcal{E}(f_j)$ where $\alpha_1, ..., \alpha_n \in \mathbb{R}$ and $f_1, ..., f_n \in L^2([0,T])$. By the mean value theorem we can write for $\theta \in [0,1]$ that

$$\begin{aligned}
T_h \sum_{j=1}^n \alpha_j \mathcal{E}(f_j) - \sum_{j=1}^n \alpha_j \mathcal{E}(f_j) &= \sum_{j=1}^n \alpha_j \mathcal{E}(f_j) \left(e^{\langle h, f_j \rangle} - 1 \right) \\
&= \sum_{j=1}^n \alpha_j \mathcal{E}(f_j) e^{\theta \langle h, f_j \rangle} \langle h, f_j \rangle \\
&= T_{\theta h} D_h \sum_{j=1}^n \alpha_j \mathcal{E}(f_j)
\end{aligned}$$

where $D_{\theta h}$ stands for the Malliavin derivative in the direction θh. We now take the $\mathcal{L}^p(W,\mu)$ norm to get

$$\left\| T_h \sum_{j=1}^n \alpha_j \mathcal{E}(f_j) - \sum_{j=1}^n \alpha_j \mathcal{E}(f_j) \right\|_p = \left\| T_{\theta h} D_h \sum_{j=1}^n \alpha_j \mathcal{E}(f_j) \right\|_p$$

$$
\begin{aligned}
&\leq C(h)\left\| D_h \sum_{j=1}^{n} \alpha_j \mathcal{E}(f_j)\right\|_q \\
&\leq C(h)|h| \left\| \left| D \sum_{j=1}^{n} \alpha_j \mathcal{E}(f_j)\right|_{L^2([0,T])}\right\|_q \\
&\leq C(h)|h| \left\| \sum_{j=1}^{n} \alpha_j \mathcal{E}(f_j)\right\|_{\mathbb{D}^{1,q}}.
\end{aligned}
$$

∎

We now continue the analysis of the term

$$
\left\| T_{-K_\varepsilon(t,\cdot)} V_t^\varepsilon - T_{-1_{[0,t]}(\cdot)} V_t^\varepsilon \right\|_q .
$$

It is not difficult to see that Assumption 1.2 implies that for any $\varepsilon > 0$ and $t \in [0,T]$ the random variable V_t^ε belongs to $\mathbb{D}^{1,q}$ for all $q \geq 1$. Moreover, the $\mathbb{D}^{1,q}$-norm of V_t^ε is bounded uniformly with respect to ε (observe that V_t, which corresponds to the case $\varepsilon = 0$, is related to an Itô type SDE which possesses the required smoothness). Therefore,

$$
\begin{aligned}
\left\| T_{-K_\varepsilon(t,\cdot)} V_t^\varepsilon - T_{-1_{[0,t]}(\cdot)} V_t^\varepsilon \right\|_q &= \left\| T_{-1_{[0,t]}(\cdot)} \left(T_{1_{[0,t]}(\cdot) - K_\varepsilon(t,\cdot)} V_t^\varepsilon - V_t^\varepsilon \right) \right\|_q \\
&\leq C \left\| T_{1_{[0,t]}(\cdot) - K_\varepsilon(t,\cdot)} V_t^\varepsilon - V_t^\varepsilon \right\|_r \\
&\leq C |K_\varepsilon(t,\cdot) - 1_{[0,t]}(\cdot)| \|V_t^\varepsilon\|_{\mathbb{D}^{1,r}}.
\end{aligned}
$$

for $r > q$. Combining all the estimates above we conclude

$$
\begin{aligned}
\|Y_t^\varepsilon - Y_t\|_p &\leq \|\mathcal{F}_1\|_p + \|\mathcal{F}_2\|_p + \|\mathcal{F}_3\|_p \\
&\leq C \left(\mathcal{S}_q \left(\sqrt{2} \sup_{s\in[0,T]} |K_\varepsilon(s,\cdot) - 1_{[0,s]}(\cdot)| \right) + |K_\varepsilon(t,\cdot) - 1_{[0,t]}(\cdot)| \right) \\
&\leq C \left(\mathcal{S}_q \left(\sqrt{2} \sup_{s\in[0,T]} |K_\varepsilon(s,\cdot) - 1_{[0,s]}(\cdot)| \right) + \sup_{t\in[0,T]} |K_\varepsilon(t,\cdot) - 1_{[0,t]}(\cdot)| \right) \\
&\leq C \mathcal{S}_q \left(\sqrt{2} \sup_{s\in[0,T]} |K_\varepsilon(s,\cdot) - 1_{[0,s]}(\cdot)| \right).
\end{aligned}
$$

The proof of Theorem 1.5 is now complete.

4 Proof of Theorem 1.7

We first note that the solution $\{A_t\}_{t\in[0,T]}$ of

$$
\frac{dA_t}{dt} = b(A_t) + A_t \cdot g(t) \quad A_0 = x,
$$

where $g : [0,T] \to \mathbb{R}$ is a continuous function, can be represented as

$$
A_t = G_t \cdot \exp\left\{ \int_0^t g(s) ds \right\} \tag{4.1}
$$

where $\{G_t\}_{t\in[0,T]}$ solves

$$\frac{dG_t}{dt} = b\left(G_t \cdot \exp\left\{\int_0^t g(s)ds\right\}\right) \cdot \exp\left\{-\int_0^t g(s)ds\right\}. \tag{4.2}$$

Moreover, recalling the argument from the previous section, we know that the solution $\{Y_t^\varepsilon\}_{t\in[0,T]}$ of (1.14) can be represented as

$$Y_t^\varepsilon = V_t^\varepsilon \diamond \mathcal{E}_\varepsilon(t)^{\diamond-1}$$

where

$$\frac{dV_t^\varepsilon}{dt} = b\left(V_t^\varepsilon \cdot (\mathcal{E}_\varepsilon(t))^{-1}\right) \cdot \mathcal{E}_\varepsilon(t).$$

Since by definition

$$\begin{aligned} (\mathcal{E}_\varepsilon(t))^{-1} &= \exp\left\{\int_0^T K_\varepsilon(t,s)dB_s + \frac{1}{2}|K_\varepsilon(t,\cdot)|^2\right\} \\ &= \exp\left\{B_t^\varepsilon + \frac{1}{2}|K_\varepsilon(t,\cdot)|^2\right\} \\ &= \exp\left\{\int_0^t \left(\frac{dB_s^\varepsilon}{ds} + \frac{1}{2}\frac{d|K_\varepsilon(s,\cdot)|^2}{ds}\right) ds\right\} \end{aligned}$$

a comparison with (4.1) and (4.2) shows that, by choosing

$$g(t) = \frac{1}{2}\frac{d}{dt}|K_\varepsilon(t,\cdot)|^2 + \frac{dB_t^\varepsilon}{dt},$$

we can write

$$V_t^\varepsilon = S_t^\varepsilon \cdot \mathcal{E}_\varepsilon(t)$$

where $\{S_t^\varepsilon\}_{t\in[0,T]}$ is the process defined in the statement of Theorem 1.7. Therefore,

$$\begin{aligned} Y_t^\varepsilon &= V_t^\varepsilon \diamond \mathcal{E}_\varepsilon(t)^{\diamond-1} \\ &= (S_t^\varepsilon \cdot \mathcal{E}_\varepsilon(t)) \diamond \mathcal{E}_\varepsilon(t)^{\diamond-1} \\ &= T_{-K_\varepsilon(t,\cdot)}\left(S_t^\varepsilon \cdot \mathcal{E}_\varepsilon(t)\right) \cdot \mathcal{E}_\varepsilon(t)^{\diamond-1} \\ &= T_{-K_\varepsilon(t,\cdot)}S_t^\varepsilon \cdot T_{-K_\varepsilon(t,\cdot)}\mathcal{E}_\varepsilon(t) \cdot \mathcal{E}_\varepsilon(t)^{\diamond-1} \\ &= T_{-K_\varepsilon(t,\cdot)}S_t^\varepsilon \cdot \exp\left\{-\int_0^T K_\varepsilon(t,s)dB_s - \frac{1}{2}|K_\varepsilon(t,\cdot)|^2 + |K_\varepsilon(t,\cdot)|^2\right\} \cdot \mathcal{E}_\varepsilon(t)^{\diamond-1} \\ &= T_{-K_\varepsilon(t,\cdot)}S_t^\varepsilon. \end{aligned}$$

Here, in the third equality, we utilized Gjessing Lemma. The proof of Theorem 1.7 is complete.

4.1 Alternative proof

We are now going to prove a technical result of independent interest that will be used to obtain a different and more direct proof of Theorem 1.7.

Proposition 4.1 *Let $\{X_t\}_{t\in[0,T]}$ be a stochastic process such that:*

- *the function $t \mapsto X_t$ is differentiable*
- *the random variable X_t belongs to $\mathcal{L}^p(W,\mu)$ for some $p>1$ and all $t \in [0,T]$.*

If the function $h : [0,T]^2 \to \mathbb{R}$ is such that

- *for almost all $s \in [0,T]$ the function $t \mapsto h(t,s)$ is continuously differentiable*
- *for all $t \in [0,T]$ the functions $h(t,\cdot)$ and $\partial_t h(t,\cdot)$ belong to $L^2([0,T])$*

then

$$\begin{aligned}\frac{d}{dt}(T_{h(t,\cdot)}X_t) &= T_{h(t,\cdot)}\frac{dX_t}{dt} + T_{h(t,\cdot)}X_t \cdot \int_0^T \partial_t h(t,s)dB_s \\ &\quad -T_{h(t,\cdot)}X_t \diamond \int_0^T \partial_t h(t,s)dB_s.\end{aligned}$$

Proof. To simplify the notation we set

$$\delta(h(t,\cdot)) := \int_0^T h(t,s)dB_s \quad \text{and} \quad \delta(\partial_t h(t,\cdot)) := \int_0^T \partial_t h(t,s)dB_s.$$

According to Gjessing Lemma we know that

$$T_{h(t,\cdot)}X_t \diamond \mathcal{E}(h(t,\cdot)) = X_t \cdot \mathcal{E}(h(t,\cdot))$$

or equivalently,

$$T_{h(t,\cdot)}X_t = (X_t \cdot \mathcal{E}(h(t,\cdot))) \diamond \mathcal{E}(-h(t,\cdot)). \tag{4.3}$$

We now use the chain rule for the Wick product to get

$$\begin{aligned}\frac{d}{dt}(T_{h(t,\cdot)}X_t) &= \frac{d}{dt}(X_t \cdot \mathcal{E}(h(t,\cdot))) \diamond \mathcal{E}(-h(t,\cdot)) + (X_t \cdot \mathcal{E}(h(t,\cdot))) \diamond \frac{d}{dt}\mathcal{E}(-h(t,\cdot)) \\ &= \left(\frac{dX_t}{dt}\cdot \mathcal{E}(h(t,\cdot))\right) \diamond \mathcal{E}(-h(t,\cdot)) + \left(X_t \cdot \frac{d}{dt}\mathcal{E}(h(t,\cdot))\right) \diamond \mathcal{E}(-h(t,\cdot)) \\ &\quad + (X_t \cdot \mathcal{E}(h(t,\cdot))) \diamond \frac{d}{dt}\mathcal{E}(-h(t,\cdot)) \\ &= \left(\frac{dX_t}{dt}\cdot \mathcal{E}(h(t,\cdot))\right) \diamond \mathcal{E}(-h(t,\cdot)) \\ &\quad + \left(X_t \cdot \mathcal{E}(h(t,\cdot)) \cdot \frac{d}{dt}\left(\delta(h(t,\cdot)) - \frac{1}{2}|h(t,\cdot)|^2\right)\right) \diamond \mathcal{E}(-h(t,\cdot)) \\ &\quad + (X_t \cdot \mathcal{E}(h(t,\cdot))) \diamond \mathcal{E}(-h(t,\cdot)) \diamond \frac{d}{dt}\delta(-h(t,\cdot))\end{aligned}$$

Observe that according to identity (4.3) we can write

$$\left(\frac{dX_t}{dt}\cdot \mathcal{E}(h(t,\cdot))\right) \diamond \mathcal{E}(-h(t,\cdot)) = T_{h(t,\cdot)}\frac{dX_t}{dt}.$$

Therefore, the last chain of equalities becomes

$$\begin{aligned}
\frac{d}{dt}(T_{h(t,\cdot)}X_t) &= T_{h(t,\cdot)}\frac{dX_t}{dt} \\
&\quad + (X_t \cdot \mathcal{E}(h(t,\cdot)) \cdot (\delta(\partial_t h(t,\cdot)) - \langle h(t,\cdot), \partial_t h(t,\cdot)\rangle)) \diamond \mathcal{E}(-h(t,\cdot)) \\
&\quad - T_{h(t,\cdot)}X_t \diamond \delta(\partial_t h(t,\cdot)) \\
&= T_{h(t,\cdot)}\frac{dX_t}{dt} + T_{h(t,\cdot)}\left(X_t \cdot (\delta(\partial_t h(t,\cdot)) - \langle h(t,\cdot), \partial_t h(t,\cdot)\rangle)\right) \\
&\quad - T_{h(t,\cdot)}X_t \diamond \delta(\partial_t h(t,\cdot)) \\
&= T_{h(t,\cdot)}\frac{dX_t}{dt} + T_{h(t,\cdot)}X_t \cdot T_{h(t,\cdot)}(\delta(\partial_t h(t,\cdot)) - \langle h(t,\cdot), \partial_t h(t,\cdot)\rangle) \\
&\quad - T_{h(t,\cdot)}X_t \diamond \delta(\partial_t h(t,\cdot)) \\
&= T_{h(t,\cdot)}\frac{dX_t}{dt} + T_{h(t,\cdot)}X_t \cdot \delta(\partial_t h(t,\cdot)) - T_{h(t,\cdot)}X_t \diamond \delta(\partial_t h(t,\cdot)).
\end{aligned}$$

The proof is complete. ∎

By means of Proposition 4.1, we are now able to prove identity (1.17) from Theorem 1.7 via a direct verification. More precisely, let $\{S_t^\varepsilon\}_{t\in[0,T]}$ be the process in the statement of Theorem 1.7. Then, using equation (1.16) we get

$$\begin{aligned}
\frac{d}{dt}T_{-K_\varepsilon(t,\cdot)}S_t^\varepsilon &= T_{-K_\varepsilon(t,\cdot)}\frac{dS_t^\varepsilon}{dt} - T_{-K_\varepsilon(t,\cdot)}S_t^\varepsilon \cdot \int_0^T \partial_t K_\varepsilon(t,s)dB_s \\
&\quad + T_{-K_\varepsilon(t,\cdot)}S_t^\varepsilon \diamond \int_0^T \partial_t K_\varepsilon(t,s)dB_s \\
&= T_{-K_\varepsilon(t,\cdot)}\left(b(S_t^\varepsilon) + \frac{1}{2}\frac{d}{dt}|K_\varepsilon(t,\cdot)|^2 \cdot S_t^\varepsilon + S_t^\varepsilon \cdot \frac{dB_t^\varepsilon}{dt}\right) \\
&\quad - T_{-K_\varepsilon(t,\cdot)}S_t^\varepsilon \cdot \frac{dB_t^\varepsilon}{dt} + T_{-K_\varepsilon(t,\cdot)}S_t^\varepsilon \diamond \frac{dB_t^\varepsilon}{dt} \\
&= b\left(T_{-K_\varepsilon(t,\cdot)}S_t^\varepsilon\right) + \frac{1}{2}\frac{d}{dt}|K_\varepsilon(t,\cdot)|^2 \cdot T_{-K_\varepsilon(t,\cdot)}S_t^\varepsilon \\
&\quad + T_{-K_\varepsilon(t,\cdot)}S_t^\varepsilon \cdot \left(\frac{dB_t^\varepsilon}{dt} - \int_0^T \partial_t K_\varepsilon(t,s)K_\varepsilon(t,s)ds\right) \\
&\quad - T_{-K_\varepsilon(t,\cdot)}S_t^\varepsilon \cdot \frac{dB_t^\varepsilon}{dt} + T_{-K_\varepsilon(t,\cdot)}S_t^\varepsilon \diamond \frac{dB_t^\varepsilon}{dt} \\
&= b\left(T_{-K_\varepsilon(t,\cdot)}S_t^\varepsilon\right) + T_{-K_\varepsilon(t,\cdot)}S_t^\varepsilon \diamond \frac{dB_t^\varepsilon}{dt}.
\end{aligned}$$

This is imply that $\{T_{-K_\varepsilon(t,\cdot)}S_t^\varepsilon\}_{t\in[0,T]}$ solves equation (1.14).

References

[1] V. I. Bogachev, *Gaussian Measures*, American Mathematical Society, Providence, 1998.

[2] Z. Brezniak and F. Flandoli, Almost sure approximation of Wong-Zakai type for stochastic partial differential equations, *Stoch. Proc. and their Appl.* **55** (1995) 329-358.

[3] P. Da Pelo and A. Lanconelli, On a new probabilistic representation for the solution of the heat equation, *Stochastics* **84** (2012) 171-181.

[4] P. Da Pelo, A. Lanconelli and A. I. Stan, A Hölder-Young-Lieb inequality for norms of Gaussian Wick products, *Inf. Dim. Anal. Quantum Prob. Related Topics* **14** (2011) 375-407.

[5] P. Da Pelo, A. Lanconelli and A. I. Stan, An Itô formula for a family of stochastic integrals and related Wong-Zakai theorems, *Stoch. Proc. and their Appl.* **123** (2013) 3183-3200.

[6] I. Gyöngy and G. Michaletzky, On Wong-Zakai Approximations with δ-Martingales, *Proceedings: Mathematical, Physical and Engineering Sciences* **460**(2004) 309-324.

[7] I. Gyöngy and A. Shmatkov, Rate of Convergence of WongZakai Approximations for Stochastic Partial Differential Equations, *Appl. Math. Optim.* **54** (2006) 315-341.

[8] I. Gyöngy and P. R. Stinga, Rate of Convergence of WongZakai approximations for Stochastic Partial Differential Equations, *Seminar on Stochastic Analysis, Random Fields and Applications VII* **67** (2013) Birkhäuser, Basel.

[9] M. Hairer and E. Pardoux, A WongZakai theorem for stochastic PDEs, *J. Math. Soc. Japan.* **67** (2015) 1551-1604.

[10] H. Holden, B. Øksendal, J. Ubøe and T.-S. Zhang, *Stochastic Partial Differential Equations - II Edition*, Springer, New York, 2010.

[11] Y. Hu, *Analysis on Gaussian spaces*, World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ, 2017.

[12] Y. Hu, G. Kallianpur and J. Xiong, An approximation for Zakai equation, *Applied Math. Optimiz.* **45** (2002) 2344.

[13] Y. Hu and D. Nualart, Rough path analysis via fractional calculus, *Trans. Amer. Math. Soc.* **361** (2009) 2689-2718.

[14] Y. Hu and B. Øksendal, Wick approximation of quasilinear stochastic differential equations, *Stochastic analysis and related topics* **V** Birkhäuser (1996) 203-231.

[15] S. Janson, *Gaussian Hilbert spaces*, Cambridge Tracts in Mathematics, 129. Cambridge University Press, Cambridge, 1997.

[16] I. Karatzas and S. E. Shreve, *Brownian motion and stochastic calculus*, Springer-Verlag, New York, 1991.

[17] F. Konecny, A Wong-Zakai approximation of stochastic differential equations, *Journal of Multivariate Analysis* **13** (1983) 605-611.

[18] A. Lanconelli and A. I. Stan, Some norm inequalities for Gaussian Wick Products, *Stochastic Analysis and Applications* **28** (2010) 523-539.

[19] J. A. Londono and A. M. Villegas, Numerical performance of some Wong-Zakai type approximations for stochastic differential equations, *Int. J. Pure and Applied Math.* **107** (2016) 301-315.

[20] N.Naganuma, Exact convergence rate of the WongZakai approximation to RDEs driven by Gaussian rough paths, *Stochastics* **7** (2016) 1041-1059.

[21] D. Nualart, *Malliavin calculus and Related Topics - II Edition*, Springer, New York, 2006.

[22] D. W. Stroock and S. R. S. Varadhan, On the support of diffusion processes with applications to the strong maximum principle, *Proceedings* 6*-th Berkeley Symposium Math. Statist. Probab.* **3** (1972) University of California Press, Berkeley, 333-359.

[23] G. Tessitore and G. J. Zabczyk, Wong-Zakai approximations of stochastic evolution equations, *Journal evol. equ.* **6** (2006) 621-655.

[24] E. Wong and M. Zakai: On the relation between ordinary and stochastic differential equations, *Intern. J. Engr. Sci.* **3** (1965) 213-229.

[25] E. Wong and M. Zakai, Riemann-Stieltjes approximations of stochastic integrals, *Z. Wahrscheinlichkeitstheorie verw. Geb.* **12** (1969) 87-97.